RISK MANAGEMENT – CURRENT ISSUES AND CHALLENGES

Edited by **Nerija Banaitiene**

INTECHOPEN.COM

Risk Management – Current Issues and Challenges
http://dx.doi.org/10.5772/2568
Edited by Nerija Banaitiene

Contributors
Werner Hürlimann, Rossella Agliardi, Petar Popivanov, Angela Slavova, A. Gonen, Sabrina Grimaldi, Carlo Rafele, Anna Corinna Cagliano, Pako Thupayagale, Marius Ioan Podean, Dan Benta, Robert Elliott Allinson, Daniel Ekwall, Omar Gaci, Hervé Mathieu, Jean-Pierre Deutsch, Laurent Gomez, David R. Comings, Wendy W. Ting, Gabriela Dobrotă, Emilia Vasile, Ion Croitoru, Óscar Lázaro, Agustín Moyano, Mikel Uriarte, Alicia González, Teresa Meneu, Juan Carlos Fernández-Llatas, Vicente Traver, Benjamín Molina, Carlos Palau, Óscar López, Etxahun Sánchez, Saioa Ros, Antonio Moreno, María González, José Antonio Palazón, Miguel Sepulcre, Javier Gozálvez, Luis Collantes, Gonzalo Prieto, Amparo Marin de la Barcena, Marin Andreica, Stere Farmache, Madalina Ecaterina Andreica, Octavian Stroie, Gurudeo Anand Tularam, Gowri Sameera Attili, Jordi Botet, Pedro Maria-Sanchez, Nerija Banaitiene, Audrius Banaitis, Satit Aditto, Christopher Gan, Gilbert V. Nartea, R. Martínez, D. Hemming, L. Malone, N. Bermudez, G. Cockfield, A. Diongue, J. Hansen, A. Hildebrand, K. Ingram, G. Jakeman, M. Kadi, G. R. McGregor, S. Mushtaq, P. Rao, R. Pulwarty, O. Ndiaye, G. Srinivasan, Eh. Seck, N. White, R. Zougmoré, Á.G. Muñoz, D. Ruiz, P. Ramírez, G. León, J. Quintana, A. Bonilla, W. Torres, M. Pastén, O. Sánchez, Andrew Charles, Yuriy Kuleshov, David Jones

Published by InTech
Janeza Trdine 9, 51000 Rijeka, Croatia

Publishing Process Manager Mirna Cvijic
Typesetting InTech Prepress, Novi Sad
Cover InTech Design Team

First published September, 2012
Printed in Croatia

A free online edition of this book is available at www.intechopen.com
Additional hard copies can be obtained from orders@intechopen.com

Risk Management – Current Issues and Challenges, Edited by Nerija Banaitiene
p. cm.
ISBN 978-953-51-0747-7

Contents

Preface

Companies face risks every day, they are part of normal business life. There are many risks — both threats and opportunities — which may impact on a company's resources, projects and profitability. Risk means different things to different businesses and organizations. Undoubtedly, the risk represents both a potential threat and potential opportunity for businesses.

Every business and decision involves a certain amount of risk. Risk might cause a loss to a company. This does not mean, however, that businesses cannot take risks. As disengagement and risk aversion may result in missed business opportunities, which will lead to slower growth and reduced prosperity of a company. In today's increasingly complex and diverse environment, it is crucial to find the right balance between **risk aversion and risk taking**. To do this it is essential to understand the complex, out of the whole range of economic, technical, operational, environmental and social risks associated with the company's activities. However, risk management is about much more than merely avoiding or successfully deriving benefit from opportunities. Risk management is the identification, assessment, and prioritization of risks. Lastly, risk management helps a company to handle the risks associated with a rapidly changing business environment. When risk management does receive attention, it is often in response to unforeseen (and usually negative) events.

The impact of the global economic crisis has varied from one country to another: not all countries, sectors and organizations were affected in the same hard way by it. Even, the impact of the financial crisis is varying widely across companies within the same sector. In today's post-crisis economy effective risk management is a critical component of any successful management strategy. In complex and rapidly changing situations, as today's supply chains and partnership arrangements tend to be, management needs to consider all risks within the enlarged business connections.

Understanding of the risk management is vital for both practitioners and researchers. The emergence of new insights into approaches and models can help address multifaceted risk management issues.

There are five parts in this book of 23 chapters. The papers are organized according to theoretical, methodological and practical issues and areas of risk management: Part 1 provides new insights into theoretical approaches and models for risk management, Part 2 deals with risk and supply chain management, Part 3 focuses on specific aspects of enterprise risk management, Part 4 examines risk management practice across different projects and industries, and Part 5 discusses emerging issues related to climate change and climate risk management. The authors touched on a wide range of risk management issues. Consequently, in the context of the thematic classification scheme, some papers fall into more than one category.

I consider it an honour and privilege to have had the opportunity to edit this book. I am particularly grateful to all the authors for their outstanding contributions, and to Mirna Cvijic, the publishing process manager at InTech, for her kind assistance in publishing this book.

Dr. Nerija Banaitiene
Department of Construction Economics and Property Management,
Faculty of Civil Engineering,
Vilnius Gediminas Technical University,
Lithuania

Section 1

Approaches and Models for Risk Management

Chapter 1

Biometric Solvency Risk for Portfolios of General Life Contracts (II) The Markov Chain Approach

Werner Hürlimann

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/48374

1. Introduction

The main theoretical goal of the present exposé is to extend the results presented in Hürlimann [1] to the Markov chain model of life insurance, which enables modeling all single life/multiple life traditional contracts subject to biometric risk with multiple causes of decrement. In particular, a complete risk modeling of single-life insurance products with mortality and disability risks requires the specification of a Markov model with three states. As novel illustration we offer to the interested practitioner an in-depth treatment of endowment contracts with waiver of premium by disability.

The present investigation is restricted to biometric risks encountered in traditional insurance contracts within a discrete time Markov chain model. The current standard requirements for the Solvency II life risk module have been specified in QIS5 [2], pp.147-163. QIS5 prescribes a solvency capital requirement (SCR), which only depends on the time of valuation (=time at which solvency is ascertained) but not on the portfolio size (=number of policies). It accounts explicitly for the uncertainty in both trends (=systematic risk) and parameters (=parameter risk) but not for the random fluctuations around frequency and severity of claims (=process risk). In fact, the process risk has been disregarded as not significant enough, and, in order to simplify the standard formula, it has been included in the systematic/parameter risk component. For the purpose of internal models and improved risk management, it appears important to capture separately or simultaneously all risk components of biometric risks. A more detailed account of our contribution follows.

As starting point, we recall in Section 2 the general solvency rule for the prospective liability risk derived in [1], Section 2, which has resulted in two simple liability VaR & CVaR target capital requirements. In both stochastic models, the target capital can be decomposed into a solvency capital component (liability risk of the current period) and a risk margin component (liability risk of future periods), where the latter must be included (besides the

INTECH
open science | open minds

best estimate liabilities) in the technical provisions. This general decomposition is in agreement with the current QIS5 specification. The proposed approach is then applied to determine the biometric solvency risk capital for a portfolio of general traditional life contracts within the Markov chain model of life insurance. For this, we assume that the best estimate liabilities of a general life contract coincide with the so-called "net premium reserves". After introduction of the Markov chain approach to life insurance in Section 3, we recall in Section 4 the ubiquitous backward recursive actuarial reserve formula and the theorem of Hattendorff. Based on this we determine in Section 5 the conditional mean and variance of a portfolio's prospective liability risk (=random present value of future cash-flows at a given time of valuation) and use a gamma distribution approximation to obtain the liability VaR & CVaR solvency capital as well as corresponding solvency capital ratios. These first formulas include only the process risk and not the systematic risk. To include the latter risk in solvency investigations we propose either to shift the biometric transition probabilities, as done in Section 6.2, or apply a stochastic model, which allows for random biometric transition probabilities, as explained in Section 6.3. Section 7 illustrates numerically and graphically the considered VaR & CVaR solvency capital models for a cohort of endowment contracts with waiver of premium by disability and compares them with the current Solvency II standard approach. Finally, Section 8 summarizes, concludes and provides an outlook for possible alternatives and extensions.

2. A general prospective approach to the liability risk solvency capital

Starting point is a multi-period discrete time stochastic model of insurance. Given is a time horizon T and a probability space (Ω, F, P) endowed with a filtration $(F_t)_{t=0}^{T}$ such that $F_0 = \{\Omega, \varnothing\}$ and $F_T = F$. Let $L^{\infty}(F_t)$ be the space of essentially bounded random variables on (Ω, F, P) and B^{∞} the space of essentially bounded stochastic processes on (Ω, F, P) which are adapted to the filtration $(F_t)_{t=0}^{T}$. The basic discrete time stochastic processes are

A_t, L_t : the *assets* and *actuarial liabilities* at time t

In a total balance sheet approach, their values depend upon the stochastic processes in B^{∞}, which describe the random cash-in and cash-out flows of any type of insurance business:

P_{t-1} : *loaded premiums* to be paid at time $t-1$ (assumed invested at time $t-1$)

X_t : *insurance costs* to be paid at time t (includes insurance benefits, expenses and bonus payments paid during the time period $(t-1,t]$)

R_t : *accumulation factor* for return on investment for the time period $(t-1,t]$

We assume that X_t is F_t-measurable and R_t is F_{t-1}-measurable. The random *cumulated accumulation factor* for return over the period $(s,t], 0 \le s < t \le T,$ is denoted by $R_{s,t} = \prod_{j=s+1}^{t} R_j$.

Since R_t is F_{t-1}-measurable $R_{s,t}$ is F_{t-1}-measurable, and therefore $\{R_{s,t}, t > s\}$ is a predictable stochastic process. The quantity $D_{s,t} = R_{s,t}^{-1}$ is called random discount rate. Consider the F_{t+j+1}-measurable discrete time stochastic process $\left\{CF_{t+j}, j = 0,1,...,T-t-1\right\}$ of *future insurance cash-flows* defined by

$$CF_{t+j} = D_{t+j,t+j+1} \cdot X_{t+j+1} - P_{t+j} \quad j = 0,1,...,T-t-1. \tag{1}$$

The actuarial liabilities at time t, also called time-t *prospective insurance liability*, coincide with the random present value of all future insurance cash-flows at time t given by

$$L_t = \sum_{j=0}^{T-t-1} D_{t,t+j} CF_{t+j}, \quad t = 0,1,...,T-1. \tag{2}$$

Using (1)-(2) and the relationship $D_{t,t+1+k} = D_{t,t+1} \cdot D_{t+1,t+1+k}$, one obtains the recursive equation $L_{t+1} = \left(L_t + P_t\right) \cdot R_{t+1} - X_{t+1}$, $t = 1,...,T-1$. On the other hand, the random assets over the time horizon $\left[0,T\right]$ satisfy by definition the recursive equation $A_{t+1} = \left(A_t + P_t\right) \cdot R_{t+1} - X_{t+1}$, $t = 1,...,T-1$. Through subtraction it follows that

$$A_{t+1} - L_{t+1} = (A_t - L_t) \cdot R_{t+1}, \quad t = 1,...,T-1, \tag{3}$$

which implies the following equivalent probabilistic conditions (use that trivially $L_T = 0$)

$$P\left(A_T \geq 0 \middle| F_t\right) \geq 1-\varepsilon, \tag{4}$$

$$P\left(A_{t+\tau} \geq L_{t+\tau}, \tau = 1,2,...,T-t-1 \middle| F_t\right) \geq 1-\varepsilon, \tag{5}$$

$$P\left(A_t \geq L_t \middle| F_t\right) \geq 1-\varepsilon. \tag{6}$$

Given a default probability $\varepsilon > 0$, the *liability VaR solvency criterion* (6) says that at time t the initial (deterministic) capital requirement A_t should exceed the random present value of future cash-flows with a probability of at least $1-\varepsilon$. By (4)-(5) this criterion automatically implies that assets will exceed liabilities with the same probability at each future time over the time horizon T. Let $A_t^{VaR} = VaR_{1-\varepsilon}\left[L_t \middle| F_t\right]$ be a minimum solution to (6), and assume that the best estimate insurance liabilities at time t coincide with the *net premium reserves* (in the sense defined later in (35), that is let $E\left[L_t \middle| F_t\right] = V_t^Z$. Then, the *liability VaR solvency capital* $SC_t^{VaR} = A_t^{VaR} - V_t^Z = VaR_{1-\varepsilon}\left[L_t \middle| F_t\right] - V_t^Z$ represents the capital available at time t to meet the insurance risk liabilities with high probability. A *risk margin* is added to this capital requirement (recall that in Solvency II the sum of the best estimate insurance liabilities and the risk margin determines the Technical Provisions). The *liability VaR target capital* is the sum of the liability VaR solvency capital and the risk margin defined by

$$TC_t^{VaR} = SC_t^{VaR} + RM_t^{VaR}. \tag{7}$$

The cost-of-capital risk margin with cost-of-capital rate $i_{CoC} = 6\%$ is defined by

$$RM_t^{VaR} = i_{CoC} \cdot \sum_{k=1}^{T-t} v_f^k \cdot SC_{t+k}^{VaR}, \tag{8}$$

where T denotes the time horizon, and v_f is the risk-free discount rate. For comparison with other solvency rules, one considers the *VaR solvency capital ratio* at time t defined by

$$SR_t^{VaR} = SC_t^{VaR} / V_t^Z. \tag{9}$$

Alternatively, let $CVaR_{1-\varepsilon}\left[L_t \middle| F_t\right] = E\left[L_t \middle| L_t > VaR_{1-\varepsilon}\left[L_t \middle| F_t\right], F_t\right]$ be the conditional value-at-risk of the random present value of future cash-flows at the confidence level $1-\varepsilon$ given the information available at time t. The *liability CVaR target capital* $TC_t^{CVaR} = CVaR_{1-\varepsilon}\left[L_t \middle| F_t\right] - V_t^Z + RM_t^{CVaR} = SC_t^{CVaR} + RM_t^{CVaR}$ also meets the insurance risk liabilities and it defines the *CVaR solvency capital ratio* at time t:

$$SR_t^{CVaR} = SC_t^{CVaR} / V_t^Z. \tag{10}$$

3. The Markov chain approach to general life contracts

Consider the Markov chain model of a *general life insurance* (GLIFE) contract with state space S and arbitrary payments. The *state space* S is the finite set of states a contract can be during its lifetime. Payments are induced by two kinds of events:

Type 1: payments induced by being in a certain state
Type 2: payments induced by a jump of state

The *payment function vector* of a contract at time $k = 0,1,2,...$ is expressed as a vector $a(k) = \left\{a_i(k), a_{ij}(k) \middle| i \neq j \in S\right\}$, where the *payment functions* are defined by

Type 1: $a_i(k), i \in S$, is the payment if the contract at time k is in state i.

Type 2: $a_{ij}(k), i \neq j \in S, k \geq 1$, is the payment if the contract was in state i at time $k-1$ and is in state state j at time k. For convenience set $a_{ij}(0) = 0$ for $i \neq j \in S$ and $a_{ij}(k) = 0$ for $i = j \in S$.

For better interpretation one splits the payment $a_i(k)$ into a *benefit part* and a *premium part* such that $a_i(k) = b_i(k) - \pi_i(k), \quad i \in S, k = 0,1,2,...$, where $\pi_i(k) \geq 0$ denotes the non-negative premium paid at time k when the contract is in state i. Note that in most applications one has $\pi_i(k) = 0$ if the state i is different from the state of being "active" (premiums are only paid in this situation). Restricting the attention to biometric risk only, we assume throughout a flat term structure of interest rates with annual interest rate i and discount factor $v = 1/(1+i)$. The state of a GLIFE contract over time is described by the discrete time stochastic process $(X_k)_{k=0,1,2,...}$ with values in S. The event $\{X_k = s\}$ means that the contract at time k is in state s. We assume that $(X_k)_{k=0,1,2,...}$ is a *Markov chain*, which implies that the joint distributions of the random states can be represented in terms of the *one-step transition probability matrix* $p(k) = (p_{ij}(k))_{i,j\in S}$, which is defined by

$$p_{ij}(k) = P(X_{k+1} = j | X_k = i), \quad k = 0,1,2,... \tag{11}$$

The set $(S, a(k), X_k, k = 0,1,2,...)$ defines the Markov chain model widely discussed in life insurance (Amsler [3]; Hoem [4], [5]; Koller [6]; Milbrodt & Helbig [7]; Wolthuis [8]; etc.). Now, using the indicator function $I(\cdot)$, consider the *random cash-flow* of the GLIFE contract in year $(k, k+1]$ valued at time $k = 0,1,2,...$, which is defined by

$$C_k = \sum_{i\in S}\{b_i(k) - \pi_i(k)\} \cdot I(X_k = i) + v \cdot \sum_{i\neq j\in S} a_{ij}(k+1) \cdot I(X_k = i \wedge X_{k+1} = j). \tag{12}$$

The *insurance loss* random variable of a GLIFE contract is defined by

$$L = \sum_{k=0}^{\infty} v^k C_k. \tag{13}$$

This identifies the insurance loss with the random present value of all future cash-flows. Furthermore, for an arbitrary non-negative integer $\tau = 0,1,...$, one defines the *time-τ prospective loss* random variable

$$L_\tau = (1+i)^\tau \cdot \sum_{k=\tau}^{\infty} v^k C_k, \tag{14}$$

whose (conditional) expected value defines the time-τ actuarial reserve

$$V_\tau = E[L_\tau | X_\tau] = \sum_{k\in S} V_\tau^k, \quad V_\tau^k = E[L_\tau | X_\tau = k]. \tag{15}$$

The quantity V_τ^k is called *state-k time-τ actuarial reserve*. In particular, one has $L_0 = L$ and $V_0 = E[L]$ is the *initial actuarial reserve*, which is not assumed to vanish.

4. Backward recursive reserve formula and the theorem of Hattendorff

In a first step, we derive a recursion formula for the actuarial reserves. Recall the recursion formula for the random prospective loss

$$L_\tau = C_\tau + vL_{\tau+1}. \tag{16}$$

Assume that the contract is in state . $k \in S$ at time τ. Inserting (16) into (15) yields

$$V_\tau^k = E\left[C_\tau \middle| X_\tau = k\right] + v \cdot E\left[L_{\tau+1} \middle| X_\tau = k\right]. \tag{17}$$

Using (12) the first expectation in (17) can be rewritten as

$$v \cdot \sum_{i \neq j \in S} a_{ij}(\tau+1) \cdot P\left(X_\tau = i \wedge X_{\tau+1} = j \middle| X_\tau = k\right) + \sum_{i \in S} \left\{b_i(\tau) - \pi_i(\tau)\right\} \cdot P\left(X_\tau = i \middle| X_\tau = k\right)$$
$$= v \cdot \sum_{j \neq k \in S} a_{kj}(\tau+1) \cdot p_{kj}(\tau) + b_k(\tau) - \pi_k(\tau).$$

The second expectation equals

$$E\left[L_{\tau+1} \middle| X_\tau = k\right] = \sum_{j \in S} E\left[L_{\tau+1} \middle| X_{\tau+1} = j\right] \cdot P\left(X_{\tau+1} = j \middle| X_\tau = k\right) = \sum_{j \in S} V_{\tau+1}^j \cdot p_{kj}(\tau).$$

Inserting both expressions into (17) and using the made convention $a_{kk}(\tau+1) = 0$ as well as the relationship $\sum_{j \in S} p_{kj}(\tau) = 1$, one obtains the backward recursive reserve formula

$$V_\tau^k = \sum_{j \in S} p_{kj}(\tau) \cdot \left\{v \cdot V_{\tau+1}^j + v \cdot a_{kj}(\tau+1) + b_k(\tau)\right\} - \pi_k(\tau). \tag{18}$$

The actuarial reserve at time τ given the contract is in state $k \in S$ equals the one-year discounted sum over all possible states of the

- actuarial reserves at time $\tau + 1$,
- payments at time $\tau + 1$ due to a jump in states,
- payments at time τ due if being in a certain state,

which is weighted by the one-step transition probabilities and reduced by the premium paid at time τ when the contract is in state k. The representation (18) is a discrete version of *Thiele's differential equation*. Thiele's differential equation is a simple example of a Kolmogorov backward equation, which is a basic tool for determining conditional expected values in intensity-driven Markov processes, e.g. Norberg [9].

Let us rearrange (18) in order to obtain the Markov chain analogue of the classical decomposition of the premium into risk premium and saving premium (Gerber [10], [11]), Section 7.5, equation (5.3), and [1], equation (19).

Theorem 4.1 The premium $\pi_k(\tau)$ at time τ if the contract is in state $k \in S$ is the sum of a *saving premium* $\pi_k^S(\tau)$ and a *risk premium* $\pi_k^R(\tau)$, which are defined as follows:

$$\pi_k^S(\tau) = v \cdot V_{\tau+1}^k - V_\tau^k \tag{19}$$

$$\pi_k^R(\tau) = b_k(\tau) - v \cdot V_{\tau+1}^k + v \cdot \sum_{j \in S} p_{kj}(\tau) \cdot \left\{a_{kj}(\tau+1) + V_{\tau+1}^j\right\} \tag{20}$$

Proof. Making use of the recursion (18) and the relationship $\sum_{j\in S} p_{kj}(\tau)=1$, one obtains

$$\pi_k^S(\tau)+\pi_k^R(\tau)=v\cdot V_{\tau+1}^k-V_\tau^k+b_k(\tau)-v\cdot V_{\tau+1}^k+v\cdot\sum_{j\in S}p_{kj}(\tau)\cdot\left\{a_{kj}(\tau+1)+V_{\tau+1}^j\right\}$$
$$=-V_\tau^k+\sum_{j\in S}p_{kj}(\tau)\cdot\left\{v\cdot a_{kj}(\tau+1)+v\cdot V_{\tau+1}^j+b_k(\tau)\right\}=\pi_k(\tau),$$

which shows the desired decomposition. ◊

The saving premium represents the expected change in actuarial reserve at time τ for a contract in state k while the risk premium is the expected value at time τ of a contract in state k needed to cover the insurance risk in time period $(\tau,\tau+1]$. Rewrite the latter as

$$\pi_k^R(\tau)=b_k(\tau)+v\cdot\sum_{j\in S}p_{kj}(\tau)\cdot\left\{a_{kj}(\tau+1)+V_{\tau+1}^j-V_{\tau+1}^k\right\}. \tag{21}$$

This is the sum of the benefit payment at time τ for a contract in state k and the probability weighted sum of the *sums at risk* $\left\{a_{kj}(\tau+1)+V_{\tau+1}^j-V_{\tau+1}^k\right\}$ due to transitions from state k to state j at time $\tau+1$. The sum at risk is the amount credited to the insured's contract upon a transition, namely the lump sum payable immediately plus the adjustment of the actuarial reserve. The obtained results constitute a discrete time version of those mentioned in Norberg [12], p.10.

To evaluate the mean and variance of the random insurance loss (13) of a GLIFE contract, we follow the martingale approach to the *Theorem of Hattendorff* (Bühlmann [13]; Gerber [14]; Gerber et al. [15]; Hattendorff [16]; Kremer [17]; Patatriandafylou & Waters [18]; etc.). For this consider the set of $\tau+1$ contract states $S_\tau=\left\{X_\eta,\eta=0,1,...,\tau\right\}$ at time τ and the sequence of random variables

$$Y_\tau=E\left[L\middle|S_\tau\right],\quad \tau=1,2,...,\quad Y_0=E\left[L\right]=V_0. \tag{22}$$

The discrete time stochastic process $\left\{Y_\tau\right\}$ is a martingale with respect to $\left\{S_\tau\right\}$. The martingale differences $Y_{\tau+1}-Y_\tau=v^\tau\Lambda_\tau$, $\tau=0,1,2,...$, represent the *discounted one-year insurance losses* and form a sequence of uncorrelated random variables such that

$$E\left[\Lambda_\eta\right]=0,\quad Cov\left[\Lambda_\eta,\Lambda_\tau\right]=0,\quad 0\le\eta<\tau,\quad L-V_0=\sum_{\tau=0}^{\infty}v^\tau\Lambda_\tau. \tag{23}$$

Through detailed calculation one obtains the following result.

Theorem 4.2 The variance of the random insurance loss of a GLIFE contract is determined by the following formulas

$$Var\left[L\right]=\sum_{\tau=0}^{\infty}v^{2\tau}\cdot Var\left[C_\tau\right], \tag{24}$$

$$Var\left[C_\tau\right]=\sum_{k\in S}E\left[C_\tau^2\middle|X_\tau=k\right]\cdot P\left(X_\tau=k\right)-\pi^S(\tau)^2,\quad \pi^S(\tau)=\sum_{k\in S}\pi_k^S(\tau), \tag{25}$$

$$\begin{aligned}E\left[C_\tau^2\middle|X_\tau=k\right]&=\left[b_k(\tau)-\pi_k(\tau)\right]^2\\&+\sum_{j\in S}va_{kj}(\tau+1)\cdot\left[b_k(\tau)-\pi_k(\tau)+va_{kj}(\tau+1)\right]\cdot p_{kj}(\tau).\end{aligned} \tag{26}$$

Proof. Similarly to Gerber et al. [15], formula (89), one has

$$Y_\tau=\sum_{j=0}^{\tau-1}v^jC_j+v^\tau E\left[L_\tau\middle|S_\tau\right]. \tag{27}$$

Using (15) one obtains

$$\begin{aligned}v^\tau\Lambda_\tau&=Y_{\tau+1}-Y_\tau=v^\tau\left(C_\tau+vE\left[L_{\tau+1}\middle|S_{\tau+1}\right]-E\left[L_\tau\middle|S_\tau\right]\right)\\&=v^\tau\left(C_\tau+\sum_{k\in S}\left\{vE\left[L_{\tau+1}\middle|X_{\tau+1}=k\right]-E\left[L_\tau\middle|X_\tau=k\right]\right\}\right)=v^\tau\left(C_\tau+\pi^S(\tau)\right),\end{aligned} \tag{28}$$

$$\Lambda_\tau=C_\tau+\pi^S(\tau). \tag{29}$$

Since $E\left[\Lambda_\tau\right]=0$ one gets $E\left[C_\tau\right]=-\pi^S(\tau)$ and further

$$Var\left[\Lambda_\tau\right]=Var\left[C_\tau\right]=E\left[C_\tau^2\right]-E\left[C_\tau\right]^2=\sum_{k\in S}E\left[C_\tau^2\middle|X_\tau=k\right]\cdot P\left(X_\tau=k\right)-\pi^S(\tau)^2, \tag{30}$$

which is (25). To obtain (26) one uses (12) and the convention $a_{ii}(\tau+1)=0$ to get

$$\begin{aligned}C_\tau^2&=\sum_{i\in S}\left[b_i(\tau)-\pi_i(\tau)\right]^2\cdot I\left(X_\tau=i\right)\\&+\sum_{i,j\in S}va_{ij}(\tau+1)\cdot\left[b_i(\tau)-\pi_i(\tau)+va_{ij}(\tau+1)\right]\cdot I\left(X_\tau=i\wedge X_{\tau+1}=j\right),\end{aligned} \tag{31}$$

which implies that

$$E\left[C_\tau^2\middle|X_\tau=k\right]=\left[b_k(\tau)-\pi_k(\tau)\right]^2+\sum_{j\in S}va_{kj}(\tau+1)\cdot\left[b_k(\tau)-\pi_k(\tau)+va_{kj}(\tau+1)\right]\cdot p_{kj}(\tau). \tag{32}$$

Remark 4.1 In the single life case, the variance formulas in Theorem 4.2 should be compared with the ones for the GLIFE contract with one and multiple causes of decrement in [1], formulas (24)-(26). One can ask if the formula (25) is equivalent to the following one (at least in the single life case)

$$Var\left[C_{\tau}\right]=\sum_{k\in S}\left[\sum_{j\in S}\left(a_{kj}(\tau+1)+V_{\tau+1}^{j}-V_{\tau+1}^{k}\right)^{2}v^{2}p_{kj}(\tau)-\pi_{k}^{R}(\tau)^{2}\right]\cdot P\left(X_{\tau}=k\right). \tag{33}$$

5. The liability VaR & CVaR solvency capital for portfolios of GLIFE contracts

We begin with risk calculations for a single GLIFE contract, and use them to determine the liability VaR & CVaR solvency capital for a portfolio of GLIFE contracts.

5.1. Risk calculations for a single GLIFE contract

Given is a single GLIFE contract with random future cash-flows $\{C_k\}$ defined by (12). We assume that the state space contains a unique distinguished "void" state $\{X_k=\phi\}$ meaning that the contract has terminated at time k. We assume *contract survival*, i.e. a contract is still alive at time of valuation t, which implies that the conditional event $E_t=\{X_t\neq\phi\}$ is fulfilled. We note that the random present value of future cash-flows at time t defined by

$$Z_t=\sum_{j=0}^{\infty}v^{j}C_{t+j},\quad t=0,1,..., \tag{34}$$

coincides with the time-t prospective loss defined in (14), that is $Z_t=L_t,\ t=0,1,...$. Therefore, the expected value given contract survival equals

$$V_t^{Z}=E\left[Z_t\left|E_t\right.\right]=\sum_{k\in S}V_t^{k}\cdot P\left(X_t=k\left|E_t\right.\right)=\sum_{k\in S}V_t^{k}\cdot\frac{P\left(X_t=k\right)}{P\left(E_t\right)}. \tag{35}$$

In contrast to (15) the reserve defined in (35) is state independent and called *net premium reserve*, see Bowers et al. [19], Chap.17.7, p. 500, for a special case. Following Section 2, this value can been chosen as best estimate of the contract liabilities.

Remarks 5.1 (i) The motivation for state-independent reserves is second-to-die life insurance, where during lifetime the insurer may not be informed about the first death. An endowment with waiver of premium during disability, which is our illustration in Section 7, seems to contradict this concept because it cannot be argued that the insurer is unaware of the state occupied while the premium is being waived. However, at a given arbitrary time of valuation (including starting dates of contracts) future states of contracts are unknown, and therefore it is reasonable in a first step to assume state independent reserves for the design of a general method. Later refinement might be necessary to cover all possible cases.

(ii) State independent reserves have been introduced by Frasier [20] for the last-survivor status, see also The Actuary [21] and Margus [22]. The choice between state independent and state dependent reserves depends upon loss recognition in the balance sheet (recognition or not of a status change). With state independent reserves, the insurance company administers the contract as if it had no knowledge of any decrements, as long as the contract is not terminated. Only the latter situation is considered in the present work.

In a first step, we determine the mean and variance of the conditional distribution of Z_t given E_t. Similarly to [1], Section 5.1, the variance formulas (24)-(26) generalize to an arbitrary discrete time $t = 1,2,...$. Formula (23) generalizes as

$$L_t - V_t \cdot I(E_t) = \sum_{k=t}^{\infty} v^{k-t} \Lambda_k . \tag{36}$$

Noting further that $Z_t = L_t, t = 0,1,...$, one obtains from (36) the following conditional variance formulas (conditional version of Theorem 4.2).

Theorem 5.1 The conditional variance is determined by the following formulas

$$\begin{aligned}
&Var\left[Z_t \middle| E_t\right] = \sum_{\tau=0}^{\infty} v^{2\tau} \cdot Var\left[C_{t+\tau} \middle| E_t\right] = \sum_{\tau=0}^{\infty} v^{2\tau} \cdot \left(\frac{E\left[C_{t+\tau}^2\right] \cdot P(E_t) - E\left[C_{t+\tau}\right]^2}{P(E_t)^2} \right), \\
&E\left[C_{t+\tau}^2\right] = \sum_{k \in S} E\left[C_{t+\tau}^2 \middle| X_{t+\tau} = k\right] \cdot P\left(X_{t+\tau} = k\right), \quad E\left[C_{t+\tau}\right] = \pi^S(t+\tau) = \sum_{k \in S} \pi_k^S(t+\tau) \\
&E\left[C_{t+\tau}^2 \middle| X_{t+\tau} = k\right] = \left[b_k(t+\tau) - \pi_k(t+\tau)\right]^2 \\
&+ \sum_{j \in S} va_{kj}(t+\tau+1) \cdot \left[b_k(t+\tau) - \pi_k(t+\tau) + va_{kj}(t+\tau+1)\right] \cdot p_{kj}(t+\tau).
\end{aligned} \tag{37}$$

As shown in the next Subsection, these formulas can be used to determine the target capital and solvency capital ratio of a portfolio of GLIFE contracts using appropriate approximations for the distribution of the random present value of future cash-flows associated to this portfolio under the condition that the contracts are still alive.

5.2. Solvency capital and solvency capital ratio for a portfolio of GLIFE contracts

Towards the ultimate goal of solvency evaluation for an arbitrary life insurance portfolio, we consider now a set of n policyholders alive at time t. From Section 3 one knows that the i-th contract $i \in \{1,...,n\}$ is characterized by the following data elements:

- contract duration t_i at time t
- state space $S^{(i)}$
- states $\left(X_k^{(i)}\right)_{k=0,1,2,...}$ of the contract over time with values in $S^{(i)}$
- condition for contract survival $E_{t_i}^{(i)} = \left\{X_{t_i}^{(i)} \neq \phi\right\}$ at time t
- one-step transition probabilities $p_{rs}^{(i)}(k) = P\left(X_{k+1}^{(i)} = s \middle| X_k^{(i)} = r\right), \quad k = 0,1,2,...$, defining the Markov chain model of the contract

- payment function vector $a^{(i)}(k)=\left\{a_r^{(i)}(k),a_{rs}^{(i)}(k)\middle| r\neq s\in S^{(i)}\right\}$ of the contract at time $k=0,1,2,...$ with the two types of payment:

Type 1: $a_r^{(i)}(k), r\in S^{(i)}$, is the payment if the contract at time k is in state r.

Type 2: $a_{rs}^{(i)}(k), r\neq s\in S^{(i)}, k\geq 1$, is the payment if the contract was in state r at time $k-1$ and is in state state s at time k

- splitting $a_r^{(i)}(k)=b_r^{(i)}(k)-\pi_r^{(i)}(k),\quad r\in S^{(i)}, k=0,1,2,...$ into a benefit & premium part

To the i-th contract one associates its random future cash-flows $\left\{C_k^{(i)}\right\}$ as defined in (12), the corresponding $L_{t_i}^{(i)}$ time-t_i random prospective loss (14) and time-t_i net premium reserve $V_{t_i}^{Z,(i)}=E\left[L_{t_i}^{(i)}\middle| E_{t_i}^{(i)}\right]$ obtained from (35). The random present value of future cash-flows of the portfolio is obtained by summing (34) over all contracts and is given by

$$Z_t=\sum_{i=1}^{n}\sum_{j=0}^{\infty}v^j C_{t_i+j}^{(i)}=\sum_{j=0}^{\infty}v^j\left(\sum_{i=1}^{n}C_{t_i+j}^{(i)}\right),\quad t=0,1,.... \tag{38}$$

Similarly, summing the individual net premium reserves, one gets the *portfolio reserve*

$$V_t^Z=\sum_{i=1}^{n}V_{t_i}^{Z,(i)}. \tag{39}$$

Following Section 2, one defines the *portfolio VaR solvency capital*

$$SC_t^{VaR}=VaR_{1-\varepsilon}\left[Z_t\middle| E_{t_i}^{(i)}, i=1,...,n\right]-V_t^Z, \tag{40}$$

as well as the *portfolio CVaR solvency capital*

$$SC_t^{CVaR}=CVaR_{1-\varepsilon}\left[Z_t\middle| E_{t_i}^{(i)}, i=1,...,n\right]-V_t^Z, \tag{41}$$

and the corresponding solvency capital ratios

$$SR_t^{VaR}=SC_t^{VaR}/V_t^Z,\quad SR_t^{CVaR}=SC_t^{CVaR}/V_t^Z. \tag{42}$$

To determine these quantities it is necessary to determine the distribution of Z_t conditional on contract survivals at time t, and under the assumption that the remaining lifetimes of all contracts are independent of each other. From Theorem 5.1 we have

$$\begin{aligned}&E\left[Z_t\middle| E_{t_i}^{(i)}, i=1,...,n\right]=V_t^Z,\quad Var\left[Z_t\middle| E_{t_i}^{(i)}, i=1,...,n\right]=\sum_{\tau=0}^{\infty}v^{2\tau}\cdot\sum_{i=1}^{n}Var\left[C_{t_i+\tau}^{(i)}\right]\\&=\sum_{\tau=0}^{\infty}v^{2\tau}\cdot\sum_{i=1}^{n}\frac{E\left[(C_{t_i+\tau}^{(i)})^2\right]\cdot P(E_{t_i}^{(i)})-E\left[C_{t_i+\tau}^{(i)}\right]^2}{P(E_{t_i}^{(i)})^2},\end{aligned} \tag{43}$$

$$
\begin{aligned}
&E\left[C^{(i)}_{t_i+\tau}\right]=\pi^{S,(i)}(t_i+\tau)=\sum_{k\in S^{(i)}}\pi_k^{S,(i)}(t_i+\tau),\\
&E\left[(C^{(i)}_{t_i+\tau})^2\right]=\sum_{k\in S^{(i)}}E\left[(C^{(i)}_{t_i+\tau})^2\middle|X^{(i)}_{t_i+\tau}=k\right]\cdot P\left(X^{(i)}_{t_i+\tau}=k\right),\\
&E\left[(C^{(i)}_{t_i+\tau})^2\middle|X^{(i)}_{t_i+\tau}=k\right]=\left[b_k^{(i)}(t_i+\tau)-\pi_k^{(i)}(t_i+\tau)\right]^2\\
&+\sum_{j\in S^{(i)}}va_{kj}^{(i)}(t_i+\tau+1)\cdot\left[b_k^{(i)}(t_i+\tau)-\pi_k^{(i)}(t_i+\tau)+va_{kj}^{(i)}(t_i+\tau+1)\right]\cdot p_{kj}^{(i)}(t_i+\tau).
\end{aligned}
\tag{44}
$$

Based on the conditional mean and variance we approximate the distribution function of Z_t by a gamma distribution as in [1], Section 5. Denote this approximation by $F_t(x)=\Pr\left(Z_t\le x\middle|E^{(i)}_{t_i}, i=1,...,n\right)$. Then, recalling the gamma distribution function, one has

$$
F_t(x)=G(\beta_t x;\alpha_t)=\frac{1}{\Gamma(\alpha_t)}\cdot\int_0^{\beta_t x}t^{\alpha_t-1}e^{-t}dt,\quad \alpha_t=\frac{1}{k_t^2},\quad \beta_t=\frac{1}{k_t^2\mu_t},
\tag{45}
$$

where μ_t, k_t are the conditional mean and coefficient of variation of Z_t (obtained from (43)-(44)). In this setting, the solvency capital ratio formulas (42) take the forms

$$
SR_t^{VaR}=z_{1-\varepsilon}\left(k_t^{-2}\right)\cdot k_t^2-1,\ SR_t^{CVaR}=z_{1-\varepsilon}\left(k_t^{-2}\right)\cdot k_t^2\cdot\frac{g\left(z_{1-\varepsilon}\left(k_t^{-2}\right);k_t^{-2}\right)}{\varepsilon},
\tag{46}
$$

where $z_{1-\varepsilon}(\alpha):=G^{-1}(1-\varepsilon;\alpha)$ denotes the $(1-\varepsilon)$-quantile of the standardized gamma $G(x;\alpha)$ and $g(x;\alpha)=G'(x;\alpha)$ denotes its probability density. The limiting results for a portfolio of infinitely growing size are similar to those in [1], Remark 5.1. If the coefficients of variation tend to zero, the gamma distributions converge to normal distributions and the solvency capital ratios converge to zero. This holds under the following assumption. Whenever insured contracts are independent and identically distributed, and if the portfolio size is large enough, then the ratio of observed state transitions to portfolio size is close to the given rates of transition with high probability. This assumption is related to the *process risk*, which describes the random fluctuations in the biometric transition probability matrix. However, if the ratio of observed state transitions to portfolio size is not close to the given rates of transition, even for large portfolio sizes, *systematic risk* exists, e.g. Olivieri & Pitacco [23], Section 2.1. In this situation, the rates of transition are uncertain and assumed to be random, and we consider stochastic models that include the process and systematic risk components. This is the subject of Section 6.3.

6. Comparing the standard approach with variants of the stochastic approach

Since the present Section has some overlap with [1], Section 6, it is treated more briefly, but can be read independently. Facts peculiar to the Markov chain approach are added whenever felt necessary. Recall that *biometric risks* in QIS5 accounts for the uncertainty in trends and parameters, the so-called *systematic/parameter risk*, but not for the *process risk*. We note that the solvency capital models of Section 5.2 only apply to the process risk. For full coverage of the process and systematic risk components, these solvency models are revised and extended. For this, we either shift the biometric transition probability matrix (see Section 6.2) or apply a stochastic biometric model with random biometric rates of transition (see Section 6.3). For completeness we briefly recall the QIS5 standard approach.

6.1. Solvency II standard approach

To value the net premium reserves a biometric "best estimate" life table is chosen. In the Markov chain model the life table is replaced by the one-step transition probabilities $p_{ij}(k) = P(X_{k+1} = j | X_k = i), \quad k = 0,1,2,...$. Given is a single life policy at time of valuation t with net premium reserve V_t^Z. Denote by $V_t^{Z,\Delta}$ the value of the reserves subject to a biometric shock Δ. The one-year *solvency capital requirement* (SCR) for this single policy is

$$SCR_t = V_t^{Z,\Delta} - V_t^Z. \tag{47}$$

Similarly to the decomposition (7) the *Solvency II target capital* (upper index S2 in quantities) is understood as the sum of the SCR and a risk margin defined by

$$TC_t^{S2} = SCR_t + RM_t. \tag{48}$$

$$RM_t = i_{CoC} \cdot \sum_{k=1}^{T-t} v_f^k \cdot SCR_{t+k}, \; i_{CoC} = 6\%, \tag{49}$$

where T denotes the time horizon, which may depend on the life policy, and v_f is the risk-free discount rate. Since Solvency II uses a total balance sheet approach, the defined single policy quantities must be aggregated on a portfolio and/or line of business level. For comparison with internal models it is useful to consider the solvency capital ratio at time t under the Solvency II standard approach defined by the quotient

$$SR_t^{S2} = SCR_t / V_t^Z. \tag{50}$$

By using a matrix of transition shocks $\Delta = (\Delta_{ij})$, we obtain formulas for the Markov chain model. Consider the *shifted biometric transition probabilities* defined by

$$p_{ij}^{\Delta_{ij}}(k), \tag{51}$$

which is associated to $p_{ij}(k)$, for a permanent shift of amount Δ_{ij} over all contracts and years $k = 0,1,2,...$. In the current specification one has $\Delta_{AD} = 0.15$ (permanent 15% increase in mortality rates at each age for jumping from the alive state "A" to the dead state "D" for the mortality risk), $\Delta_{AD} = -0.20$ (permanent 20% decrease in mortality rates at each age for jumping from "A" to "D" for the longevity risk), and $\Delta_{AJ} = 0.35$ respectively $\Delta_{AJ} = 0.25$ (increase of 35% in disability rates for the next year, respectively a permanent 25% increase in disability rates at each age in following years for jumping from "A" to the disability state "J" for the disability risk). To calculate the portfolio reserve V_t^Z in (39) and the corresponding shifted value $V_t^{Z,\Delta}$ under a matrix of transition shocks $\Delta = \left(\Delta_{ij}\right)$, we use (34) and the backward recursion formula (18) to get

$$\begin{aligned} V_t^Z &= \sum_{i=1}^{n} V_{t_i}^{Z,(i)} = \sum_{i=1}^{n}\sum_{k\in S} V_{t_i}^{k,(i)} \cdot P\left(X_{t_i}^{(i)} = k \middle| E_{t_i}^{(i)}\right), \\ V_{t_i}^{k,(i)} &= \sum_{j\in S} p_{kj}(t_i)\cdot\left\{v\cdot V_{t_i+1}^{j,(i)} + v\cdot a_{kj}^{(i)}(t_i+1) + b_k^{(i)}(t_i)\right\} - \pi_k^{(i)}(t_i), \end{aligned} \tag{52}$$

$$\begin{aligned} V_t^{Z,\Delta} &= \sum_{i=1}^{n} V_{t_i}^{Z,(i),\Delta} = \sum_{i=1}^{n}\sum_{k\in S} V_{t_i}^{k,(i),\Delta} \cdot P\left(X_{t_i}^{(i),\Delta} = k \middle| E_{t_i}^{(i),\Delta}\right), \\ V_{t_i}^{k,(i),\Delta} &= \sum_{j\in S} p_{kj}^{\Delta_{kj}}(t_i)\cdot\left\{v\cdot V_{t_i+1}^{j,(i),\Delta} + v\cdot a_{kj}^{(i)}(t_i+1) + b_k^{(i)}(t_i)\right\} - \pi_k^{(i),\Delta}(t_i). \end{aligned} \tag{53}$$

Similarly to (47)-(50) and using (52)-(53) we obtain the risk capital formulas

$$\begin{aligned} &SCR_t = V_t^{Z,\Delta} - V_t^Z, \quad TC_t^{S2} = SCR_t + RM_t, \\ &RM_t = i_{CoC} \sum_{k=1}^{\max\{\omega-(x_i+t_i)\}} v_f^k \cdot SCR_{t+k}, \quad SR_t^{S2} = \frac{SCR_t}{V_t^Z}. \end{aligned} \tag{54}$$

6.2. Stochastic approach: Shifting the biometric transition probability matrix

Following the Sections 5.2 and 6.1, we consider the "shifted" random present value Z_t^{Δ} of future cash-flows of the portfolio at time t with conditional mean and variance

$$\begin{aligned} &E\left[Z_t^{\Delta} \middle| E_{t_i}^{(i),\Delta}, i=1,...,n\right] = V_t^{Z,\Delta}, \quad Var\left[Z_t^{\Delta} \middle| E_{t_i}^{(i),\Delta}, i=1,...,n\right] = \sum_{\tau=0}^{\infty} v^{2\tau}\cdot\sum_{i=1}^{n} Var\left[C_{t_i+\tau}^{(i),\Delta}\right] \\ &= \sum_{\tau=0}^{\infty} v^{2\tau}\cdot\sum_{i=1}^{n}\frac{E\left[(C_{t_i+\tau}^{(i),\Delta})^2\right]\cdot P(E_{t_i}^{(i),\Delta}) - E\left[C_{t_i+\tau}^{(i),\Delta}\right]^2}{P(E_{t_i}^{(i),\Delta})^2}, \end{aligned} \tag{55}$$

$$\begin{aligned} &E\left[C_{t_i+\tau}^{(i),\Delta}\right] = \pi^{S,(i),\Delta}(t_i+\tau) = \sum_{k\in S^{(i)}} \pi_k^{S,(i),\Delta}(t_i+\tau), \\ &E\left[(C_{t_i+\tau}^{(i),\Delta})^2\right] = \sum_{k\in S^{(i)}} E\left[(C_{t_i+\tau}^{(i),\Delta})^2 \middle| X_{t_i+\tau}^{(i),\Delta} = k\right]\cdot P\left(X_{t_i+\tau}^{(i),\Delta} = k\right), \\ &E\left[(C_{t_i+\tau}^{(i),\Delta})^2 \middle| X_{t_i+\tau}^{(i),\Delta} = k\right] = \left[b_k^{(i)}(t_i+\tau) - \pi_k^{(i),\Delta}(t_i+\tau)\right]^2 \\ &+ \sum_{j\in S^{(i)}} va_{kj}^{(i)}(t_i+\tau+1)\cdot\left[b_k^{(i)}(t_i+\tau) - \pi_k^{(i),\Delta}(t_i+\tau) + va_{kj}^{(i)}(t_i+\tau+1)\right]\cdot p_{kj}^{(i),\Delta_{kj}}(t_i+\tau). \end{aligned} \tag{56}$$

The distribution of Z_t^{Δ} conditional on contract survivals at time t is again approximated by a gamma distribution denoted by

$$F_t^{\Delta}(x) = G(\beta_t^{\Delta} x; \alpha_t^{\Delta}), \quad \alpha_t^{\Delta} = (k_t^{\Delta})^{-2}, \beta_t^{\Delta} = (k_t^{\Delta})^{-2} / \mu_t^{\Delta}, \tag{57}$$

where the conditional mean and coefficient of variation $\mu_t^{\Delta}, k_t^{\Delta}$ of Z_t^{Δ} are obtained from the formulas (55)-(56). Making use of (46) and (47) one sees that the portfolio VaR & CVaR solvency capitals under the shifted biometric transition probability matrix are given by the expressions

$$SC_t^{\Delta,VaR} = VaR_{1-\varepsilon}\left[Z_t^{\Delta} \middle| E_{t_i}^{(i),\Delta}, i = 1,...,n \right] - V_t^{Z,\Delta} = SCR_t + \left(z_{1-\varepsilon}\left(\left[k_t^{\Delta} \right]^{-2} \right) \cdot \left[k_t^{\Delta} \right]^2 - 1 \right) \cdot V_t^{Z,\Delta}, \tag{58}$$

$$\begin{aligned} SC_t^{\Delta,CVaR} &= CVaR_{1-\varepsilon}\left[Z_t^{\Delta} \middle| E_{t_i}^{(i),\Delta}, i = 1,...,n \right] - V_t^{Z,\Delta} \\ &= SCR_t + z_{1-\varepsilon}\left(\left[k_t^{\Delta} \right]^{-2} \right) \cdot \left[k_t^{\Delta} \right]^2 \cdot \frac{g\left(z_{1-\varepsilon}\left(\left[k_t^{\Delta} \right]^{-2} \right); \left[k_t^{\Delta} \right]^{-2} \right)}{\varepsilon} \cdot V_t^{Z,\Delta}. \end{aligned} \tag{59}$$

The observations in [1], Section 6.2, hold for the Markov chain model. By small coefficients of variation the gamma distributions converge to normal distributions, and the corresponding solvency capitals converge to those of normal distributions such that

$$SC_t^{\Delta,VaR} = SCR_t + \Phi^{-1}(1-\varepsilon) k_t^{\Delta} V_t^{Z,\Delta}, \quad SC_t^{\Delta,CVaR} = SCR_t + \frac{\varphi\left[\Phi^{-1}(1-\varepsilon) \right]}{\varepsilon} k_t^{\Delta} V_t^{Z,\Delta}. \tag{60}$$

Asymptotically, the solvency capital ratios tend to the following minimum values

$$\lim_{k_t^{\Delta} \to 0} SR_t^{\Delta,VaR} = \lim_{k_t^{\Delta} \to 0} SR_t^{\Delta,CVaR} = \frac{SCR_t}{V_t^Z}. \tag{61}$$

By vanishing coefficients of variation the VaR & CVaR solvency capital ratios converge to the Solvency II solvency capital ratio. In this situation, the process risk has been fully diversified away, and, as expected, only the parameter/systematic risks remain.

6.3. Stochastic approach: Poisson-gamma model of biometric transition

For simplicity let us fix the states i, j of the transition probabilities $p_{ij}(k), k = 0,1,2,...$. In case the ratio of observed state transitions to portfolio size is not close to the given rates of transition, even for large portfolio sizes, systematic risk exists. In this situation, the transition rates are uncertain and assumed to be random. This situation is modelled similarly to [1], Section 6.3. We assume a Bayesian Poisson-Gamma model such that the number of transitions is conditional Poisson distributed with a Gamma distributed random transition probability, which results in a negative binomial distribution for the unconditional distribution of the number of transitions. Then, we consider a Poisson-Gamma model with time-dependence of the type introduced in Olivieri & Pitacco [23], which up-dates its parameters to experience. Given is a fixed time t and biometric transition probabilities $p_{ij}(k), k=0,1,2,...$, for the given fixed states, which is based on an initial cohort of size ℓ_t at time t. Let $D_{t+\tau-1}$ denote the random number of transitions produced by the cohort in the time period $(t+\tau-1,t+\tau]$, $\tau=1,2,...$. For the first time period $\tau=1$, we assume that there is no experience available and that the random number of transitions is conditional Poisson distributed such that

$$D_t \sim Po(\ell_t Q_t), \quad Q_t \sim Gamma\left(\alpha, \frac{\beta}{p_{ij}(t)}\right). \tag{62}$$

It follows that the unconditional distribution of the number of transitions in the first time period is negative binomially distributed such that

$$D_t \sim NB\left(\alpha, \frac{\theta_1}{\theta_1+1}\right), \quad \theta_1 = \frac{\beta}{p_{ij}(t)}. \tag{63}$$

In contrast to the expected number of transitions $\ell_t p_{ij}(t)$ predicted by the biometric transition probability matrix, one has

$$E[D_t] = \frac{\alpha}{\beta} \cdot \ell_t p_{ij}(t).. \tag{64}$$

To model a systematic deviation from the expectation, one assumes that the quotient α/β is different from one, for example greater than one for transitions produced by the mortality and disability risks and less than one for those produced by the longevity risk. Suppose that at time $t+1$, the number of transitions d_t observed in the cohort over the first time period is available, and let $\ell_{t+1} = \ell_t - d_t$ be the up-dated cohort size. A calculation shows that the posterior distribution of Q_t conditional on the information $D_t = d_t$ is Gamma distributed

$$Q_t | d_t \sim Gamma\left(\alpha + d_t, \frac{\beta + \ell_t p_{ij}(t)}{p_{ij}(t)}\right), \tag{65}$$

which shows that the initial structural systematic risk parameters are up-dated as follows:

$$(\alpha, \beta) \rightarrow (\alpha + d_t, \beta + \ell_t p_{ij}(t)). \tag{66}$$

Passing to the second time period $(t+1,t+2]$, we assume similarly to the first period that

$$D_{t+1} | d_t \sim Po(\ell_{t+1} Q_{t+1}), \quad Q_{t+1} | d_t \sim Gamma\left(\alpha + d_t, \frac{\beta + \ell_t p_{ij}(t)}{p_{ij}(t+1)}\right). \tag{67}$$

This implies a so-called predictive distribution of negative binomial type

$$D_{t+1}\left|d_t \sim NB\left(\alpha+d_t, \frac{\theta_2}{\theta_2+1}\right), \quad \theta_2=\frac{\beta+\ell_t p_{ij}(t)}{\ell_{t+1}p_{ij}(t)}.\right. \tag{68}$$

Iterating the above Bayesian scheme, one generalizes as follows. At time $t+\tau-1, \tau \geq 2$, having observed the annual number of transitions $d_t, d_{t+1}, ..., d_{t+\tau-2}$, the up-dated cohort size for the next time period $\left(t+\tau-1, t+\tau\right]$ is obtained from the recursion $\ell_{t+k-1}=\ell_{t+k-2}-d_{t+k-2}$, $k=2,3,...,\tau$. The corresponding predictive negative binomial distribution of the number of transitions is then given by

$$D_{t+\tau-1}\left|d_t, d_{t+1},...,d_{t+\tau-2} \sim NB\left(\alpha+\sum_{k=0}^{\tau-2}d_{t+k}, \frac{\theta_\tau}{\theta_\tau+1}\right), \quad \theta_\tau=\frac{\beta+\sum_{k=1}^{\tau-1}\ell_{t+k-1}p_{ij}(t+k-1)}{\ell_{t+\tau-1}p_{ij}(t+\tau-1)}.\right. \tag{69}$$

The conditional expected number of transitions compares with the given expected number as follows:

$$\begin{aligned} &E\left[D_{t+\tau-1}\left|d_t, d_{t+1},...,d_{t+\tau-2}\right.\right]=\frac{\alpha+\sum_{k=1}^{\tau-1}d_{t+k-1}}{\beta+\sum_{k=1}^{\tau-1}\ell_{t+k-1}p_{ij}(t+k-1)}\cdot \ell_{t+\tau-1}p_{ij}(t+\tau-1) \\ &\lesseqgtr E\left[D_{t+\tau-1}\right]=\ell_t \cdot {}_{\tau-1}p_{ij}(t)p_{ij}(t+\tau-1). \end{aligned} \tag{70}$$

If biometric experience is consistent with what is expected, the quotient of both expected values remains constant over time. On the other hand, if experience is better (worse) than expected, the same quotient will increase (decrease) over time.

In practice one proceeds as follows. Given a fixed time t, consider for each pair of fixed states the Poisson-Gamma transition probabilities obtained from (64) and (70) defined by

$$\begin{aligned} &p_{ij}^{PG}(t)=\frac{\alpha}{\beta}\cdot p_{ij}(t), \quad p_{ij}^{PG}(t+\tau)=\frac{\alpha+\sum_{k=1}^{\tau}d_{t+k-1}}{\beta+\sum_{k=1}^{\tau}\ell_{t+k-1}p_{ij}(t+k-1)}\cdot p_{ij}(t+\tau), \quad \tau=1,2,..., \\ &{}_k p_{ij}^{PG}(t+\tau)={}_{k-1}p_{ij}^{PG}(t+\tau)\cdot\left(1-p_{ij}^{PG}(t+\tau+k-)\right), \ {}_0 p_{ij}^{PG}(0)=1, \quad \tau=0,1,... \end{aligned} \tag{71}$$

Replacing everywhere in the formulas (55)-(59) the superscripts Δ_{ij} by the superscript PG for the relevant transition probabilities and using (71) in calculations, we obtain portfolio VaR & CVaR solvency capital formulas under the Poisson-Gamma model of biometric transition similar to (58) and (59). Similar limiting results apply. An implementation requires a detailed specification. To be consistent with the Solvency II standard approach, one can assume that future transitions deviate systematically from the biometric life table according to the prescribed shock $\Delta=\Delta_{ij}$ for the given fixed states, such that

$$\ell_{s+t+k} = \ell_{s+t+k-1} - d_{s+t+k-1}, \quad d_{s+t+k-1} = (1-\Delta)\cdot \ell_{s+t+k-1} q_{s+t+k-1}, \quad k = 1,2,\dots \,. \tag{72}$$

This choice is consistent with the expected number of transitions in the first period $E\left[D_{s+t}\right] = d_{s+t}$ if in (64) one sets $\alpha = (1-\Delta)\beta$. Assume further that $\beta = 100$, which implies a coefficient of variation for D_{s+t} equal to 10%. One shows that the choice (72) with $\alpha = (1-\Delta)\beta$ implies that the transition probabilities (71) coincide with the corresponding shifted entries in the biometric life table. In this special case, we observe that the stochastic model of Section 6.3 provides the same results as the shift method of Section 6.2. In general, the stochastic model of Section 6.3 is more satisfactory and flexible because it allows the use of effective observed numbers of transitions as time elapses.

7. The endowment contract with waiver of premium by disability

For a clear and simple Markov chain illustration we restrict the attention to a single cohort of identical n-year endowment contracts with waiver of premium in the event of disability and fixed one-unit of sum insured payable upon death or survival at maturity date. The treatment of other similarly complex disability contracts is left to future research. For some further possibilities consult Example 2.1 in Christiansen et al. [24].

7.1. Markov model for mortality and disability risks

A complete risk model for single-life insurance products with mortality and disability risks requires the specification of a Markov model with three states. A policyholder aged x at contract issue changes state at time $t > 0$ according to the following diagram

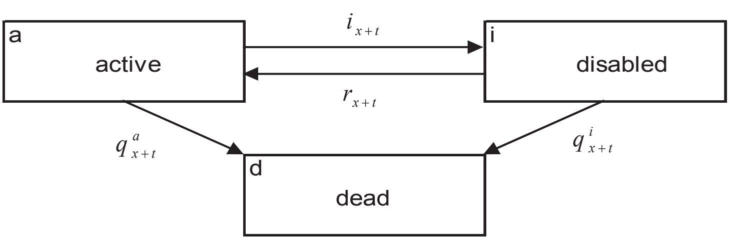


Figure 1. Markov chain states and their jump probabilities

The possible state changes occur with the following probabilities

i_{x+t} : one-year probability of disability at time t

q^a_{x+t} : one-year probability of active mortality at time t

q^i_{x+t} : one-year probability of disabled mortality at time t

r_{x+t} : one-year probability of recovery at time t

The states are $S = \{1,2,3\} = \{a,i,d\}$. The one-step transition probabilities are given by

$$p_{ij}(k)=\begin{pmatrix} 1-i_{x+k}-q^a_{x+k} & i_{x+k} & q^a_{x+k} \\ r_{x+k} & 1-r_{x+k}-q^i_{x+k} & q^i_{x+k} \\ 0 & 0 & 1 \end{pmatrix},\ i,j\in\{1,2,3\},\ k=0,1,2,... \tag{73}$$

For a n-year endowment contract with waiver of premium by disability without recovery from disability, one has $r_{x+k}=0,\ k=1,2,...,n-1$. This simplifying assumption is sometimes made in practice and justified in economic environments with a small number of disabled persons, for which the probability of recovery can be neglected. For example, the Swiss Federal Insurance Pension applies such a model and uses a biometric life table called "EVK Table", where EVK is the abbreviation for "**E**idgenössische **V**ersicherungs**K**asse", e.g. Koller [6], p.129, or Chuard [25] for a detailed historical background.

7.2. State dependent actuarial reserves and net level premiums

The net level premium of the n-year endowment with waiver of premium and one unit of sum insured for a life in the active state at age x is denoted by $P^a=P^a(x:n)$, where the upper index indicates that the premium is only due if the contract remains in the active life state. In our notations the *payment functions* of this contract are defined by

Type 1:

$$a_1(k)=\begin{cases} -P^a,\ k=0,1,...,n-1 \\ 1,\ k=n \end{cases},\quad a_2(k)=\begin{cases} 0,\ k=1,2,...,n-1 \\ 1,\ k=n \end{cases},\quad a_3(k)=0$$

Type 2:

$$a_{12}(k)=\begin{cases} 0,\ k=1,2,...,n-1 \\ 1,\ k=n \end{cases},\quad a_{13}(k)=1,\ k=1,2,...,n,\quad a_{21}(k)=0.$$

To describe the state and time dependent actuarial reserves $V^i_k,\ i\in\{1,2,3\},\ k=0,1,...,n$, one needs the survival probabilities of staying in the active or disabled state. Denote by ω the maximum attainable age. Then, the *active survival probabilities* (probability a life in the active state at age x will attain age $x+k$ in the active state without disablement) are given by

$${}_kp^a_x={}_{k-1}p^a_x\cdot\left(1-q^a_{x+k-1}-i_{x+k-1}\right),\quad {}_0p^a_x=1,\quad k=1,...,\omega-x. \tag{74}$$

Similarly, the *disabled survival probabilities* (probability a life in the disabled state at age x will attain age $x+k$ in the disabled state without recovery) are given by

$${}_kp^i_x={}_{k-1}p^i_x\cdot\left(1-q^i_{x+k-1}\right),\quad {}_0p^i_x=1,\quad k=1,...,\omega-x \tag{75}$$

Corresponding to these survival probabilities one associates n-year life annuities for a life aged x being in the active or disabled state whose actuarial present values are defined by

$$a^a(x:n)=\sum_{k=0}^{n-1} v^k \cdot {}_k p_x^a, \text{ for the } n-\textit{year active life annuity} \tag{76}$$

$$a^i(x:n)=\sum_{k=0}^{n-1} v^k \cdot {}_k p_x^i, \text{ for the } n-\textit{year disabled life annuity} \tag{77}$$

The actuarial present value (APV) of future benefits for the n -year endowment with waiver of premium and one unit of sum insured for a life in the active (respectively disabled) state at age x is denoted by $A^a(x:n)$ (respectively $A^i(x:n)$). Using the backward recursive formulas for the state dependent actuarial reserves let us determine formulas for the evaluation of the introduced APVs. In particular, an explicit formula for the net level premium is derived. The backward recursive reserve formulas are given by

$$\begin{aligned} &V_k^1 = v\cdot\left(q_{x+k}^a + p_{x+k}^a V_{k+1}^1 + i_{x+k} V_{k+1}^2\right) - P^a, \quad k=0,1,...,n-1, \quad V_n^1 = 1, \\ &V_k^2 = v\cdot\left(q_{x+k}^i + p_{x+k}^i V_{k+1}^2\right), \quad k=2,3,...,n-1, \quad V_n^2 = 1. \end{aligned} \tag{78}$$

One has $V_0^2=0$ because the life is in the state "a" at contract issue, $V_1^2=0$ because the life can only be in the state "i" after at least one year and then no actuarial reserve is available, and $V_k^3=0, \quad k=0,1,...,n$, because no actuarial reserve is required in case the insured life has died. Since actuarial reserves represent differences between APVs of future benefits and future premiums one has further the relationships

$$\begin{aligned} &V_k^1 = A^a(x+k:n-k) - P^a(x:n)\cdot a^a(x+k:n-k), \quad k=0,1,...,n-1, \\ &V_k^2 = A^i(x+k:n-k), \quad k=2,3,...,n-1. \end{aligned} \tag{79}$$

On the other hand the APVs of the active life annuities in (79) satisfy the recursions

$$a^a(x+k:n-k) = 1 + v\cdot p_{x+k}^a \cdot a^a(x+k+1:n-k-1), \quad k=0,1,...,n-1. \tag{80}$$

Inserting (79) and (80) into (78) one obtains the backward recursions for APVs

$$\begin{aligned} &A^a(x+n-k:k) = v\cdot\begin{pmatrix} q_{x+n-k}^a + p_{x+n-k}^a A^a(x+n-k+1:k-1) \\ +i_{x+n-k} A^i(x+n-k+1:k-1) \end{pmatrix}, \quad k=2,3,...,n, \\ &A^i(x+n-k:k) = v\cdot\left(q_{x+n-k}^i + p_{x+n-k}^i A^i(x+n-k+1:k-1)\right), \quad k=2,3,...,n-1, \end{aligned} \tag{81}$$

with the starting values $A^a(x+n-1:1)=A^i(x+n-1:1)=v$. One sees that the second relationship in (81) is satisfied by the formula

$$A^i(x+n-k:k) = 1 - d\cdot a^i(x+n-k:k), \quad k=1,2,...,n-1, \tag{82}$$

which reminds one of the usual formula for an endowment insurance with disability as single cause of decrement, e.g. Gerber [10], formula (2.15), p.37. Inserting (77) and rearranging one obtains the corresponding explicit sum representation

$$A^i(x+n-k:k)=\sum_{j=1}^{k-1} v^j \cdot {}_{j-1}p^i_{x+n-k} \cdot q^i_{x+n-k+j-1} + v^k \cdot {}_{k-1}p^i_{x+n-k}, \quad k=2,3,...,n-1. \tag{83}$$

Using these results and proceeding through backward induction, one obtains the following explicit formula for the evaluation of the APV of future benefits for a life in the active state

$$\begin{aligned} A^a(x:n) &= \sum_{k=1}^{n-1} v^k \cdot {}_{k-1}p^a_x \cdot q^a_{x+k-1} + v^n \cdot {}_{n-2}p^a_x \cdot (1-q^a_{x+n-2}) \\ &+ \sum_{k=2}^{n-1} v^k \cdot \left(\sum_{j=0}^{k-2} {}_j p^a_x \cdot i_{x+j} \cdot {}_{k-2-j}p^i_{x+j+1} \right) \cdot q^i_{x+k-1} + v^n \cdot \sum_{j=0}^{n-3} {}_j p^a_x \cdot i_{x+j} \cdot {}_{n-2-j}p^i_{x+j+1} \end{aligned} \tag{84}$$

The net level premium is determined by the actuarial equivalence principle, which states that at contract issue $V_0^1=0$. Using (79) one obtains the explicit formula (use (76), (84))

$$P^a(x:n)=A^a(x:n)\,/\,a^a(x:n). \tag{85}$$

7.3. Conditional mean and variance of the prospective insurance loss

We determine the conditional mean and variance given survival of the time-t prospective insurance loss $Z_t=\sum_{k=0}^{n-1-t} v^k C_{t+k}$, $t=0,1,...,n-1$. Besides the state dependent actuarial reserves (79), we consider the net premium reserve (35), which coincides with the conditional mean

$$\begin{aligned} V_t^Z &:= E\left[Z_t \middle| E_t\right] = V_t^1 \cdot P\left(X_t=a \middle| E_t\right) + V_t^2 \cdot P\left(X_t=i \middle| E_t\right) \\ &= V_t^1 \cdot \frac{P\left(X_t=a\right)}{P(E_t)} + V_t^2 \cdot \frac{P\left(X_t=i\right)}{P(E_t)} = \frac{V_t^1 \cdot {}_t p^a_x + V_t^2 \cdot i_x \cdot {}_{t-1}p^i_{x+1}}{{}_t p^a_x + i_x \cdot {}_{t-1}p^i_{x+1}}, \quad t=0,1,...,n-1, \end{aligned} \tag{86}$$

$$\begin{aligned} &P\left(X_t=a\right)={}_t p^a_x, \quad t=0,1,2,...,n-1, \\ &P\left(X_t=i\right)=i_x \cdot {}_{t-1}p^i_{x+1}, \quad t=1,2,...,n-1, \quad P\left(X_0=i\right)=0, \\ &P(E_t)={}_t p^a_x + i_x \cdot {}_{t-1}p^i_{x+1}, \quad t=1,2,...,n-1, \quad P(E_0)=0. \end{aligned} \tag{87}$$

Furthermore, a calculation based on Theorem 5.1 yields the following conditional variances

$$\begin{aligned} &(\sigma_t^Z)^2 := Var\left[Z_t \middle| E_t\right] = \sum_{k=0}^{n-1-k} v^{2k} \cdot \left(\frac{E\left[C_{t+k}^2\right] \cdot P(E_k) - \pi^S(t+k)^2}{P(E_k)^2} \right), \quad t=0,1,...,n-1, \\ &E\left[C_0^2\right] = P^a(x:n)^2 + v\left(v+P^a(x:n)\right) q^a_x, \\ &E\left[C_k^2\right] = \left[P^a(x:n)^2 + v\left(v+P^a(x:n)\right) q^a_{x+k}\right] {}_k p^a_x + v^2 i_x \cdot {}_{k-1}p^i_{x+1} q^i_{x+k}, \quad k=1,...,n-2, \\ &E\left[C_{n-1}^2\right] = \left[P^a(x:n)^2 + v\left(v+P^a(x:n)\right)(1-p^a_{x+n-1})\right] {}_{n-1}p^a_x + v^2 i_x \cdot {}_{n-2}p^i_{x+1} q^i_{x+n-1}, \end{aligned} \tag{88}$$

where the savings premiums are determined by the formulas

$$\begin{aligned} &\pi^S(0) = vV_1^1, \\ &\pi^S(k) = v\left(V_{k+1}^1 + A^i(x+k+1:n-k-1)\right) - \left(V_k^1 + A^i(x+k:n-k)\right), \quad k = 1,...,n-1, \end{aligned} \tag{89}$$

and the probabilities $P(E_k)$ are defined in (87). Neglecting the probabilistic terms of second order $i_x \cdot {}_{k-1}p_{x+1}^i q_{x+k}^i$, one obtains the following simpler approximations to (88):

$$\begin{aligned} &E\left[C_k^2\right] = \left\{P^a(x:n)^2 + v\left(v + P^a(x:n)\right)q_{x+k}^a\right\} \cdot {}_k p_x^a, \quad k = 0,...,n-2, \\ &E\left[C_{n-1}^2\right] = \left\{P^a(x:n)^2 + v\left(v + P^a(x:n)\right)(1 - p_{x+n-1}^a)\right\} \cdot {}_{n-1}p_x^a. \end{aligned} \tag{90}$$

7.4. Numerical illustration

The Markov chain parameterization of the present contract type has been given at the beginning of Section 7.2. We assume that all the policyholders are aged x at time $t = 0$. Our construction of the biometric life table with mortality and disability risk factors is based on the classical textbook Saxer [26], Section 2.5. Besides the one-year probabilities introduced in Section 7.1, one considers further the partial or independent rates of decrement, see Saxer [26], Section 2.4, or Bowers et al. [19], Section 9.5, denoted by

${}^*q_{x+t}^a$: one-year independent rate of active mortality at time t

${}^*i_{x+t}$: one-year independent rate of disability at time t

The independent rates of decrement are linked to the probabilities of active mortality and disability through the relationship, e.g. Saxer [26], formulas (2.5.1) and (2.5.2),

$$1 - q_{x+t}^a - i_{x+t} = \left(1 - {}^*q_{x+t}^a\right) \cdot \left(1 - {}^*i_{x+t}\right). \tag{91}$$

For the purpose of illustration only and by lack of another reference, we base our calculations on Table 1, which is obtained by combining the Tables 4 and 5 in Saxer [26], p.240. The entries ${}^*q_{x+t}^a$, ${}^*i_{x+t}$ and q_{x+t}^i are taken from the "EVK Table 1950" and the entry i_{x+t} is taken from the "VZ Table 1950", where VZ stands for "**V**ersicherungskasse **Z**ürich". The missing entries between the 5-year ages are linearly interpolated.

While the standard solvency capital ratio does not depend on the initial cohort size, this is the case for the stochastic approaches. The age at contract issue is either $x = 30$ or $x = 40$ and the contract duration is $n = 20$. We compare the stochastic approach with the standard approach for the contract times $t \in \{0,1,2,3,4,5,10,15,18\}$. We use the shifted biometric life table with Solvency II standard like specifications, namely at each age 20% decrease for the probability to die as active (longevity risk) respectively 15% increase for the probability to die as disabled (mortality risk), 35% increase for the first year probability to disable and then 25% increase at each future age (disability risk). The interest rate and the risk-free interest rate is 3%. Table 2 displays shifted coefficients of variation under varying cohort sizes. The values are sufficiently small so that the normal approximation to the gamma distribution can be applied. Table 3, which is based on (60), displays the cohort size dependent solvency capital ratios and their limiting values (61) for a portfolio of infinitely growing size. The chosen confidence level is 99.5% for VaR and 99% for CVaR (the accepted level, which corresponds to a 99.5% Solvency II calibration).

x	*qax	qax	*ix	ix	qix
30	0.183%	0.106%	0.064%	0.141%	4.348%
31	0.181%	0.097%	0.064%	0.148%	4.324%
32	0.178%	0.087%	0.064%	0.154%	4.300%
33	0.176%	0.078%	0.063%	0.161%	4.276%
34	0.173%	0.069%	0.063%	0.167%	4.252%
35	0.171%	0.060%	0.063%	0.174%	4.228%
36	0.174%	0.059%	0.072%	0.186%	4.204%
37	0.177%	0.059%	0.080%	0.198%	4.180%
38	0.180%	0.059%	0.089%	0.210%	4.156%
39	0.183%	0.058%	0.097%	0.222%	4.132%
40	0.186%	0.058%	0.106%	0.234%	4.108%
41	0.207%	0.079%	0.129%	0.257%	4.084%
42	0.227%	0.099%	0.153%	0.280%	4.061%
43	0.248%	0.120%	0.176%	0.304%	4.037%
44	0.268%	0.141%	0.200%	0.327%	4.014%
45	0.289%	0.161%	0.223%	0.350%	3.990%
46	0.335%	0.218%	0.285%	0.401%	3.966%
47	0.382%	0.275%	0.347%	0.452%	3.942%
48	0.428%	0.331%	0.408%	0.504%	3.919%
49	0.475%	0.388%	0.470%	0.555%	3.895%
50	0.521%	0.444%	0.532%	0.606%	3.871%
51	0.586%	0.532%	0.694%	0.743%	3.848%
52	0.650%	0.620%	0.856%	0.881%	3.824%
53	0.715%	0.708%	1.019%	1.018%	3.801%
54	0.779%	0.795%	1.181%	1.156%	3.777%
55	0.844%	0.883%	1.343%	1.293%	3.754%
56	0.909%	0.892%	1.765%	1.766%	3.731%
57	0.974%	0.900%	2.187%	2.240%	3.707%
58	1.040%	0.908%	2.609%	2.713%	3.684%
59	1.105%	0.916%	3.031%	3.187%	3.660%
60	1.170%	0.923%	3.453%	3.660%	3.637%

Table 1. One-step transition probabilities for the mortality and disability Markov chain

In the present case study, we observe that for all cohort sizes and contract times, the current standard approach prescribes almost negligible solvency capital ratios. For small cohort sizes and early contract times, the discrepancies between the stochastic and standard approach increase with age and contract duration attaining solvency capital ratios above 200% for small cohort sizes with 100 insured lives. In fact, as already explained, the current QIS5 specification neglects the process risk. Moreover, we note that the chosen results for the normal distribution are only approximate, especially for small cohort sizes. In this respect, we think that the displayed figures are most likely lower bounds due to the fact that often a normal approximation rather underestimates than overestimates risk. A more detailed analysis of this point is left as open issue for further investigation (however, the use of the gamma approximation makes no big difference). On the other hand, solvency capital ratios of cohort sizes exceeding 10'000 policyholders and late contract times tend more and more to the lower limiting bound as expected from the central limit theorem. Fig. 2 visualizes these findings. In virtue of the made confidence level calibration, the VaR & CVaR solvency capital ratios are of the same order of magnitude. Finally, the considered example points out to another difficulty. Though almost negligible in absolute value, we note that the standard solvency capital ratios change their signs repeatedly over the time axis. In this respect, one can ask whether fixed transition shifts are the "crucial scenarios". As a response to this "biometric worst- and best-case scenarios" are proposed in Christiansen [27], [28].

Cohort size	Contract Time								
	0	1	2	3	4	5	10	15	18
(x,n)=(30,20)									
100	53.815%	26.466%	17.602%	13.150%	10.483%	8.714%	4.628%	2.525%	1.464%
500	24.067%	11.836%	7.872%	5.881%	4.688%	3.897%	2.070%	1.129%	0.655%
1'000	17.018%	8.369%	5.566%	4.158%	3.315%	2.756%	1.463%	0.798%	0.463%
10'000	5.381%	2.647%	1.760%	1.315%	1.048%	0.871%	0.463%	0.252%	0.146%
100'000	1.702%	0.837%	0.557%	0.416%	0.332%	0.276%	0.146%	0.080%	0.046%
(x,n)=(40,20)									
100	91.895%	45.158%	30.663%	23.298%	18.838%	15.846%	8.844%	5.223%	3.176%
500	41.097%	20.195%	13.713%	10.419%	8.425%	7.087%	3.955%	2.336%	1.421%
1'000	29.060%	14.280%	9.696%	7.368%	5.957%	5.011%	2.797%	1.652%	1.004%
10'000	9.190%	4.516%	3.066%	2.330%	1.884%	1.585%	0.884%	0.522%	0.318%
100'000	2.906%	1.428%	0.970%	0.737%	0.596%	0.501%	0.280%	0.165%	0.100%

Table 2. Coefficients of variation of the shifted prospective insurance loss

Cohort size	Contract Time								
	0	1	2	3	4	5	10	15	18
(x,n)=(30,20)									
100	138.7%	68.9%	45.7%	34.1%	27.1%	22.5%	11.8%	6.4%	3.8%
500	62.0%	31.1%	20.6%	15.3%	12.2%	10.1%	5.2%	2.8%	1.7%
1'000	43.9%	22.1%	14.6%	10.9%	8.6%	7.1%	3.7%	2.0%	1.2%
10'000	13.9%	7.3%	4.8%	3.5%	2.8%	2.3%	1.1%	0.6%	0.4%
100'000	4.4%	2.6%	1.7%	1.2%	0.9%	0.7%	0.3%	0.1%	0.1%
Limiting QIS5 ratio SCRt/VZt	0.0%	0.5%	0.3%	0.1%	0.1%	0.0%	-0.1%	-0.1%	0.0%
QIS5 TC ratio = (SCRt + RMt)/VZt	-0.3%	0.3%	0.1%	0.0%	0.0%	-0.1%	-0.1%	-0.1%	0.0%
(x,n)=(40,20)									
100	235.6%	117.2%	79.3%	60.1%	48.5%	40.7%	22.7%	13.4%	8.1%
500	105.2%	52.6%	35.5%	26.9%	21.7%	18.2%	10.1%	6.0%	3.6%
1'000	74.3%	37.4%	25.2%	19.0%	15.3%	12.8%	7.1%	4.2%	2.5%
10'000	23.3%	12.1%	8.1%	6.0%	4.8%	4.0%	2.2%	1.3%	0.8%
100'000	7.1%	4.1%	2.7%	1.9%	1.5%	1.2%	0.7%	0.4%	0.2%
Limiting QIS5 ratio SCRt/VZt	-0.3%	0.4%	0.2%	0.0%	0.0%	-0.1%	-0.1%	0.0%	0.0%
QIS5 TC ratio = (SCRt + RMt)/VZt	-0.7%	0.2%	0.0%	-0.1%	-0.1%	-0.1%	-0.1%	0.0%	0.0%

Table 3. VaR solvency capital ratios for the endowment with waiver of premium

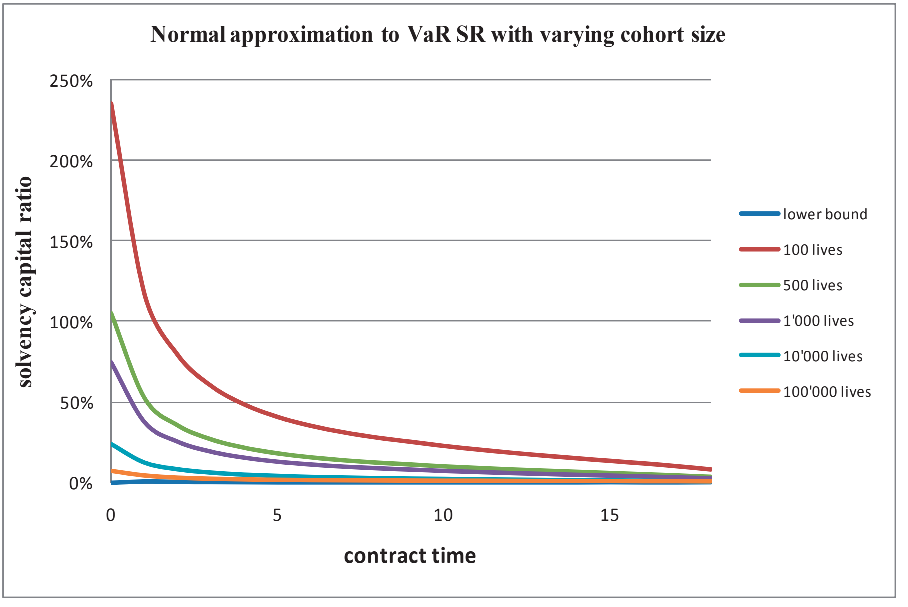


Figure 2. Time evolution of VaR solvency capital ratios, $(x,n)=(40,20)$

8. Conclusions and outlook

Let us summarize the present work. We have derived a general solvency rule for the prospective liability, which has resulted in two simple liability VaR & CVaR target capital requirements. The proposed approach has been applied to determine the biometric solvency risk capital for a portfolio of general traditional life contracts within the Markov chain model of life insurance. Our main actuarial tools have been the backward recursive actuarial reserve formula and the theorem of Hattendorff. Based on this we have determined the conditional mean and variance of a portfolio's prospective liability risk and have used a gamma approximation to obtain the liability VaR & CVaR solvency capital. Since our first formulas include only the process risk and do not take into account the possibility of systematic risk, we have proposed either to shift the biometric transition probabilities, or apply a stochastic model, which allows for random biometric transition probabilities.

Similarly to [1], Section 8, the adopted general methodology is in agreement with several known facts as (i) the process risk is negligible for portfolios with increasing size and has a small impact on medium to large insurers (ii) all else equal, process risk will increase (decrease) with higher (lower) coefficients of variation (aggregated effect of both decrement rates and sums at risk). Another interesting observation has been made at the end of Section 6.3 that the model with shifted biometric transitions can be viewed as a sub-model of the model with Poisson-Gamma time dependent biometric transitions.

Moreover, a detailed analysis for a single cohort of identical endowment contracts with waiver of premium by disability has been undertaken in Section 7. Besides a complete Markov chain specification, which seems to be missing in the literature, the numerical illustration has shown, as expected, that the cohort size is a main driving factor of process risk. Due to the statistical law of large numbers, the larger the cohort size the less solvency capital is actually required. In contrast to the life annuity "longevity risk" study in [1], the stochastic approach penalizes almost all insurers (except the very large ones) because the current standard approach prescribes almost negligible solvency capital ratios and does not measure explicitly the process risk effects.

The interested actuary might challenge the proposed approach with alternatives from other regulatory environments than Solvency II. Moreover, it is important to point out that a lot of technical issues remain to be settled properly. They are not only regulatory specific but also related to the complex mathematics of related software products and go beyond the Markov chain model. Today's life insurance contracts include many embedded options and are henceforth even more complex. A challenging issue is the definition of capital requirements for unit-linked contracts without and with guarantee and variable annuities with guaranteed minimum benefits (so-called variable GMXB annuities).

Author details

Werner Hürlimann
Wolters Kluwer Financial Services / FRSGlobal, Switzerland

9. References

[1] Hürlimann, W. Biometric solvency risk for portfolios of general life contracts (I) The single life multiple decrement case. North Amer. Actuarial J. 2010; 14(4): 400-419.
[2] QIS5 (2010). Technical Specifications QIS5 – CEIOPS Quantitative Impact Study 5. 2010 July 5. Available from: http://www.ceiops.org.
[3] Amsler, M.-H. Sur la modélisation des risques vie par les chaînes de Markov.Trans. 23rd Int. Congress of Actuaries. 1988; vol. 3: 1-17.
[4] Hoem, J.M. Markov chain models in life insurance. Blätter DGVFM. 1969; 9: 91-107.
[5] Hoem, J.M. The versatility of the Markov chain as a tool in the mathematics of life Insurance. Trans. 23rd Int. Congress of Actuaries. 1988; vol. 3: 171-202.
[6] Koller, M. Stochastische Modelle in der Lebensversicherung. Berlin: Springer; 2000.

[7] Milbrodt, H., Helbig, M. Mathematische Methoden der Personenversicherung. Berlin: De Gruyter; 1999.
[8] Wolthuis, H. Life Insurance Mathematics (The Markovian Model). Brüssel: CAIRE; 1994.
[9] Norberg, R. Thiele differential equation, Springer Online Reference Works. 2001.Available from:
http://eom.springer.de/T/t130080.htm.
[10] Gerber, H.U. Lebensversicherungsmathematik. Berlin: Springer; 1986.
[11] Gerber, H.U. Life Insurance Mathematics. 3rd ed. Berlin: Springer; 1997.
[12] Norberg, R. Life insurance mathematics, In: Teugels, J., Sundt, B., editors. Encyclopedia of Actuarial Science. New York: J. Wiley; 2004.
[13] Bühlmann, H. A probabilistic approach to long-term insurance (typically life insurance). Trans. 20th Int. Congress of Actuaries. 1976; vol. 4: 267-276.
[14] Gerber, H.U. An Introduction to Mathematical Risk Theory. Huebner Foundation Monograph, vol. 8. Homewood, Illinois: R.D. Irwin; 1979.
[15] Gerber, H.U; Leung, B.P.K., Shiu, E.S.W. Indicator function and Hattendorff theorem. North Amer. Actuarial J. 2003; 7(1): 38-47.
[16] Hattendorff, K. Das Risiko bei der Lebensversicherung. Maisius' Rundschau der Versicherungen. 1868; 18: 169-183.
[17] Kremer, E. Stochastic Life-Insurance Mathematics. Lecture Notes No. 5 – Special Issue. Hamburg: Verein Förderung angew. Math. Statistik & Risikotheorie; 1979.
[18] Patatriandafylou, A., Waters, H.R. Martingales in life insurance. Scandinavian Actuarial Journal. 1984; 2: 210-230.
[19] Bowers, N.L.; Gerber, H.U.; Hickman, J.C.; Jones, D.A., Nesbitt, C.J. Actuarial Mathematics. Itascal, IL: Society of Actuaries; 1986.
[20] Frasier, W. M. Second to Die Joint Life Cash Values and Reserves. 1978; The Actuary: 4.
[21] Second to Die. Letters. 1978 June; The Actuary: 3.
[22] Margus, P. Generalized Frasier claim rates under survivorship life insurance policies.North Amer. Actuarial J. 2002; 6(2): 76-94.
[23] Olivieri, A., Pitacco, E. Stochastic mortality: the impact on target capital.ASTIN Bulletin. 2009; 39(2): 541-563.
[24] Christiansen, M.; Denuit, M., Lazar, D. The Solvency II square-root formula for systematic biometric risk. Insurance: Math. and Economics. 2012; 50(2): 257-265.
[25] Chuard, C. (2008). Zur Geschichte der Sterbetafeln. Vortrag an der GV der Schweiz. Kammer der Pensionskassenexperten, Bern. 2008 April 23. Available from:
www.chuard.com/new/d/wpcontent/uploads/2009/03/hist_sterbetafeln.pdf
[26] Saxer, W. Versicherungsmathematik I. Series of Comprehensive Studies in Mathematics, vol. 79 (Reprint 1979). Berlin: Springer; 1955.
[27] Christiansen, M. A sensitivity analysis of typical life insurance contracts with respect to the technical basis. Insurance: Math. and Economics. 2008; 42(2): 787-796.

[28] Christiansen, M. Biometric worst-case scenarios for multi-state life insurance policies. Insurance: Math. and Economics. 2010; 47(2): 190-197.

Chapter 2

Boundary-Value Problems for Second Order PDEs Arising in Risk Management and Cellular Neural Networks Approach

Rossella Agliardi, Petar Popivanov and Angela Slavova

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/49936

1. Introduction

This work deals with the Dirichlet problem for some PDEs of second order with non-negative characteristic form. One main motivation is to study some boundary-value problems for PDEs of Black-Scholes type arising in the pricing problem for financial options of barrier type. Barrier options on stocks have been traded since the end of the Sixties and the market for these options has been dramatically expanding, making barrier options the most popular ones among the exotic. The class of standard barrier options includes 'in' barriers and 'out' barriers, which are activated (knocked in) and, respectively, extinguished (knocked out) if the underlying asset price crosses the barrier before expiration. Moreover, each class includes 'down' or 'up' options, depending on whether the barrier is below or above the current asset price and thus can be breached from above or below. Therefore there are eight types of standard barrier options, depending on their 'in' or 'out', 'down' or 'up', and 'call' or 'put' attributes. It is possible to include a cash rebate, which is paid out at option expiration if an 'in' ('out') option has not been knocked in (has been knocked out, respectively) during its lifetime. One can consider barrier options with rebates of several types, terminal payoffs of different forms (e.g. power options), more than one underlying assets and/or barriers, and allow for time-dependent barriers, thus enriching this class still further. On the other hand, a large variety of new exotic barriers have been designed to accommodate investors' preferences. Another motivation for the study of such options is related to credit risk theory. Several credit-risk models build on the barrier option formalism, since the default event can be modeled throughout a signalling variable hitting a pre-specified boundary value (See [3],[8] among others). As a consequence, a substantial body of academic literature provides pricing methods for valuating barrier options, starting from the seminal work of [18], where an exact formula is offered for a down-and-out European call with zero rebate. Further extensions are provided - among others - in [22] for the different types of standard barrier options, in [16] for simultaneous ’down’ and ’up’ barriers with exponential dependence on time, in [10] for two boundaries via Laplace transform, in [12] and [7] for partial barrier and rainbow options, in [17] for multi-asset options with an outside barrier, in [5] in a most comprehensive setting employing the image solution method. Many analytical formulas for barrier options are collected also in handbooks (see [11], for example).

For analytical tractability most literature assumes that the barrier hitting is monitored in continuous time. However there exist some works dealing with the discrete version, i.e. barrier crossing is allowed only at some specific dates -typically at daily closings. (See [1] and [15], for a survey). Furthermore, a recent literature relaxes the Brownian motion assumption and considers a more general Lévy framework. For example, [4] study barrier options of European type assuming that the returns of the underlying asset follows a Lévy process from a wide class. They employ the Wiener-Hopf factorization method and elements of pseudodifferential calculus to solve the related boundary problem. This book chapter adopts a classical Black-Scholes framework. The problem of pricing barrier options is reducible to boundary value problems for a PDE of Black-Scholes type and with pre-specified boundaries. The value at the terminal time T is assigned, specifying the terminal payoff which is paid provided that an ’in’ option is knocked in or an ’out’ option is not knocked out during its lifetime. The option holder may be entitled or not to a rebate. From a mathematical point of view, the boundary condition can be inhomogeneous or homogeneous. While there are several types of barrier options, in this work we will focus on ’up’ barriers in view of the relationships between the prices of different types of vanilla options (see [25]). Moreover, the case of floating barriers of exponential form can be easily accommodated by substitution of the relevant parameters (see [25], Chapter 11), thus we confine ourselves to the case of constant barriers. On the other hand, we work within a general framework that allows for multi-asset options, a generic payoff and rebate. Furthermore, we tackle some regularity questions and the problem of existence of generalized solutions. In Section 2 the (initial) boundary value problem is studied in a multidimensional framework generalizing the Black-Scholes equation and analytical solutions are obtained, while a comparison principle is provided in Section 4. Section 3 presents some applications in Finance: our general setting incorporates several known pricing expressions and, at the same time, allows to generate new valuation formulas. Section 5 and the Appendix study the existence and regularity of generalized solutions to the boundary value problems for a class of PDEs incorporating the Black-Scholes type. We build on the approach of Oleinik and Radkevič and adapt the method to the PDEs of interest in the financial applications.

2. Generalizations of the Black-Scholes equation in the multidimensional case: (initial) boundary value problems

Consider in $\mathbf{R}^1_t \times \mathbf{R}^n_x$ the following generalization of the Black-Scholes equation:

$$Lu = u_t + \sum_{i,j=1}^{n} a_{ij} x_i x_j u_{x_i x_j} + \sum_{i=1}^{n} b_i x_i u_{x_i} + cu = f(t,x), \tag{1}$$

where $0 \leq t \leq T$ and $x_j \geq 0, 1 \leq j \leq n$.

This is the Cauchy problem:

$$\begin{cases} Lu = f(t,x), \\ u|_{t=T} = u_0(x) \end{cases}, \tag{2}$$

$$\begin{cases} x_j \geq 0, \quad 1 \leq j \leq n \\ 0 \leq t \leq T \end{cases}$$

and this is the boundary value problem:

$$\begin{cases} Lu = f \\ u|_{t=T} = u_0(x) \\ u|_{x_j=a_j} = g_j(t,x)|_{x_j=a_j}, 1 \leq j \leq n, \end{cases} \tag{3}$$

$$\begin{cases} 0 \leq t \leq T \\ 0 \leq x_j \leq a_j,\ a_j > 0 \\ 1 \leq j \leq n \end{cases}$$

In (1) $a_{ij} = a_{ji} = const, b_i = const, c = const$ and

$$\sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j \geq c_0|\xi|^2, c_0 = const > 0. \tag{4}$$

Our first step is to make in the non-hypoelliptic PDE L the change of the space variables:

$$y_j = ln x_j, 1 \leq j \leq n, \tau = T - t \Rightarrow \frac{\partial u}{\partial t} = -\frac{\partial u}{\partial \tau}, y_j \in \mathbf{R}^1 \tag{5}$$

$\frac{\partial u}{\partial x_i} = e^{-y_i}\frac{\partial u}{\partial y_i}, \frac{\partial^2 u}{\partial x_i \partial x_j} = e^{-y_i - y_j}[\frac{\partial^2 u}{\partial y_i \partial y_j} - \delta_{ij}\frac{\partial u}{\partial y_i}]$, δ_{ij} being the Kronecker symbol.

Thus, (1) takes the form:

$$\tilde{L}u = -\frac{\partial u}{\partial \tau} + \sum_{i,j=1}^{n} a_{ij}\frac{\partial^2 u}{\partial y_i \partial y_j} + \sum_{i=1}^{n}\frac{\partial u}{\partial y_i}(b_i - a_{ii}) + cu = f, \tag{6}$$

i.e.

$$\frac{\partial u}{\partial \tau} = \sum_{i,j=1}^{n} a_{ij}\frac{\partial^2 u}{\partial y_i \partial y_j} + \sum_{i=1}^{n}\tilde{b}_i\frac{\partial u}{\partial y_i} + cu - f; \tilde{b}_i = b_i - a_{ii}$$

In the case (2) we have

$$\begin{cases} \tilde{L}u = f, & 0 \leq \tau \leq T \\ u|_{\tau=0} = \widetilde{u}_0(y) = u_0(e^{y_1}, \ldots e^{y_n}), \ y \in \mathbf{R}^n, \end{cases} \tag{7}$$

while in the case (3)

$$\begin{cases} \tilde{L}u = f, & 0 \leq \tau \leq T \\ u|_{\tau=0} = \widetilde{u}_0(y) \\ u|_{y_j=\tilde{a}_j} = g_j|_{y_j=\tilde{a}_j} \end{cases} \tag{8}$$

Denote $D = \{0 \leq \tau \leq T, -\infty < y_j \leq lna_j = \tilde{a}_j, 1 \leq j \leq n\}$, $x_j = e^{y_j}$, $1 \leq j \leq n \Rightarrow f(t,x) = f(T-\tau, e^{y_1}, \ldots, e^{y_n})$.

In (6) we make the change of the unknown function $u : u = v(\tau, y)e^{\sum \alpha_i y_i + \beta\tau}$ in $(\tau, y) \in D$. Thus, after standard computations we get:

$$v_\tau + \beta v = \sum_{i,j} a_{ij} v_{y_i y_j} + \sum_{i,j} a_{ij}(\alpha_i v_{y_j} + \alpha_j v_{y_i}) + \tag{9}$$

$$+ \sum_{i=1}^{n} \tilde{b}_i v_{y_i} + \sum_{i,j=1}^{n} a_{ij}\alpha_i\alpha_j v + \sum_{i=1}^{n} \tilde{b}_i \alpha_i v + cv - fe^{-\sum \alpha_i y_i - \beta\tau}.$$

Let us take

$$\beta = \sum_{i,j} a_{ij}\alpha_i\alpha_j + \sum_i \tilde{b}_i \alpha_i + c \tag{10}$$

and put $f_1 = -fe^{-\sum_i \alpha_i y_i - \beta\tau}$. Put $A = (a_{ij})_{i,j=1}^n$, $A^* = A$, $\alpha = (\alpha_1, \ldots, \alpha_n)$. Then the scalar product $(A\alpha, \nabla_y v) = \sum_{i,j} a_{ij}\alpha_j \frac{\partial v}{\partial y_i} = \sum_{i,j} a_{ji}\alpha_i \frac{\partial v}{\partial y_j} = \sum_{i,j} a_{ij}\alpha_i \frac{\partial v}{\partial y_j}$, i.e. we assume that

$$2(A\alpha, \nabla_y v) + (\tilde{b}, \nabla_y v) = 0 \iff \tag{11}$$

$$2A\alpha + \tilde{b} = 0,$$

where $\tilde{b} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ is given, $det A \neq 0$.

In conclusion we solve the algebraic system (11): $\alpha = -\frac{1}{2}A^{-1}(\tilde{b})$ and then we define β by (10). This way (9) takes the form:

$$v_\tau = \sum_{i,j=1}^{n} a_{ij} v_{y_i y_j} + f_1(\tau, y) \tag{12}$$

The Cauchy problem (12) has initial condition

$$v_0(y) = v|_{\tau=0} = \tilde{u}_0(y)e^{-\sum_i \alpha_i y_i}; \tilde{u}_0 \equiv u_0(e^{y_1}, \ldots, e^{y_n}), y \in \mathbf{R}^n.$$

To find a formula (Poisson type) for the solution of the Cauchy problem (12), $v|_{\tau=0} = v_0(y)$ we must use some auxiliary results from the linear algebra. So let $Mu = \sum_{i,j=1}^n a_{ij} v_{y_i y_j}$. Then the change of the independent variables $y = Bz \iff z = B^{-1}y$, $B^{-1} = (\beta_{li})_{l,i=1}^n$ leads to $\frac{\partial^2}{\partial y_i \partial y_j} = \sum_{k,l=1}^n \beta_{li}\beta_{kj} \frac{\partial^2}{\partial z_k \partial z_l}$, i.e.

$$Mu = \sum_{k,l=1}^{n} (\sum_i (\sum_j a_{ij}\beta_{kj})\beta_{li}) \frac{\partial^2 u}{\partial z_k \partial z_l}.$$

One can easily guess that $\sum_i(\sum_j a_{ij}\beta_{kj})\beta_{li} = \tilde{c}_{kl}$ are the elements of the matrix $B^{-1}A(B^{-1})^*$ and of course $(B^{-1})^* = (B^*)^{-1}$. On the other hand consider the elliptic quadratic form $(Ax, x) = (C^*ACy, y)$ after the nondegenerate change $x = Cy$. As we know one can find such a matrix C that

$$C^*AC = I_n, \tag{13}$$

I_n being the unit matrix. Put now $C = (B^{-1})^* \Rightarrow C^* = B^{-1}$. Then $C^*AC = I_n \Rightarrow B^{-1}A(B^{-1})^* = I_n \Rightarrow Mu = \sum_{k=1}^{n} \frac{\partial^2 u}{\partial z_k^2}$.

This way the change $y = (C^{-1})^* z \Rightarrow z = B^{-1}y$ transforms the Cauchy problem (12) to:

$$\begin{cases} \frac{\partial v}{\partial \tau} = \sum_{k=1}^{n} \frac{\partial^2 v}{\partial z_k^2} + \tilde{f}_1(\tau, z), & 0 \leq \tau \leq T \\ v|_{\tau=0} = v_0((C^{-1})^* z) \equiv \tilde{v}_0(z), \; z \in \mathbf{R}^n. \end{cases} \tag{14}$$

The solution of the Cauchy problem (14) is given by the formula

$$v(\tau, z) = \frac{1}{(2\sqrt{\pi\tau})^n} \int_{\mathbf{R}^n} \tilde{v}_0(\lambda) e^{-\frac{|z-\lambda|^2}{4\tau}} d\lambda + \tag{15}$$

$$+ \int_{\mathbf{R}^n} \int_0^{\tau} \frac{\tilde{f}_1(\Theta, \lambda)}{[2\sqrt{\pi(\tau - \Theta)}]^n} e^{-\frac{|z-\lambda|^2}{4(\tau-\Theta)}} d\lambda d\Theta,$$

$z \in \mathbf{R}^n$, $\lambda \in \mathbf{R}^n \Rightarrow |z - \lambda|^2 = \sum_{i=1}^{n} (z_i - \lambda_i)^2$ (see [6] or [21]).

Going back to the old coordinates (τ, x) and the old function $u = ve^{\sum \alpha_i y_i + \beta\tau}$, we find $u(t, x)$-the solution of (2); $t = T - \tau$, $y_j = lnx_j$, $z = B^{-1}y = B^{-1}(lnx_1, \ldots, lnx_n)$; $u = vx_1^{\alpha_1} \ldots x_n^{\alpha_n} e^{\beta(T-t)}$.

We shall concentrate now on (3), $n = 2$.

Remark 1. To simplify the things, consider the quadratic form (elliptic) $Q = a_{11}\xi^2 + 2a_{12}\xi\eta + a_{22}\eta^2$, $a_{11} > 0$, $a_{22} > 0$, $a_{12}^2 - a_{11}a_{22} < 0$, $Q = (A\begin{pmatrix} \xi \\ \eta \end{pmatrix}, \begin{pmatrix} \xi \\ \eta \end{pmatrix})$.

Then $Q = \frac{1}{a_{11}}(a_{11}\xi + a_{12}\eta)^2 + b\eta^2$; $b = a_{22} - \frac{a_{12}^2}{a_{11}} > 0$. The change

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{a_{11}} & \frac{a_{12}}{\sqrt{a_{11}}} \\ 0 & \sqrt{b} \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \tag{16}$$

leads to $Q = x^2 + y^2$. Moreover, the first quadrant $\xi \geq 0, \eta \geq 0$ is transformed under the linear transformation with matrix $D = \begin{pmatrix} \sqrt{a_{11}} & \frac{a_{12}}{\sqrt{a_{11}}} \\ 0 & \sqrt{b} \end{pmatrix}$, $D^{-1} = \begin{pmatrix} \frac{1}{\sqrt{a_{11}}} & -\frac{a_{12}}{a_{11}\sqrt{b}} \\ 0 & \frac{1}{\sqrt{b}} \end{pmatrix}$ into angle between the rays (straight lines) $l_1 : \begin{vmatrix} x \geq 0 \\ y = 0 \end{vmatrix}$ and $l_2 : \begin{vmatrix} x = \frac{a_{12}}{\sqrt{a_{11}}}\eta \\ y = \sqrt{b}\eta \geq 0 \end{vmatrix}$ with opening φ_0. Evidently, $(D^{-1})^* A D^{-1} = I_2$.

Consequently, the transformation D is not orthogonal for $a_{12} \neq 0$.

Let us now consider the boundary value problem (8). The above-proposed procedure yields:

$$\begin{cases} v_\tau = \sum_{i,j=1}^{2} a_{ij} v_{y_i y_j} + f_1(\tau, y) \\ v|_{\tau=0} = v_0(y) = u_0(e^{y_1}, e^{y_2}) e^{-\sum \alpha_i y_i} \\ v|_{y_1=\tilde{a}_1} = g_1(T-\tau, e^{y_1}, e^{y_2})|_{y_1=\tilde{a}_1} e^{-\beta\tau} a_1^{-\alpha_1} e^{-\alpha_2 y_2} \equiv \tilde{g}_1(\tau, y_2) \\ v|_{y_2=\tilde{a}_2} = g_2(T-\tau, e^{y_1}, e^{y_2})|_{y_2=\tilde{a}_2} e^{-\beta\tau} a_2^{-\alpha_2} e^{-\alpha_1 y_1} \equiv \tilde{g}_2(\tau, y_1) \end{cases} \tag{17}$$

$$\begin{cases} -\infty < y_j < ln a_j = \tilde{a}_j \\ 0 \le \tau \le T \end{cases}$$

The change $\begin{cases} \lambda_j = \tilde{a}_j - y_j \ge 0, j = 1,2 \\ \tau = \tau \end{cases}$ in (17) yields:

$$\begin{cases} \tilde{v}_\tau = \sum_{i,j=1}^{2} a_{ij} \tilde{v}_{\lambda_i \lambda_j} + \tilde{f}_1(\tau, \lambda) \\ \tilde{v}|_{\tau=0} = \tilde{v}_0(\lambda) = v_0(\tilde{a}_1 - \lambda_1, \tilde{a}_2 - \lambda_2) e^{-\sum_{i=1}^{2} \alpha_i (\tilde{a}_i - \lambda_i)} \\ \tilde{v}|_{\lambda_1=0} = \tilde{g}_1(\tau, \tilde{a}_2 - \lambda_2) \\ \tilde{v}|_{\lambda_2=0} = \tilde{g}_2(\tau, \tilde{a}_1 - \lambda_1), \end{cases} \tag{18}$$

$\Omega = \{0 \le \tau \le T, \lambda_j \ge 0, j = 1,2\}$, Ω is a wedge with opening $\frac{\pi}{2}$.

Now we use the linear transformation described in Remark 1, that maps the first quadrant $\lambda_1 \ge 0, \lambda_2 \ge 0$ onto the angle between the rays l_1 and l_2 in the plane $0_{z_1 z_2}$ and we obtain:

$$\begin{cases} w_\tau = w_{z_1 z_1} + w_{z_2 z_2} + f(\tau, z) \\ w|_{\tau=0} = w_0(z) \\ w|_{z_1=0} = \tilde{g}_1(\tau, z_1), (\tau, z) \in \tilde{\Omega} \\ w|_{l_2} = \tilde{g}_2(\tau, z_1, z_2)|_{(z_1,z_2)} \in l_2, \end{cases} \tag{19}$$

$l_1 : \begin{cases} z_1 = 0 \\ z_2 = \frac{\lambda_2}{\sqrt{b}} \end{cases}$, $l_2 : \begin{cases} z_1 = \frac{\lambda_1}{\sqrt{a_{11}}} \\ z_2 = \frac{-a_{12}}{\sqrt{b a_{11}}} \lambda_1 \end{cases}$, $\tilde{\Omega}$ is a wedge with opening φ_0, i.e. $\tilde{\Omega} = [0, T] \times \Gamma$, Γ being the interior of the angle between l_1, l_2.

In fact, $\lambda = Bz \iff z = B^{-1}\lambda$ and $B^{-1}A(B^{-1})^* = I_2$ implies that $\sum_{i,j=1}^{2} a_{ij} \frac{\partial^2}{\partial\lambda_i \partial\lambda_j}$ is transformed in $\frac{\partial^2}{\partial z_1^2} + \frac{\partial^2}{\partial z_2^2}$. According to Remark 1: $(D^{-1})^* A D^{-1} = I_2$. Taking $B^{-1} = (D^{-1})^*$, i.e. $B = D^*$ we obtain that $\{\lambda_1 \ge 0, \lambda_2 \ge 0\}$ is mapped onto the angle φ_0 between the rays l_1, l_2. Of course, there are three possibilities: $\varphi_0 = \frac{\pi}{2}, 0 < \varphi_0 < \frac{\pi}{2}, \frac{\pi}{2} < \varphi_0 < \pi$.

From now on we shall make polar coordinates change in (19): $\begin{cases} z_1 = r cos\varphi \\ z_2 = r sin\varphi \end{cases}$ and to fix the ideas let $0 < \varphi_0 < \frac{\pi}{2}$, $\begin{cases} r \ge 0 \\ \frac{\pi}{2} - \varphi_0 \le \varphi \le \frac{\pi}{2} \end{cases}$, φ_0 is the angle between l_2 and l_1.

The new change $\Phi = \varphi - (\frac{\pi}{2} - \varphi_0) \Rightarrow 0 \le \Phi \le \varphi_0$ and $\frac{\partial}{\partial\Phi} = \frac{\partial}{\partial\varphi}$. To simplify the notation we shall write again (r, φ) instead of (r, Φ), $0 \le \Phi \le \varphi_0$. Thus we have a wedge type initial-boundary value problem for (19) with unknown function $w(\tau, r, \varphi)$:

$$\begin{cases} w_\tau = \frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{1}{r^2}\frac{\partial^2 w}{\partial \varphi^2} + f(\tau, r, \varphi) \\ w|_{\tau=0} = w_0(r, \varphi) \\ w|_{\varphi=0} = \tilde{\tilde{g}}_1(\tau, r) \\ w|_{\varphi=\varphi_0} = \tilde{\tilde{g}}_2(\tau, r) \end{cases} \tag{20}$$

$r \geq 0$, $0 \leq \varphi \leq \varphi_0$, $l_1 : \{\varphi = 0, r \geq 0\}$, $l_2 : \{\varphi = \varphi_0, r \geq 0\}$, $r \leftrightarrow \xi$, $\varphi \leftrightarrow \eta$, $0 \leq \Theta \leq \tau$, $0 \leq \xi \leq \infty$, $0 \leq \eta \leq \varphi_0$, $0 < \varphi_0 < \pi$. Then

$$\begin{aligned} w(\tau, r, \varphi) = \int_0^\tau \int_0^{\varphi_0} \int_0^\infty f(\Theta, \xi, \eta) G(r, \varphi, \xi, \eta, \tau - \Theta) \xi d\xi d\eta d\Theta + \\ + \int_0^\tau \int_0^\infty \tilde{\tilde{g}}_1(\Theta, \xi) \frac{1}{\xi} [\frac{\partial}{\partial \eta} G(r, \varphi, \xi, \eta, \tau - \Theta)]_{\eta=0} d\xi d\Theta - \\ - \int_0^\tau \int_0^\infty \tilde{\tilde{g}}_2(\Theta, \xi) \frac{1}{\xi} [\frac{\partial}{\partial \eta} G(r, \varphi, \xi, \eta, \tau - \Theta)]_{\eta=\varphi_0} d\xi d\Theta + \\ + \int_0^{\varphi_0} \int_0^\infty w_0(\xi, \eta) G(r, \varphi, \xi, \eta, \tau) \xi d\xi d\eta, \end{aligned} \tag{21}$$

where $G(r, \varphi, \xi, \eta, \tau) = \frac{1}{\varphi_0 \tau} e^{-\frac{(r^2+\xi^2)}{4\tau}} \sum_{n=1}^\infty I_{\frac{n\pi}{\varphi_0}}(\frac{r\xi}{2\tau}) sin \frac{n\pi}{\varphi_0} \varphi sin \frac{n\pi}{\varphi_0} \eta$ and the modified Bessel function $w = I_\nu(z)$ satisfies the equation:

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} - (z^2 + \nu^2) w = 0, \nu \geq 0, \tag{22}$$

$I_\nu(z) = \sum_{m=0}^\infty \frac{(\frac{z}{2})^{2m+\nu}}{m! \Gamma(m+\nu+1)}$ (see [2]).

Remark 2. One can see that $lim_{\tau \to +0} \int_0^{\varphi_0} \int_0^\infty w_0(\xi, \eta) G(r, \varphi, \xi, \eta, \tau) \xi d\xi d\eta = w_0(r, \varphi)$, i.e. formally $lim_{\tau \to +0} \xi G(r, \varphi, \xi, \eta, \tau) = \delta(r - \xi, \varphi - \eta)$ in the sense of Schwartz distributions $D^{'}(\mathbf{R}^1_+ \times [0, \varphi_0])$, $\mathbf{R}_+ = \{\xi \geq 0\}$. G is the corresponding Green function.

Formula (21) is given in [21], pages 182 and 166 or in [6], pp.498. The proof of (21) is based on the properties of the Bessel functions and Hankel transform.

Remark 3. In the special case when $a_{12} = 0$ in (16) we obtain (18) and after the change $\tau = \tau$, $\lambda_j = \sqrt{a_{jj}} z_j$, $1 \leq j \leq 2$ (18) takes the form:

$$\left| \begin{array}{l} \frac{\partial \tilde{v}}{\partial \tau} = \frac{\partial^2 \tilde{v}}{\partial z_1^2} + \frac{\partial^2 \tilde{v}}{\partial z_2^2} + \tilde{\tilde{f}}_1(\tau, z) \\ \tilde{v}|_{\tau=0} = \tilde{v}_0(z) \\ \tilde{v}|_{z_1=0} = \tilde{\tilde{g}}_1(\tau, z_2) \\ \tilde{v}|_{z_2=0} = \tilde{\tilde{g}}_2(\tau, z_1) \end{array} \right. \tag{23}$$

$0 \leq \tau \leq T, z_j \geq 0, 1 \leq j \leq 2$. Certainly, $\varphi_0 = \frac{\pi}{2}$.

According to [21]:

$$\tilde{\tilde{v}}(\tau,z) = \int_0^{\tau}\int_0^{\infty}\int_0^{\infty} \tilde{\tilde{f}}_1(\Theta,\xi,\eta)G(\tau-\Theta,z_1,z_2,\xi,\eta))d\xi d\eta d\Theta + \tag{24}$$

$$+\int_0^{\infty}\int_0^{\infty} \tilde{\tilde{v}}_0(\xi,\eta)G(\tau,z_1,z_2,\xi,\eta)d\xi d\eta +$$

$$+\int_0^{\tau}\int_0^{\infty} \tilde{\tilde{g}}_1(\Theta,\eta)[\frac{\partial}{\partial\xi}G(\tau-\Theta,z_1,z_2,\xi,\eta)]_{\xi=0}d\eta d\Theta +$$

$$+\int_0^{\tau}\int_0^{\infty} \tilde{\tilde{g}}_2(\Theta,\xi)[\frac{\partial}{\partial\eta}G(\tau-\Theta,z_1,z_2,\xi,\eta)]_{\eta=0}d\xi d\Theta,$$

where the Green function $G(\tau,z_1,z_2,\xi,\eta) = \frac{1}{4\pi\tau}[e^{-\frac{(z_1-\xi)^2}{4\tau}} - e^{-\frac{(z_1+\xi)^2}{4\tau}}] \times [e^{-\frac{(z_2-\eta)^2}{4\tau}} - e^{-\frac{(z_2+\eta)^2}{4\tau}}]$.

3. Applications to financial options and numerical results via CNN

Here the analysis of Section 2 is applied to some problems arising in option pricing theory. Some known pricing formulas are revisited in a more general setting and some new results are offered. We apply Cellular Neural Networks (CNN) approach [24] in order to obtain some numerical results. Let us consider a two-dimensional grid with 3×3 neighborhood system as it is shown on Figure 1.

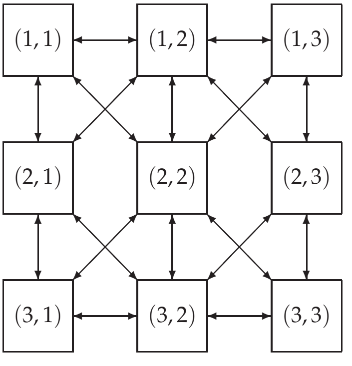


Figure 1. 3×3 neighborhood CNN.

[htb] One of the key features of a CNN is that the individual cells are nonlinear dynamical systems, but that the coupling between them is linear. Roughly speaking, one could say that these arrays are nonlinear but have a linear spatial structure, which makes the use of techniques for their investigation common in engineering or physics attractive.

We will give the general definition of a CNN which follows the original one:

Definition 1. *The CNN is a*

a). 2-, 3-, or n- dimensional array of

b). mainly identical dynamical systems, called cells, which satisfies two properties:

c). most interactions are local within a finite radius r, and

d). all state variables are continuous valued signals.

Definition 2. *An $M \times M$ cellular neural network is defined mathematically by four specifications:*

1). CNN cell dynamics;

2). CNN synaptic law which represents the interactions (spatial coupling) within the neighbor cells;

3). Boundary conditions;

4). Initial conditions.

Now in terms of definition 2 we can present the dynamical systems describing CNNs. For a general CNN whose cells are made of time-invariant circuit elements, each cell $C(ij)$ is characterized by its CNN cell dynamics :

$$\dot{x}_{ij} = -g(x_{ij}, u_{ij}, I_{ij}^s), \tag{25}$$

where $x_{ij} \in \mathbf{R}^m$, u_{ij} is usually a scalar. In most cases, the interactions (spatial coupling) with the neighbor cell $C(i+k, j+l)$ are specified by a CNN synaptic law:

$$\begin{aligned} I_{ij}^s = {} & A_{ij,kl} x_{i+k,j+l} + \\ & + \tilde{A}_{ij,kl} * f_{kl}(x_{ij}, x_{i+k,j+l}) + \\ & + \tilde{B}_{ij,kl} * u_{i+k,j+l}(t). \end{aligned} \tag{26}$$

The first term $A_{ij,kl} x_{i+k,j+l}$ of (26) is simply a linear feedback of the states of the neighborhood nodes. The second term provides an arbitrary nonlinear coupling, and the third term accounts for the contributions from the external inputs of each neighbor cell that is located in the N_r neighborhood.

It is known [24] that some autonomous CNNs represent an excellent approximation to nonlinear partial differential equations (PDEs). The intrinsic space distributed topology makes the CNN able to produce real-time solutions of nonlinear PDEs. There are several ways to approximate the Laplacian operator in discrete space by a CNN synaptic law with an appropriate A-template:

- one-dimensional discretized Laplacian template:

$$A_1 = (1, -2, 1),$$

- two-dimensional discretized Laplacian template:

$$A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Example 1 (Single-asset inside barrier options) The case of single-barrier zero-rebate down-and-out options was already priced in [18], while the case with rebate is found in [22]. A simple method for obtaining analytical formulas for barrier options is the reflection principle that has a long history in Physics and is commonly used in Finance. Here we write down the pricing formula for a general payoff and rebate and study its analytical properties. Let us consider the following boundary value problem:

$$\begin{cases} Lu = 0 \quad \text{in} \quad \Omega = \{(t,S); 0 < t < T, 0 < S < S^*\} \\ u|_{t=T} = u_0(S), 0 \le S \le S^* \\ u|_{S=S^*} = g(t), 0 \le t \le T \end{cases}$$

where $L = \partial_t + rS\partial_S + \frac{1}{2}\sigma^2 S^2 \partial_S^2 - r$, u_0 and g are continuous and $u_0(S^*) = g(T)$. Using the notation of Section 2 and taking $\alpha = \frac{1}{2} - \frac{r}{\sigma^2}$, $\beta = -r\left[\frac{1}{2} + \frac{r}{\sigma^2}\right]$, $C = \frac{\sqrt{2}}{\sigma}$ we straightforwardly obtain the following pricing formula (after changing to variables $\frac{\sigma}{\sqrt{2}}\lambda = \ln S^* - \xi$):

$$u(t,S) = \left(\frac{S}{S^*}\right)^{\alpha} \frac{e^{\beta(T-t)}}{\sqrt{2\pi}\sigma}[\frac{1}{\sqrt{(T-t)}}\int_0^{+\infty} u_0(S^* e^{-\xi})e^{\alpha\xi} \times \tag{27}$$

$$\times[\exp(-\frac{[\ln(S/S^*)+\xi]^2}{2\sigma^2(T-t)}) - \exp(-\frac{[\ln(S/S^*)-\xi]^2}{2\sigma^2(T-t)})]d\xi +$$

$$+\ln\frac{S^*}{S}\int_0^{T-t}\frac{g(T-s)}{(T-t-s)^{3/2}}e^{-\frac{\beta\sigma^2 s}{2}}\exp(-\frac{\ln^2(S/S^*)}{2\sigma^2(T-t-s)})ds]$$

Let us study the properties of $u(t,S)$ analytically. Without loss of generality we can assume $S^* = 1$ and therefore $e^{-\beta(T-t)}u(t,S) = \widetilde{u}(t,S)$ is written in the form $I_1 + I_2 + I_3$ with:

$$I_1(\tau,y) = \frac{-ye^{\alpha y}}{2\sqrt{\pi}}\int_0^{\tau}\frac{g(T-\frac{2\gamma}{\sigma^2})}{(\tau-\gamma)^{3/2}}e^{-\beta\gamma}\exp(-\frac{y^2}{4(\tau-\gamma)})d\gamma$$

$$I_2(\tau,y) = \frac{e^{\alpha y}}{2\sqrt{\pi\tau}}\int_0^{+\infty}u_0(e^{-\xi})e^{\alpha\xi}\exp(-\frac{[y+\xi]^2}{4\tau})d\xi$$

$$I_3(\tau,y) = -\frac{e^{\alpha y}}{2\sqrt{\pi\tau}}\int_0^{+\infty}u_0(e^{-\xi})e^{\alpha\xi}\exp(-\frac{[y-\xi]^2}{4\tau})d\xi$$

where $y = \ln S$ and $\tau = \frac{\sigma^2}{2}(T-t)$. We shall examine the asymptotics of $\widetilde{v}(\tau,y) = \widetilde{u}(t,S)$ for $0 < \tau < \frac{\sigma^2}{2}T$ (i.e. $0 < t < T$) fixed and for $y \longrightarrow -\infty$ (i.e. $S \to 0^+$). Put $h(\xi) = u_0(e^{-\xi})$, $\xi \ge 0$. Then:

$$I_2(\tau,y) = \frac{e^{\tau\alpha^2}}{2\sqrt{\pi\tau}}\int_0^{+\infty}h(\xi)\exp(-\frac{[y+\xi-2\alpha\tau]^2}{4\tau})d\xi = \frac{e^{\tau\alpha^2}}{\sqrt{\pi}}\int_{\frac{y-2\alpha\tau}{2\sqrt{\tau}}}^{+\infty}h(-y+2a\tau+2\eta\sqrt{\tau})e^{-\eta^2}d\eta.$$

According to Lebesgue's dominated convergence theorem, since $lim_{y\longrightarrow -\infty}h(-y+2a\tau+2\eta\sqrt{\tau}) = u_0(0)$ for each fixed η and τ, one has $lim_{y\longrightarrow -\infty}I_2(\tau,y) = e^{\tau\alpha^2}u_0(0)$. On the other hand:

$$|I_3(\tau,y)| \le \frac{const}{2\sqrt{\pi\tau}}\int_0^{+\infty}e^{\alpha(y+\xi)}(-\frac{(\xi-y)^2}{4\tau})d\xi =$$
$$=\text{const}.e^{\tau\alpha^2}\int_{\frac{-y+2\alpha\tau}{2\sqrt{\tau}}}^{+\infty}\exp[-\mu^2+2\alpha y-4\tau\alpha^2+2\alpha\mu\sqrt{\tau}]d\mu =$$
$$=\text{const}.e^{2\alpha y-2\tau\alpha^2}\int_{\frac{-y}{2\sqrt{\tau}}}^{+\infty}e^{-\varepsilon^2}d\varepsilon.$$

Thus, for fixed τ, $0 < \tau < \frac{\sigma^2}{2}T$, and $y << -1$, we have

$|I_3(\tau, y)| \leq \text{const.} e^{2\alpha y - 2\tau\alpha^2} \frac{\sqrt{\tau}}{-y} e^{-\frac{y^2}{4\tau}}$, which implies that $lim_{y \longrightarrow -\infty} I_3(\tau, y) = 0$. Finally, we observe that:

$|I_1(\tau, y)| \leq \frac{\max|g|}{2\sqrt{\pi}} |y| e^{\alpha y} \int_0^\tau \frac{e^{-\beta\tau}}{(\tau-\gamma)^{3/2}} \exp(-\frac{y^2}{4(\tau-\gamma)}) d\gamma$ as $\beta \leq 0$ implies $0 \leq -\beta\gamma \leq -\beta\tau$. The change $\theta = \frac{-y}{2\sqrt{\tau-\gamma}}$ yields

$|I_1(\tau, y)| \leq \text{const.} e^{\alpha y} \int_{\frac{-y}{2\sqrt{\tau}}}^{+\infty} e^{-\theta^2} d\theta$, that is

$|I_1(\tau, y)| \leq \text{const.} e^{\alpha y} \frac{2\sqrt{\tau}}{\sqrt{\pi}|y|} e^{-\frac{y^2}{4\tau}}$ for $y \longrightarrow -\infty$, τ fixed. Therefore we get:

$lim_{S \longrightarrow 0^+} u(t, S) = u_0(0) e^{-r(T-t)}, 0 < t < T.$

Remark 4. Assume that $u \in C^2(\overline{\Omega})$. Then, putting $S = 0$, $U(t) = u(t, 0)$, we get $U'(t) = rU$, $U(T) = u_0(0)$. Evidently, $U(t) = u_0(0)e^{-r(T-t)}$ is the only solution of that Cauchy problem. So $u\,|_{\Sigma_0}$, with $\Sigma_0 = \{0 < t < T, S = 0^+\}$, is uniquely determined by $u_0(0)$.

For this example our CNN model is the following:

$$\frac{dS_{ij}}{dt} + rS_{ij}A_1 * S_{ij} + \frac{1}{2}\sigma^2 S_{ij}^2 A_2 * S_{ij} - r = 0, \tag{28}$$

where $*$ is the convolution operator [24], $M \leq i, j \leq M$. We shall consider this model with free-boundary conditions:

$$u_{ij}(x, t) = x - k, \frac{\partial u_{ij}(x, t)}{dt} = +1,$$

$$u_{ij}(x, t) = k - x, \frac{\partial u_{ij}(x, t)}{dt} = -1.$$

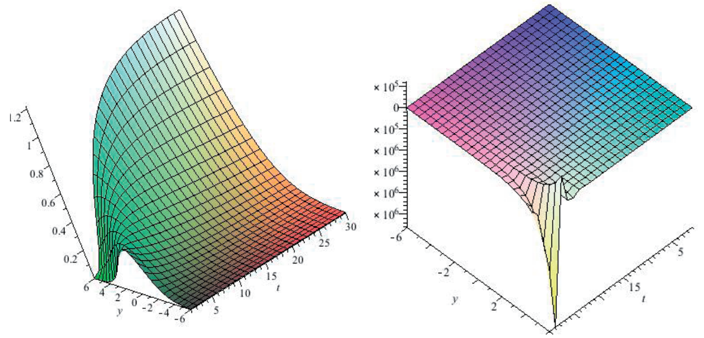
These are classical first-order contact free-boundary conditions for obstacle problems.

Based on the above CNN model (28) we obtain the following simulations for different values of the parameters:

Example 2. (Multi-asset option with single barrier) Analytic valuation formulas for standard European options with single external barrier have been provided in Heynen-Kat (1994), Kwok-Wu-Yu (1998) and Buchen (2001). Here we give a slightly more general formula in that we allow for any payoff and for both an internal and an external barrier. We confine ourselves to the case of an upstream barrier and zero rebate for simplicity of exposition. Consider the following boundary value problem in $\Omega = \{(t, S_1, S_2); 0 < t < T, 0 < S_1, 0 < S_2 < S^*\}$:

$$\begin{cases} Lu = 0 \\ u|_{t=T} = u_0(S_1, S_2) \quad 0 \leq S_2 \leq S^* \\ u|_{S_2 = S^*} = 0 \quad 0 \leq t \leq T \end{cases}$$

where $L = \partial_t + \sum_{i=1}^2 \frac{\sigma_i^2 S_i^2}{2} \partial_{S_i}^2 + \rho\sigma_1\sigma_2 S_1 S_2 \partial_{S_1 S_2}^2 + r\sum_{i=1}^2 S_i \partial_{S_i} - r$, u_0 is continuous and $u_0(S_1, S^*) = 0$. Assume that $\sigma_1, \sigma_2 > 0$, $\rho^2 < 1$. Using the notation of Section 2 and



(a) (b)

Figure 2. CNN simulations for Example 1. (a) $r = 1, 1 \le t \le 30, \sigma = 1$; (b) $r = 0.5, 1 \le t \le 30, \sigma = 1.5$.

taking $\mu_i = r - \frac{\sigma_i^2}{2}$ for $i, j = 1,2$, we have $\alpha_i = \frac{-\mu_i + \rho\mu_j\sigma_i/\sigma_j}{\sigma_i^2(1-\rho^2)}$ for $i, j = 1,2$ and $i \neq j$, $\beta = \sum_{i,j=1,2} \frac{\sigma_i\sigma_j}{2}\alpha_i\alpha_j + \sum_{i=1,2}\mu_i\alpha_i - r$. Then we have the following pricing formula:

$$u(t,S_1,S_2) = S_1^{\alpha_1}S_2^{\alpha_2}\frac{e^{\beta\tau}}{4\pi\tau}\int_{R^2} w_0(\lambda_1,\lambda_2)\exp[-\frac{\left|\frac{\sqrt{2}\ln S_1}{\sigma_1\sqrt{1-\rho^2}} - \rho\frac{\sqrt{2}\ln S_2}{\sigma_2\sqrt{1-\rho^2}} - \lambda_1\right|^2}{4\tau}]$$
$$\left\{\exp[-\frac{(\sqrt{2}\ln S_2/\sigma_2 - \lambda_2)^2}{4\tau}] - \exp[-\frac{(\sqrt{2}\ln S_2/\sigma_2 + \lambda_2)^2}{4\tau}]\right\} d\lambda_1 d\lambda_2$$

where
$w_0(\lambda_1,\lambda_2) = \exp[-\frac{\alpha_1\sigma_1}{\sqrt{2}}(\lambda_1\sqrt{1-\rho^2} + \rho\lambda_2) - \frac{\alpha_2\sigma_2}{\sqrt{2}}\lambda_2]u_0(\frac{\sigma_1}{\sqrt{2}}(\lambda_1\sqrt{1-\rho^2} + \rho\lambda_2), \frac{\sigma_2\lambda_2}{\sqrt{2}})1_{\lambda_2} < \frac{\sqrt{2}\ln S^*}{\sigma_2}$.

Splitting the integral into two integrals and changing to variables $\eta_1 = \frac{\lambda_1\sqrt{1-\rho^2} + \rho\lambda_2 - \sqrt{2}\ln S_1}{\sigma_1\sqrt{2\tau}}$, $\eta_2 = \frac{\lambda_2 - \sqrt{2}\ln S_2}{\sigma_2\sqrt{2\tau}}$ ($\eta_2 = \frac{\lambda_2 + \sqrt{2}\ln S_2}{\sigma_2\sqrt{2\tau}}$) in the first (second) integral, one gets:

$$u(t,S_1,S_2) = I_1 - I_2$$

where

$$I_1 = \frac{e^{\beta\tau}}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{\frac{\ln(S^*/S_2)}{\sigma_2\sqrt{\tau}}} \exp[-(\alpha_1\sigma_1\eta_1 + \alpha_2\sigma_2\eta_2)\sqrt{\tau}]u_0(S_1e^{\sigma_1\sqrt{\tau}\eta_1}, S_2e^{\sigma_2\sqrt{\tau}\eta_2})$$
$$\exp[-\frac{(\eta_1 - \rho\eta_2)^2}{2(1-\rho^2)} - \frac{\eta_2^2}{2}]d\eta_1 d\eta_2$$
$$I_2 = \frac{S_2e^{\beta\tau}}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{\frac{\ln(S^*/S_2)}{\sigma_2\sqrt{\tau}}} \exp[-(\alpha_1\sigma_1\eta_1 + \alpha_2\sigma_2\eta_2)\sqrt{\tau}]u_0(S_1e^{\sigma_1\sqrt{\tau}\eta_1}, S_2^{-1}e^{\sigma_2\sqrt{\tau}\eta_2})$$
$$\exp[-\frac{(-\eta_1 + \rho\eta_2 - 2\rho\ln S_2/(\sigma_2\sqrt{\tau}))^2}{2(1-\rho^2)} - \frac{\eta_2^2}{2}]d\eta_1 d\eta_2.$$

Note that $(\beta + r)(1-\rho^2) + \frac{\mu_1^2}{2\sigma_1^2} + \frac{\mu_2^2}{2\sigma_2^2} - \rho\frac{\mu_1\mu_2}{\sigma_1\sigma_2} = 0$. Then the first integral (after changing to variables $X_1 = -\eta_1 + \frac{\mu_1}{\sigma_1}\sqrt{\tau}$, $X_2 = \eta_2 - \frac{\mu_2}{\sigma_2}\sqrt{\tau}$) is written in the form:

$I_1 = \frac{e^{-r\tau}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{-\frac{\ln(S_2/S^*)+\mu_2\tau}{\sigma_2\sqrt{\tau}}} \exp[-\frac{1}{2(1-\rho^2)}(X_1^2 + X_2^2 + 2\rho X_1 X_2)] u_0(S_1 e^{\mu_1\tau - \sigma_1\sqrt{\tau}X_1},$

$S_2 e^{\mu_2\tau+\sigma_2\sqrt{\tau}X_2}) dX_1 dX_2.$

Changing to the variables $X_1 = -\eta_1 + \frac{\mu_1}{\sigma_1}\sqrt{\tau} - \frac{2\rho \ln S_2}{\sigma_2\sqrt{\tau}}$, $X_2 = \eta_2 - \frac{\mu_2}{\sigma_2}\sqrt{\tau}$, the second integral becomes:

$I_2 = \frac{e^{-r\tau}}{2\pi\sqrt{1-\rho^2}} (S_2)^{-\frac{2\mu_2}{\sigma_2^2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{\frac{\ln(S_2 S^*)-\mu_2\tau}{\sigma_2\sqrt{\tau}}} \exp[-\frac{1}{2(1-\rho^2)}(X_1^2 + X_2^2 + 2\rho X_1 X_2)]$
$u_0(S_1 S_2^{-2\rho\frac{\sigma_1}{\sigma_2}} e^{\mu_1\tau-\sigma_1\sqrt{\tau}X_1}, S_2^{-1} e^{\mu_2\tau+\sigma_2\sqrt{\tau}X_2}) dX_1 dX_2.$

In the special case of standard options one has: $u_0(S_1, S_2) = \max(\omega(S_1 - K), 0)$, $\omega = \pm 1$. Then I_1 can be written in the form:
$\omega S_1 N_2(\omega d^+, e^+; -\rho\omega) - \omega K e^{-r\tau} N_2(\omega d^-, e^-; -\rho\omega)$

where N_2 is the bivariate cumulative normal distribution function, $d^\pm = \frac{\ln(\frac{S_1}{K}) + (r \pm \frac{\sigma_1^2}{2})\tau}{\sigma_1\sqrt{\tau}}$, $e^- = -\frac{\ln(\frac{S_2}{S^*}) + \mu_2\tau}{\sigma_2\sqrt{\tau}}$, $e^+ = e^- - \rho\sigma_1\sqrt{\tau}$. Similarly I_2 is written in the form:

$\omega e^{-2\frac{\mu_2}{\sigma_2^2}\ln(\frac{S_2}{S^*})}[e^{-2\rho\frac{\sigma_1}{\sigma_2}\ln(\frac{S_2}{S^*})} S_1 N_2(\omega\widehat{d}^+, \widehat{e}^+; -\rho\omega) - K e^{-r\tau} N_2(\omega\widehat{d}^-, \widehat{e}^-; -\rho\omega)$

where $\widehat{d}^\pm = d^\pm - \frac{2\rho}{\sigma_2\sqrt{\tau}}\ln(\frac{S_2}{S^*})$, $\widehat{e}^\pm = e^\pm + \frac{2}{\sigma_2\sqrt{\tau}}\ln(\frac{S_2}{S^*})$.

Simulating CNN for multi-asset option with single barrier model, we obtain the following figure with different values of the parameter set:

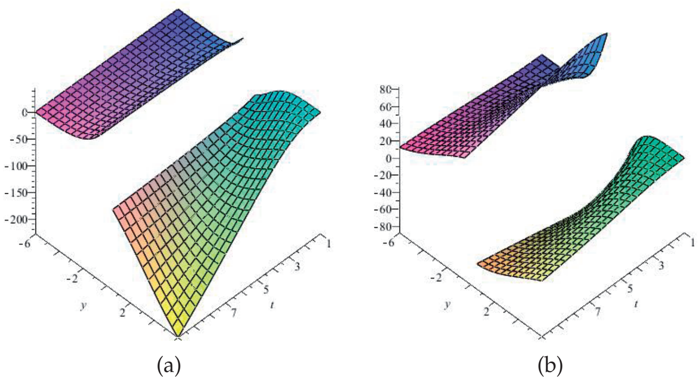


Figure 3. CNN simulations for Example 2. (a) $r = 1$, $T = 60$ days, $\sigma = 1$, $\rho = 0.05$; (b) $r = 0.5$, $T = 120$ days, $\sigma = 1.5$, $\rho = 0.06$.

Example 3. (Two-asset barrier options with simultaneous barriers) While single-asset barrier options have received substantial coverage in the literature, multi-asset options with several barriers have been discussed only in some special cases (e.g. sequential barriers, radial options, etc.). Here we show how the case of two simultaneous barriers can be valued straightforwardly from the arguments in Section 2. Let us confine ourselves to zero-rebate options for simplicity's sake, although Section 2 deals with the general case too. Then the boundary value problem takes the form:

$$\begin{cases} Lu = 0 \quad \text{in} \quad \Omega \\ u|_{t=T} = u_0(S_1, S_2) \\ u|_{S_1=S_1^*} = 0 \quad \text{and,} \quad u|_{S_2=S_2^*} = 0 \quad 0 \le t \le T \end{cases}$$

where $L = \partial_t + \sum_{i=1}^{2} \frac{\sigma_i^2 S_i^2}{2} \partial_{S_i}^2 + \rho\sigma_1\sigma_2 S_1 S_2 \partial_{S_1 S_2}^2 + r \sum_{i=1}^{2} S_i \partial_{S_i} - r$, $\Omega = \{(t, S_1, S_2);\ 0 < t < T, 0 < S_1 < S_1^*, 0 < S_2 < S_2^*\}$. Arguing as in the last part of Section 2 and taking

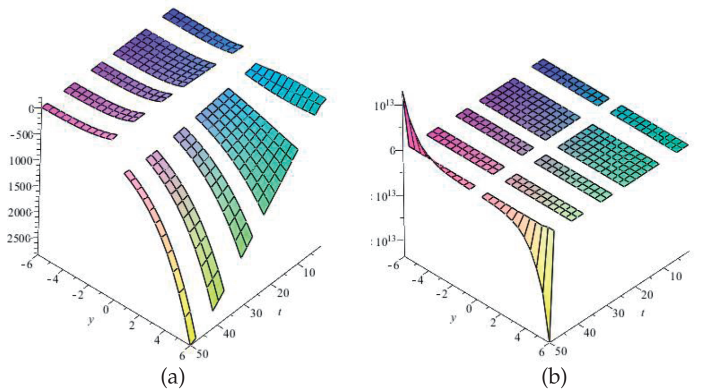
$$D = \begin{pmatrix} \sigma_1 & \rho\sigma_2 \\ 0 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix}, \rho^2 < 1, \sigma_1 > 0, \sigma_2 > 0$$

and φ_0 as the opening of the angle between $\begin{cases} x \le 0 \\ y = 0 \end{cases}$ and $\begin{cases} x = \rho\sigma_2\eta, \eta \ge 0 \\ y = \sqrt{1-\rho^2}\sigma_2\eta \end{cases}$, from (21) we have

$$w(\tau, r, \varphi) = \int_0^{\varphi_0} \int_0^{\infty} w_0(\xi, \eta) G(r, \varphi, \xi, \eta, \tau) \xi d\xi d\eta, \tag{29}$$

where $G(r, \varphi, \xi, \eta, \tau) = \frac{1}{\varphi_0 \tau} e^{-\frac{(r^2+\xi^2)}{4\tau}} \sum_{n=1}^{\infty} I_{\frac{n\pi}{\varphi_0}}(\frac{r\xi}{2\tau}) sin \frac{n\pi}{\varphi_0} \varphi sin \frac{n\pi}{\varphi_0} \eta$ and I_v is the modified Bessel function satisfying (22). Here $w_0(r, \varphi) = \widetilde{v}_0(D^* z) \ |_{z_1 = r\cos\varphi, z_2 = r\sin\varphi}$ where $\widetilde{v}_0(\lambda) = u_0(S_1^* e^{-\lambda_1}, S_2^* e^{-\lambda_2}) e^{-\Sigma\alpha_i(\ln S_i^* - \lambda_i)}$. Changing back the variables one obtains $u(t, S_1, S_2)$.

Simulating CNN for two-asset barrier options with simultaneous barriers model, we obtain the following figure with different values of the parameter set:



(a) (b)

Figure 4. CNN simulations for Example 3. (a) $r = 1$, $T = 120$ days, $\sigma = 1$, $\rho = 0.05$; (b) $r = 0.5$, $T = 180$ days, $\sigma = 1.5$, $\rho = 0.06$.

4. Comparison principle for multi-asset Black-Scholes equations

For the sake of simplicity consider

$$u_t + \sum_{i,j=1}^{2} a_{ij}x_i x_j u_{x_i x_j} + \sum_{i=1}^{2} b_i x_i u_{x_i} + cu = f, \tag{30}$$

where $(a_{ij})^* = (a_{ij})$, $(a_{ij}) > 0$, a_{ij}, b_i, c are real constants and $c < 0$ in the domain $D : \left\{ \begin{array}{l} 0 < t < T \\ 0 < x_j < a_j, \end{array} j = 1,2 \right\}$, $a_j = const > 0$. The boundary of the parallelepiped D is split into two parts: Parabolic $\Gamma = \{x_1 = a_1, 0 < x_2 < a_2, 0 < t < T\} \cup \{x_2 = a_2, 0 < x_1 < a_1, 0 < t < T\} \cup \{t = T, 0 < x_j < a_j, j = 1,2\}$ and free of boundary data part $\Gamma_1 = I \cup II \cup III$, where $I = \{0 < x_j < a_j, j = 1,2; t = 0\}$, $II = \{x_1 = 0, 0 < x_2 < a_2, 0 < t < T\}$, $III = \{x_2 = 0, 0 < x_1 < a_1, 0 < t < T\}$. The Dirichlet data are prescribed on Γ:

$$u|_\Gamma = g \tag{31}$$

Theorem 1. (Comparison principle)

Assume that u is a classical solution of (30), (31), i.e. $u \in C^2(D \cup \bar{\Gamma}_1) \cap C^0(\bar{D})$. Let v be another solution of (30), (31) belonging to $C^2(D \cup \bar{\Gamma}_1) \cap C^0(\bar{D})$. Suppose that $u|_\Gamma \le v|_\Gamma$. Then $u \le v$ everywhere in $\bar{D}$.

Proof. Put $w = u - v$. Assume that $max\ w = w(t_0, x_0) = M > 0$, $P_0 = (t_0, x_0) \in \bar{D}$. Evidently, $(t_0, x_0) \in D \cup \Gamma_1$ as $w|_\Gamma \le 0$.

Case a). $(t_0, x_0) \in D$. Having in mind that $\sum a_{ij} x_i x_j w_{x_i x_j}$ is a strictly elliptic operator in the open rectangle $\{0 < x_j < a_j, j = 1,2\}$ we shall apply the interior parabolic maximum principle (see A.Friedman, Partial Differential equations of parabolic type, Prentice Hall, Inc. (1964), Chapter II). To do this we shall work in the domain $D_1 : \left\{ \begin{array}{l} 0 < t < T \\ 0 < \varepsilon_j < x_j < a_j, j = 1,2 \end{array} \right\}$, such that $x_0 \in \Pi = (\varepsilon_1, a_1) \times (\varepsilon_2, a_2)$, $0 < t_0 < T$. Then Th1 from Chapter II of the above mentioned book gives: $w \equiv M > 0$ for $T \ge t \ge t_0$, $x \in \bar{\Pi}$ and this is a contradiction with $w \le 0$ on $t = T$.

Case b). $(t_0, x_0) \in I \Rightarrow t_0 = 0$, (1) $\left\{ \begin{array}{l} 0 < x_{10} < a_1 \\ 0 < x_{20} < a_2 \end{array} \right.$, (2) $\left\{ \begin{array}{l} 0 < x_{10} < a_1 \\ x_{20} = 0 \end{array} \right.$, (3) $\left\{ \begin{array}{l} x_{10} = 0 \\ x_{20} = 0 \end{array} \right.$ and a similar case with respect to $x_{20} \in [0, a_2)$, $x_{10} = 0$. Thus,

b). (1) x_0 is interior point of $(0, a_1) \times (0, a_2)$ and therefore $\frac{\partial w}{\partial x_j}(P_0) = 0$, $j = 1,2$, while $\sum_{i,j}^2 a_{ij} x_{i0} x_{j0} \frac{\partial^2 w}{\partial x_i \partial x_j}(P_0) \le 0$ as it is shown in Friedman book. Obviously, $w_t(P_0) \le 0$, as $w(0, x_0) = M = max_{\bar{D}} w$. As we know, (30) is satisfied on I $\Rightarrow \sum_1^2 a_{ij} x_{i0} x_{j0} \frac{\partial^2 w}{\partial x_i \partial x_j}(P_0) + cw(P_0) + w_t(P_0) = 0$ -contradiction with $c < 0$, $w(P_0) > 0$.

b). (2) Again $w_t(P_0) \le 0$ and $w_{x_1}(P_0) = 0$, $w_{x_1 x_1}(P_0) \le 0$ as P_0 is interior point for the interval $(0, a_1)$. According to (30) : $a_{11} x_{10}^2 \frac{\partial^2 w}{\partial x_1^2}(P_0) + b_1 \frac{\partial w}{\partial x_1}(P_0) + cw(P_0) + w_t(P_0) = 0 \rightarrow$ contradiction.

b). (3) Then (30) takes the form: $cw(P_0) + w_t(P_0) = 0$ - contradiction.

Case c). $(t_0, x_0) \in II \Rightarrow 0 \le t_0 < T$, $x_{10} = 0$; (1) $\begin{cases} 0 < t_0 < T \\ 0 < x_{20} < a_2 \end{cases}$, (2) $\begin{cases} t_0 = 0 \\ 0 < x_2^0 < a_2 \end{cases}$, (3) $\begin{cases} t_0 = 0 \\ x_2^0 = 0 \end{cases}$, (4) $\begin{cases} T > t_0 > 0 \\ x_2^0 = 0 \end{cases}$.

Certainly, $w_t(P_0) \le 0$ in each case (1) -(4).

c). (1) As P_0 is interior point in the rectangle $\{0 < t < T\} \times \{0 < x_2 < a_2\} \Rightarrow w_t(P_0) = 0$, $w_{x_2}(P_0) = 0$, $w_{x_2x_2}(P_0) \le 0$. According to (30) $a_2x_{20}^2 w_{x_2x_2}(P_0) + b_2x_{20}w_{x_2}(P_0) + cw(P_0) + w_t(P_0) = 0$ - contradiction.

c). (2) As $x_2^0 \in (0, a_2) \Rightarrow w_{x_2}(P_0) = 0$, $w_{x_2x_2}(P_0) \le 0$. The contradiction is obvious.

c). (3) The equation (30) takes the form:

$$cw(P_0) + w_t(P_0) = 0 \tag{32}$$

and again a contradiction.

c). (4). Then $w_t(P_0) = 0$ and according to (30) $cw(P_0) + w_t(P_0) = 0$ - contradiction.

We conclude that $M = sup_{\bar{D}} w \le 0 \Rightarrow u - v \le 0$ in $\bar{D} \Rightarrow u \le v$ in $\bar{D}$.

The comparison principle is proved.

Remark 5. The operator

$$Lu = u_t + \sum_{i,j=1}^{n} a_{ij}x_ix_ju_{x_ix_j} + \sum_{i=1}^{n} b_ix_iu_{x_i} + cu$$

is non-hypoelliptic. The constants a_{ij}, b_i, c are arbitrary. To verify this we recall that the function $s_+^a = \begin{cases} s^a, s > 0 \\ 0, \ s \le 0 \end{cases}$ considered as a Schwartz distribution in $D'(\mathbf{R}^1)$ satisfies for $Re\, a > 1$ the following identities:

$$ss_+^a = s_+^{a+1}, \frac{d}{ds}s_+^a = as_+^{a-1}, \frac{d^2}{ds^2}s_+^a = a(a-1)s_+^{a-2}.$$

Consider now the distribution $u = e^{\lambda t}u_1(x_1) \otimes \ldots \otimes u_n(x_n)$, where $\lambda = const$, $u_j(x_j) = x_j^{d_j} \in D'(\mathbf{R}^1_{x_j})$, $Re d_j > 1$. Then $u \in D'(\mathbf{R}^{n+1})$ satisfies in distribution sense $Lu = 0$ if

$$\lambda + \sum_{i \ne j}^{n} a_{ij}d_id_j + \sum_{i=j}^{n} a_{ii}d_i(d_i - 1) + \sum_{i=1}^{n} b_id_i + c = 0$$

Of course, $sing\ supp\ u = \partial\{x \in \mathbf{R}^n : x_j \ge 0, 1 \le j \le n\}$, i.e. $sing\ supp\ u$ is the boundary of the first octant of $\mathbf{R}^n_x$ multiplied by $\mathbf{R}^1_t$. The nonhypoellipticity is proved. Evidently, under (4) L is hypoelliptic in the open domain $\{x_j > 0, 1 \le j \le n\}$ as it is strictly parabolic there.

5. The approach of Fichera-Oleinik-Radkevič

In this section we revise the results of [9] and [20] for the Dirichlet problem for PDEs of second order having non-negative characteristic form; then the method is applied to some PDEs of Black-Scholes type.

To begin with consider the following equation in a bounded domain $\Omega \subset \mathbf{R}^m$ with piecewise smooth boundary Σ:

$$L(u) = \sum_{k,j=1,\dots,m} a^{kj}(x)u_{x_kx_j} + \sum_{k=1,\dots,m} b^k(x)u_{x_k} + c(x)u = f(x) \tag{33}$$

where $\sum_{k,j=1,\dots,m} a^{kj}(x)\xi_k\xi_j \geq 0$, $\forall x \in \overline{\Omega}$, $\forall \xi \in \mathbf{R}^m$; $a^{kj}(x) = a^{jk}(x)$, $\forall x \in \Omega$. Moreover, $a^{kj} \in C^2(\overline{\Omega})$, $b^k \in C^1(\overline{\Omega})$, $c \in C^0(\overline{\Omega})$. Denote the unit inner normal to Σ by $\overrightarrow{n} = (n_1, \dots, n_m)$ and let $\Sigma_3 = \left\{x \in \Sigma; \sum_{k,j=1,\dots,m} a^{kj}(x)n_kn_j > 0\right\}$ be the non-characteristic part of Σ. Define $\Sigma^0 = \left\{x \in \Sigma; \sum_{k,j=1,\dots,m} a^{kj}(x)n_kn_j = 0\right\}$, i.e. $\Sigma = \Sigma^0 \cup \Sigma_3$ and Σ^0 is the characteristic part of Σ. Following Fichera (1956) we introduce on Σ^0 the Fichera function:

$$\beta(x) = \sum_{k=1,\dots,m} (b^k(x) - \sum_{j=1,\dots,m} a^{kj}_{x_j}(x))n_k, x \in \Sigma^0 \tag{34}$$

Then we split Σ^0 into three parts, namely

$\Sigma_1 = \left\{x \in \Sigma^0; \beta(x) > 0\right\}$,

$\Sigma_2 = \left\{x \in \Sigma^0; \beta(x) < 0\right\}$,

$\Sigma_0 = \left\{x \in \Sigma^0; \beta(x) = 0\right\}$.

As it is proved in Oleinik and Radkevič (1971) the sets Σ_0, Σ_1, Σ_2, Σ_3 are invariant under smooth non-degenerate changes of the variables. More precisely, let $L(u) = f$ in Ω; after the change $y = F(x)$ it takes the form $\widetilde{L}(\widetilde{u}) = \widetilde{f}$ in $\widetilde{\Omega}$. Denote the Fichera function for $\widetilde{L}(\widetilde{u}) = \widetilde{f}$ by $\widetilde{\beta}$. Then $\widetilde{\beta} = \beta.A$ where $A > 0$ and A is continuous.

Assume now that $u \in C^2(\Omega)$ and $v \in C_0^\infty(\Omega)$. Then

$\int_\Omega L(u)vdx = \int_\Omega uL^*(v)dx$,

where

$$L^*(v) = \sum_{k,j=1,\dots,m} a^{kj}(x)v_{x_kx_j} + \sum_{k=1,\dots,m} b^{*k}(x)v_{x_k} + c^*(x)v \tag{35}$$

and $b^{*k} = 2\sum_{j=1,\dots,m} a^{kj}_{x_j} - b^k$, $c^* = \sum_{k=1,\dots,m}(\sum_{j=1,\dots,m} a^{kj}_{x_kx_j} - b^k_{x_k}) + c$. One can easily see that if we denote the Fichera function for $L^*(v)$ by β^*, then $\beta^* = -\beta$ and β is defined by (34).

Assume now that $u \in C^2(\overline{\Omega})$, $u = 0$ at $\Sigma_2 \cup \Sigma_3$, and define the following set of test functions: $\mathcal{V} = \left\{v \in C^2(\overline{\Omega}); v = 0 \text{ at } \Sigma_1 \cup \Sigma_3\right\}$. In view of the Green formula for L we get:

$$\int_\Omega (L(u)v - L^*(v)u)dx = 0 \Leftrightarrow \int_\Omega L(u)vdx = \int_\Omega uL^*(v)dx \tag{36}$$

for any u and $v \in \mathcal{V}$. Let us now recall the definitions of generalized solution.

Definition 3. *The function $u \in L^p(\Omega)$, $p \geq 1$, is called a generalized solution of the boundary value problem*

$$\begin{cases} L(u) = f & \text{in} \quad \Omega \\ u = 0 & \text{at} \quad \Sigma_2 \cup \Sigma_3 \end{cases} \tag{37}$$

if for each test function $v \in \mathcal{V}$ the following integral identity holds:

$$\int_\Omega f v dx = \int_\Omega u L^*(v) dx. \tag{38}$$

Theorem 2. (See [20],Th.1.3.1).

Suppose that $c < 0$, $c^* < 0$ in Ω and $p > 1$. Then for each $f \in L^p(\Omega)$ there exists a generalized solution $u \in L^p(\Omega)$ of (37) in the sense of (38) and such that

$$\inf_{u_0 \in Z} \|u + u_0\|_{L^p(\Omega)} \leq K \|f\|_{L^p(\Omega)} \tag{39}$$

$K = const > 0$. The set $Z = \left\{u_0 \in L^p(\Omega) : \int_\Omega u_0 L^*(v) dx = 0, \forall v \in \mathcal{V}\right\}$.

Theorem 3. (See [20], Th. 1.3.2).

Let $c < 0$ in $\overline{\Omega}$, $\frac{1}{p} + \frac{1}{q} = 1$ and $-c + (1-q)c^* > 0$ in $\overline{\Omega}$. Then for each $f \in L^p(\Omega)$ there exists a generalized solution u of (37) satisfying the a-priori estimate (39).

Theorem 4. (See [20], Th. 1.3.3).

Let $c^* < 0$ in $\overline{\Omega}$ and $-c + (1-q)c^* > 0$ in $\overline{\Omega}$, $\frac{1}{p} + \frac{1}{q} = 1$. Then for each $f \in L^p(\Omega)$ there exists a generalized solution u of (37) satisfying the estimate (39).

Conclusion. Assume that $c < 0$. Then (37) is solvable in the sense of Definition 1 for $p >> 1$ as $p \to +\infty \Rightarrow q \to 1$. On the other hand, $c^* < 0$ implies the solvability of (40) for $p \geq 1$, $p \approx 1$ as $p \to 1 \Rightarrow q \to +\infty$.

We shall now discuss the problem for existence of a generalized solution of (37) in the Sobolev space $H^1(\Omega)$ with an appropriate weight. Define the following set of test functions:

$$\mathcal{W} = \{v \in C^1(\overline{\Omega}); v \mid_{\Sigma_3} = 0\}$$

and equip $\mathcal{W}$ with the scalar product: $(u,v)_{\mathcal{H}} = \int_\Omega (\sum_{k,j} a^{kj} u_{x_j} v_{x_k} + uv) dx + \int_{\Sigma_1 \cup \Sigma_3} uv \, |\beta| \, d\sigma$. The completion of $\mathcal{W}$ with respect to the norm $\|u\|_{\mathcal{H}}$ is a real Hilbert space denoted by $\mathcal{H}$. For each two functions $u, v \in \mathcal{W}$ we consider the bilinear form $B(u,v) = -\int_\Omega [\sum_{k,j} a^{kj} u_{x_j} v_{x_k} + \sum_k (u \ell^k v_{x_k} + (\ell^k_{x_k} - c) uv)] dx - \int_{\Sigma_1} uv\beta d\sigma$, where $l^k = b^k - \sum_j a^{kj}_{x_j}$. According to the Cauchy-Schwartz inequality $|B(u,v)| \leq const[\|v\|_{H^1(\Omega)} + \|v\|_{L^2(\Sigma_1)}] \, \|u\|_{\mathcal{H}}$. Therefore, $B(u,v)$ is well defined for $v \in \mathcal{W}$ and $u \in \mathcal{H}$.

Definition 4. *Let $f \in L^2(\Omega)$. We shall say that the function $u \in \mathcal{H}$ is a generalized solution of (37) if for each $v \in \mathcal{W}$ the following identity is satisfied:*

$$\int_{\Omega} v f dx = B(u,v). \tag{40}$$

Theorem 5. (See [20], Th. 1.4.1).

Assume that $f \in L^2(\Omega)$ and $\frac{1}{2}\sum_k (b^k_{x_k} - \sum_j a^{kj}_{x_k x_j}) - c \geq c_0 > 0$ in $\overline{\Omega}$. Then the boundary value problem (37) possesses a generalized solution $u \in \mathcal{H}$ (i.e. a weak solution) in the sense of (40).

Finally we propose the existence of a generalized solution of (37) in the space $L^\infty(\Omega)$. To fix the ideas we assume that the coefficients of L and L^* belong to $C^1(\overline{\Omega})$ and Σ is thrice piecewise smooth (i.e. Σ can be split into several parts and each of them is C^3 smooth). Consider the boundary value problem:

$$\begin{cases} L(u) = f & \text{in } \Omega \\ u = g & \text{on } \Sigma_2 \cup \Sigma_3 \end{cases} \tag{41}$$

If $u \in C^2(\overline{\Omega})$ is a classical solution of (41) and $v \in \mathcal{V}$ then according to the Green formula

$$\int_{\Omega} L^*(v) u dx = \int_{\Omega} f v dx - \int_{\Sigma_3} g \frac{\partial v}{\partial \overrightarrow{\nu}} d\sigma + \int_{\Sigma_2} \beta g v d\sigma, \tag{42}$$

where $\overrightarrow{\nu} = (\nu_1, \ldots, \nu_m)$, $\nu_k = \sum_j a^{kj} n_j$, $\frac{\partial}{\partial \overrightarrow{\nu}} = \sum_k \nu_k \frac{\partial}{\partial x_k}$.

Definition 5. *We shall say that the function $u \in L^\infty(\Omega)$ is a generalized solution of (41) if for each test function $v \in \mathcal{V}$ the identity (42) is fulfilled.*

We point out that $f \in L^\infty(\Omega)$ and $g \in L^\infty(\Sigma_2 \cup \Sigma_3)$.

Theorem 6. (See [20], Th. 1.5.1).

Assume that the coefficient $c(x)$ of L is such that $c(x) \leq -c_0 < 0$ in $\overline{\Omega}$, $f \in L^\infty(\Omega)$, $g \in L^\infty(\Sigma_2 \cup \Sigma_3)$ and $\beta(x) \leq 0$ in the interior points of $\Sigma_2 \cup \Sigma_0$. Then there exists a generalized solution of (41) in the sense of Definition 5. Moreover, $|u(x)| \leq \max(\sup \frac{|f|}{c_0}, \sup |g|)$.

Remark 6. In Th.6 it is assumed that $\sum_{k,j=1,\ldots,m} a^{kj}(x)\xi_k\xi_j \geq 0$ in an $m-$dimensional neighbourhood of Σ^0, $\forall \xi \in \mathbf{R}^m$.

Theorem 7. (See [20], Th. 1.5.2).

Suppose that g is continuous in the interior points of $\Sigma_2 \cup \Sigma_3$. Then the generalized solution u of (41) constructed in Th. 6 is continuous at those points and, moreover, $u = g$ there.

As we shall deal with (degenerate) parabolic PDEs we shall have to work in cylindrical domains (rectangles in $\mathbf{R}^2$). Therefore $\Sigma = \partial\Omega$ will be piecewise smooth. Consider now the bounded domain Ω having piecewise C^3 smooth boundary Σ. The corresponding boundary value problem is:

$$\begin{cases} L(u) = f \text{ in } \Omega, \\ u = 0 \quad \text{on } \Sigma_2 \cup \Sigma_3 \end{cases} \tag{43}$$

We shall say that the point $P \in \Sigma$ is regular if locally near to P the surface Σ can be written in the form $x_k = \varphi_k(x_1, ..., x_{k-1}, x_{k+1}, ..., x_m)$, $(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_m)$ describing some neighborhood of the projection of P onto the plane $x_k = 0$. The set of the boundary points which do not possess such a representation will be denoted by B.

Definition 6. *The function $u \in L^\infty(\Omega)$ is called a generalized solution of (43) for $f \in L^\infty(\Omega)$ if for each function $v \in C^2(\overline{\Omega})$, $v = 0$ at $\Sigma_1 \cup \Sigma_3 \cup B$ the following identity holds:*
$\int_\Omega uL^*(v)dx = \int_\Omega fvdx.$

Theorem 8. (See [20], Th. 1.5.5).

Suppose that the boundary Σ of the bounded domain Ω is C^3 piecewise smooth, $f \in L^\infty(\Omega)$, $g = 0$, $c(x) \leq -c_0 < 0$ in $\overline{\Omega}$ and $\beta \leq 0$ in the interior points of $\Sigma_0 \cup \Sigma_2$. Then there exists a generalized solution u of (43) in the sense of Definition 6 and such that $|u| \leq \sup \frac{|f|}{c_0}$.

We shall not discuss here in details the problems of uniqueness and regularity of the generalized solutions. Unicity results are given by Theorems 1.6.1.-1.6.2. in [20]. For domains with C^3 smooth boundary under several restrictions on the coefficients, including $c(x) \leq -c_0 < 0$, $c^* < 0$ in $\overline{\Omega}$, $\beta \leq 0$ in the interior points of $\Sigma_0 \cup \Sigma_2$, $\beta^* = -\beta < 0$ at Σ_1, the maximum principle is valid for each generalized solution u in the sense of Definition 5:
$|u| \leq \max\left\{\sup_\Omega \frac{|f|}{c_0}, \sup_{\Sigma_3 \cup \Sigma_2} |g|\right\}.$

In Th. 1.6.9. uniqueness result is proved for the boundary value problem (43) in the class $L^\infty(\Omega)$. The existence result is given Th. 8. Regularity result is given in the Appendix.

Remark 7. Backward parabolic and parabolic operators satisfy the conditions: $a^{km} = 0$, $k = 1, ..., m$, and $b^m = \pm 1$ if $x = (x_1, ..., x_{m-1}, t)$, i.e. $t = x_m$. Put now $u = ve^{\alpha t}$ in (33). Then
$L_1(v) = \sum_{k,j=1,...,m} a^{kj} v_{x_k x_j} + \sum_{k=1,...,m} b^k v_{x_k} + (c + \alpha)v = fe^{-\alpha t}$
and
$L_1^*(w) = \sum_{k,j=1,...,m} a^{kj} w_{x_k x_j} + \sum_{k=1,...,m} b^{*k} w_{x_k} + c_1^* w$
where $c_1 = c + \alpha$, $b^{*k} = 2\sum_{j=1,...,m} a^{kj}_{x_j} - b^k$, $c_1^* = \sum_{k,j=1,...,m} a^{kj}_{x_k x_j} - \sum_{k=1,...,m} b^k_{x_k} + c + \alpha$.

Having in mind that $|c| \leq \tilde{c} = const$ we conclude that for $b^m = \pm 1$ and $\alpha \to \mp\infty$ then $c_1 \to -\infty$, $c_1^* \to -\infty$ uniformly in $(x_1, ..., x_{m-1}, t) \in \Omega$. So for parabolic (backward parabolic) equations the conditions of Theorems 2, 5 are fulfilled.

We shall illustrate the previous results by the backward parabolic equations:

$$L(u) = \frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} + rx\frac{\partial u}{\partial x} - ru = f(t, x) \tag{44}$$

which is the famous Black-Scholes equation, and

$$M(u) = \frac{\partial u}{\partial t} + x^2 \frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x} + c(x)u = f(t, x) \tag{45}$$

We shall work in the following rectangles: $\Omega_1 = \{(t, x) : 0 < t < T, 0 < x < a_1\}$, $\Omega_2 = \{(t, x) : 0 < t < T, a_2 < x < 0\}$, $\Omega = \{(t, x) : 0 < t < T, a_2 < x < a_1\}$. Under the previous notation for Ω we have: $\Sigma_1 = \{t = 0\}$, $\Sigma_2 = \{t = T\}$, $\Sigma_3 = \{x = a_1\} \cup \{x = a_2\}$. Certainly, for Ω_1, Ω_2 another part of the boundary appears, $\Sigma_0 = \{x = 0\}$.

As we know from [20] there exists an $L^p(\Omega_1)$ solution of the boundary value problem

$$\begin{cases} L(u_1) = f & \in \ \Omega_1 \\ u_1 = 0 & \text{on} \ \ \Sigma_2^{(1)} \cup \Sigma_3^{(1)} \end{cases} \tag{46}$$

According to the Definition 3: $\int_{\Omega_1} u_1 L^*(v_1)dx = \int_{\Omega_1} f v_1 dx$ for each test function $v_1 \in C^2(\overline{\Omega}_1)$, $v_1 \mid_{\Sigma_1^{(1)} \cup \Sigma_3^{(1)}} = 0$.

In a similar way there exists $u_2 \in L^p(\Omega_2)$ such that

$$\begin{cases} L(u_2) = f & \text{in} \ \ \Omega_2 \\ u_2 = 0 & \text{on} \ \ \Sigma_2^{(2)} \cup \Sigma_3^{(2)} \end{cases} \tag{47}$$

Therefore: $\int_{\Omega_2} u_2 L^*(v_2)dx = \int_{\Omega_2} f v_2 dx$ for each test function $v_2 \in C^2(\overline{\Omega}_2)$, $v_2|_{\Sigma_1^{(2)} \cup \Sigma_3^{(2)}} = 0$.

Certainly, there exists $u \in L^p(\Omega)$ such that $\int_\Omega u L^*(v)dx = \int_\Omega f v dx$ for each test function $v \in C^2(\overline{\Omega})$, $v \mid_{\Sigma_1 \cup \Sigma_3} = 0$. Evidently, $v \in C^2(\overline{\Omega})$, $v \mid_{\Sigma_1 \cup \Sigma_3} = 0 \Rightarrow v \in C^2(\overline{\Omega}_i)$, $v \mid_{\Sigma_1^{(i)} \cup \Sigma_3^{(i)}} = 0$, $i = 1, 2$. Consequently, $\int_{\Omega_1} u_1 L^*(v)dx = \int_{\Omega_1} f v dx$ and $\int_{\Omega_2} u_2 L^*(v)dx = \int_{\Omega_2} f v dx$, and thus the function

$$W = \begin{cases} u_1 \ \text{in} \ \ \Omega_1 \\ u_2 \ \text{in} \ \ \Omega_2 \end{cases} \in L^p(\Omega) \tag{48}$$

satisfies the identity $\int_\Omega f v dx = \int_{\Omega_1} f v dx + \int_{\Omega_2} f v dx = \int_\Omega W L^*(v)dx$, i.e. W is a generalized $L^p(\Omega)$ solution of

$$\begin{cases} L(W) = f & \text{in} \ \ \Omega \\ W = 0 & \text{on} \ \ \Sigma_2 \cup \Sigma_3 \end{cases} \tag{49}$$

We conclude as follows: *(a)* If u_i satisfies

$$\begin{cases} L(u_i) = f & \text{in} \ \ \Omega_i \\ u_i = 0 & \text{on} \ \ \Sigma_2^{(i)} \cup \Sigma_3^{(i)} \end{cases} \tag{50}$$

$i = 1, 2$, then (48) satisfies (49).

(b) In the special case when $f \in L^\infty(\Omega)$, $u_i \in L^\infty(\Omega_i)$, $i = 1, 2$, $u \in L^\infty(\Omega)$, u satisfies the identity $\int_\Omega f v dx = \int_\Omega u L^*(v)dx$, we have a uniqueness theorem and therefore $u = W$.

The set Σ_0 is called interior boundary of Ω.

Acknowledgments

The authors gratefully acknowledge financial support from CNR/BAS.

Author details

Rossella Agliardi
Department of Mathematics for Economic and Social Sciences, University of Bologna, Italy

Petar Popivanov and Angela Slavova
Institute of Mathematics, Bulgarian Academy of Sciences, Bulgaria

Appendix

One can find results concerning regularity of the generalized solutions of degenerate parabolic operators in cylindrical domains in [14] and [19]. For the sake of simplicity we shall consider only one example from Il'in as the conditions are simple and clear. Consider

$$N(u) = \frac{\partial u}{\partial t} + h(t,x)\frac{\partial^2 u}{\partial x^2} + g(t,x)\frac{\partial u}{\partial x} + c(t,x)u = F(t,x) \tag{51}$$

in the rectangle $Q = \{(t,x) : 0 < t < T, a_2 < x < a_1\}$ and h, g, c, $F \in C^3(\bar{Q})$. Moreover, we assume that in some domain

(i) $Q' \supset \bar{Q}$ the function $h \geq 0$ and $h \in C^2(Q')$.

(ii) Suppose that if $h(t,a_1) = 0$ ($h(t,a_2) = 0$), $0 \leq t \leq T$, then $g(t,a_1) > 0$ ($g(t,a_2) < 0$).

Moreover, we assume that the following compatibility conditions hold:

(iii) $D^{\alpha}_{t,x}F(T,a_1) = D^{\alpha}_{t,x}F(T,a_2) = 0$, $|\alpha| \leq 2$.

Define now the following parts of the boundary ∂Q:

$I = \{(t,x) : 0 < t < T, x = a_2\}, II = \{(t,x) : 0 < t < T, x = a_1),$

$III = \{(t,x) : a_2 < x < a_1, t = 0\}$ and $IV = \{(t,x) : a_2 < x < a_1, t = T\}$.

One can easily see that: $\Sigma_3 = \{(t,x) \in I \cup II : h(t,x) > 0\}, \Sigma^0 = \{(t,x) \in I \cup II : h(t,x) = 0\} \cup \{(t,x) \in III \cup IV\}$, $\beta = gn_1 + n_2 - \frac{\partial h}{\partial x}n_1$, i.e. $(t,x) \in \Sigma^0$, $(t,x) \in I \cup II \Rightarrow h(t,x) = 0 \Rightarrow \frac{\partial h}{\partial x} = 0$ and $\overrightarrow{n} = (1,0)$ on I, $\overrightarrow{n} = (-1,0)$ on II. Thus $\beta \mid_{I\cap\Sigma^0} = gn_1 = g < 0$, while $\beta \mid_{II\cap\Sigma^0} = -g < 0$. Therefore, $I \cap \Sigma^0 \subset \Sigma_2$, $II \cap \Sigma^0 \subset \Sigma_2$. Evidently, $\beta \mid_{III} = n_2 = 1 \Rightarrow III \subset \Sigma_1$, while $IV \subset \Sigma_2$; $\Sigma_0 = \emptyset$.

In conclusion, III is free of data as it is of the type Σ_1; $(I \cup II) \cap \Sigma^0$ and IV are of the type Σ_2, while $\Sigma_3 = (I \cup II) \cap \{h > 0\}$. Part of $I \cup II$ is non-characteristic, part of $I \cup II$ is of Σ_2 type. Data are prescribed on $\Sigma_2 \cup \Sigma_3$, i.e. on $I \cup II \cup IV$.

Theorem 9. (see [14]).

There exists a unique classical solution u of (51), $u \mid_{I\cup II\cup IV} = 0$ under the conditions (i), (ii), (iii). More specifically, there exists Lipschitz continuous derivatives: u, $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial x}$, $\frac{\partial^2 u}{\partial x^2} \in C^{0,\alpha}(\overline{Q})$, $0 < \alpha < 1$.

In [19] it is mentioned that under several restrictions on the coefficients the boundary value problem

$$\begin{cases} N(u) = 0 \\ u \mid_{I\cup II\cup IV} = 0 \end{cases} \tag{52}$$

possesses a unique generalized bounded solution which is Lipschitz continuous in $\overline{Q}$. The proof relies on the method of elliptic regularization.

Remark 8. If $a_2 < x < a_1$, $a_2 < 0$, $a_1 > 0$, the Black-Scholes equation (44) is with $h(t,x) = \frac{\sigma^2}{2}x^2 > 0$ on $I \cup II$, i.e. $\Sigma_3 = I \cup II$ and the equation

$$\begin{cases} L(u) = f \quad \text{in} \quad Q \\ u\mid_{I \cup II \cup IV} = 0 \end{cases}$$

possesses a unique classical solution. As we know, $u\mid_{x=0} = U(t)$ satisfies in classical sense the ODE:

$U\prime(t) - rU(t) = f(t,0)$, $U(T) = 0$. Therefore, we can consider the restrictions: $u\mid_{x>0}$, $u\mid_{x<0}$ and conclude that they are classical solutions of the respective boundary value problems with 0 data at $\Sigma_2^{(1)} \cup \Sigma_3^{(1)}$, respectively at $\Sigma_3^{(2)} \cup \Sigma_2^{(2)}$.

6. References

[1] Agliardi, R. (2011)A comprehensive mathematical approach to exotic option pricing. forthcoming in Mathematical Methods in the Applied Sciences, 19 pp.

[2] Bateman, H., Erdélyi, A. (1953) Higher transcendental functions vol.2. NY-Toronto-London, McGraw-Hill book Company.

[3] Black, F., Cox, J. C. (1976) Valuing corporate securities: some effects of bond indenture provisions. Journal of Finance. 31: 351-367.

[4] Boyarchenko, S. I., and S. Z. Levendorskiĭ (2002) Barrier options and touch-and-out options under regular Lévy processes of exponential type.Annals of Applied Probability. 12: 1261-1298.

[5] Buchen, P. (2001) Image options and the road to barriers. Risk. (Sept.)

[6] Budak, B., Samarskii, A., Tihonov, A. (1972) A Problem Book on the Equations of Mathematical Physics. Nauka, Moscow.

[7] Carr, P. (1995) Two extensions to barrier option valuation. Applied Mathematical Finance. 2: 173-209.

[8] Ericsson, J., Reneby, J. (1998) A framework for valuing corporate securities. Applied Mathematical Finance. 5: 143-163.

[9] Fichera, G. (1956) Sulle equazioni differenziali lineari ellittico-paraboliche del secondo ordine. Atti dell'Accademia Nazionale dei Lincei. Cl. Sci. Fis. Mat. Natur. Rend. Lincei. 5: 1-30.

[10] Geman, H., Yor, M. (1996) Pricing and hedging double barrier options: a probabilistic approach. Mathematical Finance. 6: 365-378.

[11] Haug, E. G. (1997) The Complete Guide to Option Pricing Formulas. Mc-Graw Hill, N.Y.

[12] Heynen, R., Kat, H. (1994) Crossing the barrier. Risk. 7: 46-51.

[13] Heynen, R., Kat, H. (1996) Discrete partial barrier options with a moving barrier. Journal of Financial Engineering. 5: 199-209.

[14] Il'in, A. (1960) Degenerate Elliptic and Parabolic Equations. Russian Math. Sbornik. vol. 50. 4: 443-498

[15] Kou, S. G. (2008) Discrete barrier and lookback options. in J. R. Birge and V. Linetsky (Eds.). Handbooks in OR & MS, Elsevier. vol. 15: 343-373.

[16] Kunitomo, N., Ikeda, M. (1992) Pricing options with curved boundaries. Mathematical Finance. 2: 275-298.

[17] Kwok, Y., Wu, L., Yu, H. (1998) Pricing multi-asset options with an external barrier. International Journal of Theoretical and Applied Finance. 1: 523-541.

[18] Merton, R. (1973) A rational theory of option pricing. Bell Journal of Economics and Management Science.4: 141-183.

[19] Oleinik, O. A. (1965) Regularity of the solutions of degenerate elliptic and parabolic equations. Dokladi Acad. Nauk USSR. 163. 3: 577-580.

[20] Oleinik, O. A., Radkevič, E. V. (1971) Second Order Equations with Non-negative Characteristic Form. Itogi Nauki, Moscow [English translation in Plenum Press, N.Y.-London (1973)].

[21] Poljanin, A. (2001) Handbook-Linear Equations of Mathematical Physics. Fizmatlit, Moscow (In Russian).

[22] Rubinstein, M., Reiner, E. (1991) Breaking down the barriers. Risk. 4: 28-35.

[23] Skipper, M. (2007) The analytic valuation of multi-asset, one-touch barrier options. Oxford Centre of Industrial and Applied Mathematics Research Report. 28 p.

[24] Slavova A. (2003) Cellular Neural Networks:Dynamics and Modeling. Kluwer Academic Publisher. 221p.

[25] Zhang, P. G. (2001) Exotic Options. World Scientific, Singapore. 2^{nd} edition.

Chapter 3

Selecting a Response Plan Under Budget Constraints

A. Gonen

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50202

1. Introduction

Usually project risk management plans deal with identifying, assessing and planning adequate responses to risks. The main problem is that there are many types of response plans, and we must be able to select the optimal one. The usual approach is to first handle the most "dangerous" risk (the risk with the maximum expected damage). However, handling this risk may also be very expensive and beyond the limitations of the allocated risk budget. The dilemma is how to select the right risks to be handled within a limited budget.

2. Risk management methodology

Risk management methodology was first described in detail by Wideman in [1]. The methodology was then improved by the PMI [2], adding details based on users' experience.

Project Risk management involves the following steps:

1. Planning Risk Assessment, which includes selecting an assessment team, setting up rules, and determining the supporting risk management tools. The risk assessment team should include representatives from all areas related to the project.
2. Risk Identification is a process of defining future events that should be considered as risk events. The list is usually generated by a brain-storming session conducted by the projects' experts. The list is then reduced to the most important risks. This step is sometimes subjective, but this issue is not relevant to this paper.
3. Risk Assessment is the quantification of identified risks, conducted in order to define priorities among the possible risk events. It usually includes the probability of the event and the severity of the damage. Later, the ranking of risks is based on these two parameters. One possible method is the Borda [3] methodology for ranking alternatives.

4. A Risk Response plan includes answers to the threats that are identified in the risk assessment phase. There are a number of ways to address these threats.
a. Avoidance - generate a course of action that eliminates the risk.
b. Transfer – transfer responsibility for the particular risk to a third party, either by utilizing insurance or, in the international arena, by forming treaties and international agreements.
c. Acceptance - a rational decision to accept a known risk without taking any action to prevent its outcome or deal with its consequences. The risk is usually dealt with when it is recognized as a risk. An acceptance of risk is recommended in situations where the consequences of the risk are less costly or less traumatic than the effort required preventing the risk.
d. Mitigation - refers to action taken to reduce either the probability of occurrence of an unfavorable event or the impact of this event. Mitigation is usually executed in the form of a plan designed to handle high-threat possible events.
e. Contingency Planning – refers to specific actions to be taken when a potential risk event occurs. In general, contingency plans should be developed in advance in preparation for the moment when risk events are realized.

Out of the five responses, only three (avoidance, transfer and mitigation) involve a real investment and require budget allocation.

5. A Control Plan is a series of course adjustments within the project's main objectives. These adjustments include scheduling and tracing the advance of risk situations. The control plan defines indicators that provide warnings regarding the realization of specific risks.

Continuously assessing program risks is the implementation of the control plan by checking any changes in the assessment of risks, and conducting a continuous search for warning signs that indicate any realization of known risks.

This part of the project plan includes the updating of the risk management plan.

The current study concentrates on allocating a budget to the response plan in an optimal manner.

This paper includes a literature review, problem definition, algorithms' definition, an example that is solved by all algorithms, and a comparison among the algorithms by simulation results. The research is quantitative and presents simulation results. Since the difference among the algorithms for different budgets are so big, statistical analysis is unnecessary.

The simulation and algorithms were verified by solving known problems and their solutions.

3. Literature review

Project risk management literature commonly describes the need to rank and prioritize project risks in order to focus the risk management effort on the higher risks. Baccarinia et al. in [4] describe the use of a methodology for the risk ranking of projects by some subjective judgment; this method has been implemented in construction projects and multi-project environments. Engert from MITRE [3] wrote a user's manual for an Excel application for risk management. The application includes a ranking method of risks based on Borda's method. The Borda method is more quantitative than the subjective judgment method, but still includes some fuzzy ranking when it combines the rank of risk probability with the rank of impact. Ochsner [5] emphasizes the limited attention to risk-based priorities and the growing consensus among industries that risk considerations need to be better integrated into decisions. He agrees that although money is not always the best way to measure risks, no better alternative has thus far been suggested. His ranking method is based on discussions with consultants and experts, assigning scores from 1 to 10 for each category. Li et al. [6] present a ranking method for multiple hazard risks; the method is based on screening all the risks with experts and weighting the risks according to frequency, severity, availability of warning, awareness, etc.

In [7] the author presents the difficulties involved in ranking risks. He utilizes the following framework: Risk = Threat×Vulnerability×Consequence, which is usually used in military operations research. The Threat×Vulnerability framework actually reflects the probability to damage a target, when the consequence is the damage impact. Our study is important in that for some qualitative measures, it presents counter examples that highlight the limitation of this measurement type. Klein [8] developed a conceptual model for analyzing alternative risk mitigation responses, while accounting for the possibility of trade-off risk among the three main success criteria: cost, duration and scope (or quality). He showed that, given the numerical estimates of risks probabilities and impacts, of all the relevant responses, mathematical techniques - such as dynamic programming or integer programming - could be applied to find the best combination of responses that minimizes project uncertainty. This approach analyzes trade-off among success criteria.

Ben-David et al. [9] analyzes a problem that is similar to the current one, but takes a different approach. Assuming that several risk mitigation responses can be implemented with different costs and different expected results, a selection of the best combination of responses is needed. All of the responses are broken down to their work elements, so that each risk can belong to several of them. The Total Risk Cost (TRC) is minimized by two heuristic algorithms; the greedy and the naïve, after which a comparison is presented. The current manuscript does not take into account the budget limitation, and assumes that as long as risk can be mitigated and it is worthwhile from the budget point of view – it will be done.

There are many studies that use subjective judgment to rank risks in different areas, industries, projects and programs. However, none of these ranking methods take into account *the response capability to risks*. There might be a huge difference between two risks that have the same probability to occur and the actual impact, when one of the risks occurs. However, for the first risk there is a mitigation plan that reduces its effect substantially and costs $1,000, while for the second risk, any type of mitigation plan costs more than $100,000. The study of Gonen et al. [10] proposes an additional criterion for the assessment of risks – that of *controllability*. The introduction of this criterion adds a third dimension to the risk evaluation process, in addition to its probability and impact. The controllability of a given risk reflects the ability to control it, mitigate it, or even prevent it. Assessing controllability may reduce the efforts and spending of managerial time and expenses on non-controllable risks and, in the end, direct the attention of management solely to controllable risks.

Controllability adds a new criterion that takes into account the response capability, but still does not propose a method to quantitatively rank the risks. In the current paper, we overcome the problem of ranking risks by utilizing a method of selecting the optimal mitigation plan for a given budget, and therefore, the risks to be mitigated or transferred.

Kutsch et al. [11] have investigated the type of risks that can be deliberately ignored. In the current study, we deal with risks that are not supposed to be ignored.

4. Defining the problem

The problem we will address in this study is the *allocation of a risk management budget among the possible responses*. The solution to this problem is not only ranking the risks to be dealt with, but also recommending the best risk response investment.

As was mentioned in Section 2, Part 4, responses to identified risks can be divided into two groups: Responses that include a real money investment - like transfer, avoidance and mitigation - and the other responses, which do not require any investment - like accepting the risk or preparing a contingency plan. Our study concentrates on the responses that require an investment and examines how to select the right set of responses when we are limited by a well-defined budget. In order to clarify these issues, let us look at the following theoretical example:

Assume there are two risks in a project - R_1 and R_2. R_1 will occur with probability P_1 and the damage in this case will be D_1. R_2 will occur with probability P_2 and the damage in this case will be D_2. In order to overcome these risks, we can either transfer the risk R_1 (by purchasing insurance), which will cost C_{11} and the policyholder's participation D_{11} or respond to risk R_1 with a mitigation plan that will cost C_{12}. After its application, the remaining probability to occur is P_{12} with damage when it occurs of D_{12}. For risk R_2, we have one mitigation plan that costs C_{21}; after its application, the remaining probability to occur is P_{21} with damage when it occurs of D_{21}. We have a risk mitigation budget of B that we can invest to handle these risks and we would like to know what our best policy is (B is usually determined by the project's customer).

In this study, we assume a linear utility function. This means that we will choose the policy that will reduce our expected cost to a minimum. The following table presents a numeric example of the dilemma described above:

Risk	P	D	Expected Damage	Response	Cost	P_{ij}	D_{ij}	Expected Damage+Cost
R1(*)	0.2	1000	200	Transfer	80	0.2	50	90
				Mitigation	50	0.1	500	100
R2(*)	0.3	700	210	Mitigation	50	0.3	200	110

Table 1. Numeric Example of the Dilemma

Let B=50. In this case, we can either choose the second or the third row. If we choose the second row, we reduce the expected damage of Risk 1 to 100 and stay with Risk 2 at an expected damage of 210. All together, the expected damage of both risks is 310. The same is true if we choose to handle Risk 2 and reduce the expected damage to 110. Since the expected damage of Risk 1 is 200, the total is 310. Let B=80. In this case, we can choose the first row or the previous option of B=50. Choosing Row 1 derives the total expected damage to 90+210=300 (the 210 is from R_2). If we choose the second row (mitigating R_1), our total expected damage will be 310, and the third row (mitigating R_2) will be the same - 310. However, in both mitigation plans we only invest 50, while in the transfer policy we invest a minimum of 80. People who are risk-averse will prefer this option, while others who are attracted to risk might prefer the second or third row. If B=100, then an additional option is open which allows us to choose Rows 2 and 3 and reduce the expected damage to 210. If B=130 and up, we can choose Rows 1 and 3 and reduce the expected damage to 200.

If we try to minimize the expected damage when B=80, then transferring R_1 would be optimal, although usually risk management methods will rank R_2 higher and recommend treating it first.

In order to define the optimal response problem, we will use the following terminology and symbolization:

There are n risks $R_1,\ldots,R_n$. For each risk R_i, the probability of its occurrence is P_i and the damage when it occurs is D_i. Therefore, for each risk R_i, the expected damage is $Q_i = P_i \cdot D_i$. Index *i will be used for risks.*

For each risk R_i, there are k responses (some can be empty; others can be transfer or mitigation) out of which we can choose, at most, **one**. This can be done by combining mitigation plans together. Index *j will be used for a response plan*.

The response j to risk R_i costs C_{ij}; after its implementation, the probability of its occurrence is P_{ij} and the corresponding damage is D_{ij}. The expected damage after its implementation is

$$Q_{ij} = P_{ij} \cdot D_{ij}.$$

A response plan is defined as “worthwhile” only if

$$Q_i \geq \left(C_{ij} + Q_{ij}\right) \; for \;\; i = 1, \ldots, n \tag{1}$$

(Only if the investment + the expected damage after the implementation are lower than the original expected damage). A response plan that is not worthwhile will not be included in the list of possible responses. Actually, the savings in selecting response j to risk R_i is:

$$Q_i - \left(C_{ij} + Q_{ij}\right).$$

Let us now define the decision variables X_{ij} as 1, if response j is selected for risk R_i, and 0, otherwise.

Only one response can be selected (if the user wants to enable selecting two responses to risk R_i, he can combine both responses into one plan with the accumulated cost). From the definition of X_{ij}, the expected value of all the risks will be:

$$\sum_{i=1}^{n}\left[(1-\sum_{j=1}^{k} X_{ij})Q_i + \sum_{j=1}^{k} X_{ij}(C_{ij}+Q_{ij})\right] \tag{2}$$

After opening the equation, it is clear that the expected value of all the risks (that we would like to minimize) is:

$$\sum_{i=1}^{n} Q_i - \sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}\left(Q_i - \left(C_{ij}+Q_{ij}\right)\right) \tag{3}$$

Since $\sum_{i=1}^{n} Q_i$ does not depend on the selection of risks to be handled, the problem can be defined as an integer programming problem, as follows:

$$max\left\{\sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}\left(Q_i - \left(C_{ij}+Q_{ij}\right)\right)\right\} \tag{4}$$

s.t.

$$\sum_{j=1}^{k} X_{ij} \leq 1 \qquad for\ i = 1, \dots, n \tag{5}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}C_{ij} \leq B \tag{6}$$

(budget constraint)

And X_{ij} can be either 0 or 1 for i=1,…,n and j=1,…, k.

5. Solving the problem

The problem can be solved by Integer Linear Programming (ILP), as was mentioned in [9, 12]. In this paper, we compare 3 heuristic algorithms that solve this ILP. The algorithms are as follows:

1. **The Most Dangerous Risk**_(MDR) method (PMI, 2008) is used to show the “naïve” solution. In the current case, the first risk to be handled is the one with maximum Q_i. For the selected risk, the most effective response is selected and the accumulated budget is increased by Cij.

For each selected risk, the response with the maximum savings (Qi-(Cij+Qij)) will always be selected. The algorithm that is used is as follows:

   1. Sort the risks according to Qi from higher to lower.
   2. For each risk, select the response j with the higher (Qi-(Cij+Qij)).
   3. Calculate the accumulated cost of applying the responses according to the sorted list.
   4. Calculate the accumulated savings.
   5. If the accumulated cost of risk responses is less than the budget, go back to Step 1.

2. **The Most Profitable Response** (MPR) method is defined as follows:
   1. Sort the responses according to (Qi-(Cij+Qij)) from higher to lower.
   2. Choose the upper risk in the sorted list that was not selected yet.

3. Calculate the accumulated cost of applying the responses according to the sorted list.
4. Calculate the accumulated savings.
5. If the accumulated cost of risk responses is less than the budget, go back to Step 1.

In this algorithm, the response savings plays a major role and the decision is made according to the possible savings.

3. **The Best Saving Ratio** (BSR) method is defined as follows:

Definition: The ratio between the savings in expected damage and the cost of the response will be called the **savings ratio.** Mathematically, it is defined as $(Q_i - (C_{ij} + Q_{ij}))/C_{ij}$. The economic meaning of this ratio is the amount of savings in expected damage per each unit of investment in the response.

The algorithm will be as follows:

1. Sort the responses according to the savings ratio (Qi-(Cij+Qij))/Cij from higher to lower.
2. Choose the upper risk in the sorted list that was not selected yet.
3. Calculate the accumulated cost of applying the responses according to the sorted list.
4. Calculate the accumulated savings.
5. If the accumulated cost of risk responses is less than the budget, go back to Step 1.

In order to clarify the three algorithms, let us demonstrate them by an example:

In the following table (Table 2) there are 6 risks; for each risk there are three possible response plans. The table includes the Pi, Di, Qi, Cij, Pij, Dij, Qij, and both the savings in expected damage + cost and the savings ratio.

Risk #	Pi	Di (in K $)	Qi	Response#	Rank	Cij	Pij	Dij	Qij	Cij+Qij	Qi-(Cij+Qij)	Xij	Saving Ratio
1	0.29	133.37	39.20	1	4	35.81	0.02	118.17	2.70	38.50	0.70	X11	0.02
	0.29	133.37	39.20	2	4	20.02	0.08	32.41	2.54	22.57	16.64	X12	0.83
	0.29	133.37	39.20	3	4	0.67	0.04	116.92	4.20	4.87	34.34	X13	51.35
2	0.85	170.91	144.68	1	1	31.06	0.25	70.48	17.31	48.37	96.31	X21	3.10
	0.85	170.91	144.68	2	1	22.28	0.74	149.28	110.87	133.15	11.53	X22	0.52
	0.85	170.91	144.68	3	1	120.89	0.11	137.49	14.88	135.77	8.91	X23	0.07
3	0.83	155.09	129.23	1	2	25.79	0.65	136.71	89.54	115.33	13.90	X31	0.54
	0.83	155.09	129.23	2	2	14.02	0.73	77.06	56.45	70.47	58.76	X32	4.19
	0.83	155.09	129.23	3	2	4.11	0.80	117.19	93.55	97.66	31.57	X33	7.68
4	0.83	19.44	16.10	1	6	12.21	0.54	6.03	3.25	15.46	0.64	X41	0.05
	0.83	19.44	16.10	2	6	4.83	0.03	8.48	0.28	5.11	10.99	X42	2.27
	0.83	19.44	16.10	3	6	3.85	0.81	2.29	1.84	5.69	10.41	X43	2.71
5	0.19	168.04	31.15	1	5	13.03	0.00	11.51	0.02	13.05	18.11	X51	1.39
	0.19	168.04	31.15	2	5	8.00	0.12	30.92	3.76	11.76	19.40	X52	2.43
	0.19	168.04	31.15	3	5	1.54	0.06	111.67	7.21	8.76	22.40	X53	14.51
6	0.58	101.83	58.83	1	3	46.87	0.00	2.99	0.01	46.89	11.94	X61	0.25
	0.58	101.83	58.83	2	3	37.98	0.30	23.31	6.90	44.88	13.94	X62	0.37
	0.58	101.83	58.83	3	3	6.46	0.56	86.35	48.06	54.52	4.30	X63	0.67

Table 2. Numeric Example to compare the three algorithms

The numeric example is generated by a simulation that will be described later. Table 2 includes all the information needed for applying the algorithms MDR, MPR and BSR.

Tables 3, 4, 5 present the MDR, MPR and BSR solutions accordingly.

In Table 3, the ranked risk =1 means the first risk to respond. The first risk that is handled is Risk number 2, since its Qi is 145 (from Table 2). The response is selected as the highest savings solution. Total handling of the 6 risks requires a budget of 90.1 and saves 236.7 in expected damages, plus the cost of applying the responses.

Table 4 shows that the selection order is different from MDR. However, the accumulated savings converges to the same amount, since at the end both algorithms use the same response plans. The difference is in the selection order.

Most Dangerous Risk (MDR)						
Ranked Risk	Risk Number	Response Number	Cost	Accumulated Budget	Savings	Accumulated Savings
1	2	1	31.06	31.06	96.31	96.31
2	3	2	14.02	45.08	58.76	155.07
3	6	2	37.98	83.06	13.94	169.01
4	1	3	0.67	83.72	34.34	203.35
5	5	3	1.54	85.27	22.40	225.75
6	4	2	4.83	90.10	10.99	236.73

Table 3. Solution of the example using the MDR algorithm

Most Profitable Response (MPR)						
Ranked Risk	Risk Number	Response Number	Cost	Accumulated Budget	Savings	Accumulated Savings
1	2	1	31.06	31.06	96.31	96.31
2	3	2	14.02	45.08	58.76	155.07
3	1	3	0.67	45.75	34.34	189.40
4	5	3	1.54	47.29	22.40	211.80
5	6	2	37.98	85.27	13.94	225.75
6	4	2	4.83	90.10	10.99	236.73

Table 4. Solution of the example using the MPR algorithm

Table 5 shows that the BSR uses different response options and therefore converges to different accumulated savings. In this example, the BSR is the worst option out of the 3 algorithms, although this result does not represent the most common situation, as will be seen later.

Best Savings Ratio (BSR)						
Ranked Risk	Risk Number	Response Number	Budget	Accumulated Budget	Savings	Accumulated Savings
1	1	3	0.67	0.67	34.34	34.34
2	5	3	1.54	2.21	22.40	56.73
3	3	3	4.11	6.32	31.57	88.30
4	2	1	31.06	37.38	96.31	184.61
5	4	3	3.85	41.22	10.41	195.02
6	6	3	6.46	47.69	4.30	199.32

Table 5. Solution of the example using the BSR algorithm

6. Comparison of the three algorithms

In order to compare the three algorithms, a scenario simulation was generated with 15 risks and 3 responses per risk. The simulation draws the probabilities and damages according to the following rules:

1. Draw P_i distributed U(0.01.0.9) (uniform between 0.01 and 0.9)
2. Draw D_i distributed U(10,200)
3. Draw P_{ij} distributed U(0,P_i)
4. Draw D_{ij} distributed U(0,Di)
5. Draw C_{ij} distributed U(0.1, Q_i-Q_{ij}) where $Q_i = P_i \cdot D_i$ and $Q_{ij}=P_{ij} \cdot Dij$

For all i=1,….15, j=1,…,3

The following chart (Figure 1) shows an example of the behavior of the three algorithms, while the budget increases, step by step.

Figure 1 is an example of a typical situation in which, for a limited budget the BSR is the best algorithm, while for an unlimited budget, the other algorithms can produce better results. This phenomenon holds in most of the simulation examples, but there are cases where the BSR is better for all budgets and cases.

In order to compare the three algorithms, 100 simulations were generated. For each simulation, the maximum needed budget was calculated. (Since the Cij are drawn, the required budget is stochastic and different in each simulation). For each simulation, the savings was calculated for an investment of 20%, 40%, 60%, 80% and 100% of the budget.

For each percentage investment, the savings was calculated for each algorithm. Later, the best algorithm was defined as the successor, for each specific budget, and the frequency of its success was calculated. The following table (Table 6) summarizes the number of successes of each algorithm

Table 6 shows that for a low budget (20 to 60 percent) the BSR is the best algorithm, while for an unlimited budget the MPR behaves better. In many cases, the MPR and MDR behave the same and reach the same savings.

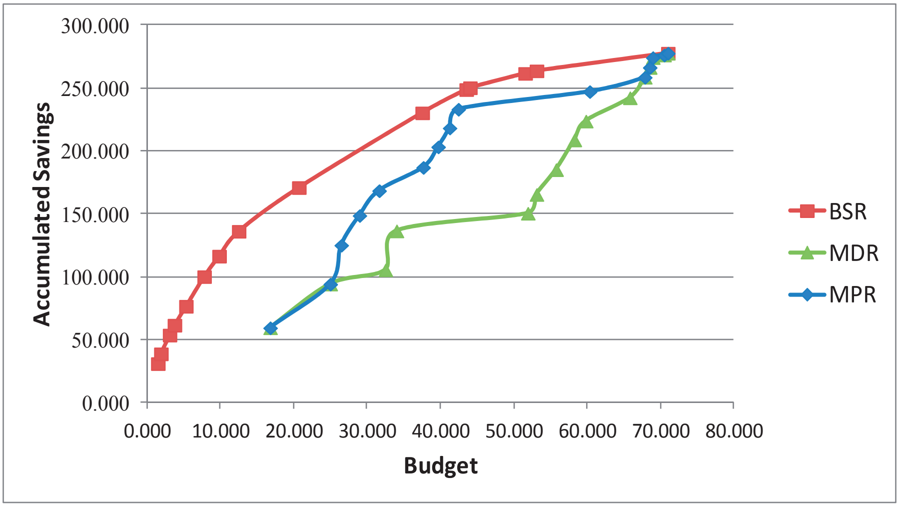


Figure 1. The savings of each algorithm regarding the limited budget

The main conclusion from Table 6 is that there is no optimal heuristic algorithm. Moreover, if only part of the risks budget can be handled, it is recommended to use the BSR algorithm.

	Percentage of budget investment				
Algorithm	20%	40%	60%	80%	100%
MDR	0	0	0	0	1
MPR	4	3	7	23	97
BSR	96	97	93	77	2

Table 6. Distribution of success of each algorithm

7. Discussion

1. Findings

The paper presented three heuristic algorithms for risk response selection. In many cases, the ranking of risks is not enough for project managers and they need to know how to invest their risk management budget among the possible responses. We observe that for a limited budget the BSR algorithm is better than the MDR or MPR method, while for a budget that can cover all the risks, the MDR or MPR are better. Currently, in most projects, the customer asks to see the risk management plan. The above method adds the selection method of risks to be mitigated. It should be an essential part of the risk management plan.

A stronger result is that *risk ranking is no longer needed*. This saves the effort of ranking risks, which is usually subjective.

2. Limitations

One limitation of the current paper is that estimating the probabilities and damages for each risk and response is usually considered to be a very difficult task. However, it is required by most of the risk management standards. Tools, like mathematical models and simulations, are available for this task and there are already many projects that include these estimations.

Another limitation is that we assume that responses with a negative expected savings cannot be selected. However, in reality, there are responses, like insurance, that are based on negative expected savings (otherwise, insurance company would not sell insurance policies).

A third limitation is the dependencies among risks. It might be that a delay in one task is not critical, while a delay in a second task, together with delay in the first task, might prove to be a severe problem.

8. Conclusion

In this article, we describe a method for how to allocate a risk management budget among the possible mitigation or transfer plans. In most of today's literature, the risk management plan usually ranks the risks and recommends handling those with high rankings. Almost no consideration is given to either response plans or response feasibility. This study proposes three heuristic algorithm approaches to budget allocation, and demonstrates the method, including a sensitivity analysis of the budget constraint. The results are encouraging and help define rules about risk management budgeting.

The model is based on the expected damage, and assumes we will always prefer to reduce expected damages plus their cost. It does not discuss the question of risk taking.

A simulated scenario with 15 risks and 45 response plans was demonstrated. The most important lesson learned from the example tested in the study is that the solution is mainly influenced by the response plan, and not only by the expected damage of the risk, as all of the ranking methods recommend. Moreover, for a limited budget, the BSR is usually the best algorithm, while for an unlimited budget the MDR or MPR algorithms are more preferable..

Author details

A. Gonen
Faculty of Management of Technology, HIT - Holon Institute of Technology, Holon, Israel

9. References

[1] Wideman R.M. 1992. Project and Program Risk Management: A guide to managing project risks and opportunities. Project Management Institute.
[2] PMI Standards Committee, 2008. A Guide to the Project Management Body of Knowledge. Project Management Institute. Newtown Square, PA.

[3] Engert P.A., Lansdowne Z. F. 1999. Risk Matrix User's Guide, Version 2.2, MITRE Bedford, Massachusetts.

[4] Baccarinia D., Archerb R., 2001. The risk ranking of projects: A methodology, International Journal of Project Management, Volume 19, Issue 3, Pages 139-145.

[5] Ochsner M., 2000. Case study: Risk prioritization and ISO 14001 at Acushnet Rubber Company, Environmental Quality Management.

[6] Li H., Apostolakis G.E., Gifun J., VanSchalkwyk W., Leite S., Barber D. 2009. Ranking the Risks from Multiple Hazards in a Small Community, Risk Analysis, Vol. 29, Issue 3, pp. 438–456.

[7] Cox L. A. 2008. Some Limitations of "Risk = Threat × Vulnerability × Consequence" for Risk Analysis of Terrorist Attacks, Risk Analysis, Volume 28, Issue 6, pages 1,749–1,761.

[8] Klein J. H. 1993. Modelling Risk Trade-Off, Journal of The Operational Research Society, Vol 44, No. 5, pp. 445-460.

[9] Ben David I., Raz T, 2001. An Integrated Approach for Risk Response Development in Project Planning, Journal of The Operational Research Society, Vol 52, pp. 14-25.

[10] Gonen A., Zeitouni N. 2010. Risk management of water resources in a changing climate in: Advances in Risk Management ed. Giancarlo Nota Ch.10, 199-222 Publishers: Sciyo.

[11] Kutsch E., Hall M., 2009. Deliberate ignorance in project risk management, International Journal of Project Management , Vol. 28, 245-255.

[12] Ben David I., Rabinowitz G., Raz T, 2004. Economic Optimization of Project Risk Management Efforts, Internet Working Paper. Available at http://www.fisher.osu.edu/fin/faculty/Ben-David/articles/BDRR2002.pdf

Chapter 4

A Framework to Select Techniques Supporting Project Risk Management

Sabrina Grimaldi, Carlo Rafele and Anna Corinna Cagliano

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50991

1. Introduction

Projects may be conceived as temporary endeavors with a finite completion date aimed at generating unique products or services [1]. Today's marketplace characterised by fierce competition requires increased accuracy and reduced time and costs in running projects [2]. In such a context, the variability of actual quality, time, and cost performance compared to the expected one crucially impacts on the success of a project and makes risk a central issue in project management [3]. It has been demonstrated that failure to deal with risk is a main cause of budget exceeding, falling behind schedule, and missing performance targets [4,5]. Additionally, in several industries, such as the construction and information and communication technology ones, the growing level of complexity, due to increased size and scope, huger investments, longer execution processes, more required resources, an augmented number of stakeholders, instable economic and political environments, and changing regulations, exacerbates the degree of risk in projects [6]. Therefore, these aspects ask for assessing and controlling risk throughout all the phases of a project. Before going into detail about project risk management, it is beneficial to recall the notions of uncertainty and risk. Uncertainty arises from either the natural variability or randomness of a system or an incomplete information or knowledge of some of its characteristics. In the first instance, uncertainty cannot be reduced by increasing data collection or knowledge, though they are valuable for assessing it, while in the second case a more accurate data collection and understanding are able to decrease the level of uncertainty [7-9]. Project risk is defined as an uncertain event or condition that, if it occurs, has either a positive or a negative effect on project objectives [1,10].

The management of risk is currently one of the main topics of interest for researchers and practitioners working in the field of project management. Different perceptions, attitudes, values regarding risk, needs, project sectors, specifications, geographical, social, economic, and political environments have led to a variety of definitions, concepts, terms, and approaches, all highlighting the need for systematically addressing uncertainty.

Since the Nineties, most of the contributions have focused on the establishment of a risk management process: significant examples are the Project Uncertainty MAnagement (PUMA) process [11],the Multi-Party Risk Management Process (MRMP)[12],the Shape, Harness and Manage Project Uncertainty (SHAMPU) process [13], the Two-Pillar Risk Management (TPRM) process [14],the risk management process developed by the Project Management Institute [1], the Project Risk Analysis and Management (PRAM) process [15], the Risk Analysis and Management for Projects (RAMP) process [16], and The Active Threat and Opportunity Management (ATOM) Risk Process [10].

An effective application of risk management processes is not disjointed from sound enabling instruments. So, another research stream is running parallel to that focusing on the overall risk management structure: the development, implementation, and evaluation of operational means to put in practice risk management [17].

However, in literature there is a scarce systematisation of the actual capabilities of such practices. In addition, there is a lack of frameworks categorising them based on a comprehensive set of the peculiar characteristics of a project, of its management process, and of its surrounding business environment, as well as on the attitude of an organisation towards risk.

In order to contribute to fill this gap, the present work puts forward a taxonomy supporting the selection of the most suitable risk management techniques in any given project scenario, with the aim of fostering knowledge creation about how to treat risky events. The research mainly focuses on projects characterised by the achievement of a final work product not completely defined at the beginning of the project itself, such as in the construction, engineering, and information and communication technology industries.

After discussing the value of communication and knowledge in risk management, a set of dimensions reflecting the most important managerial and operational conditions characterising a project is developed starting from a review of pertinent literature. Widely applied techniques to support project risk management are presented and classified according to the framework. Finally, implications, ramifications, and future research directions are elaborated and conclusions drawn.

2. Communication and knowledge creation in risk management

Identifying and assessing risk sources and their impacts on project activities as well as developing responses to risk rely on a heterogeneous knowledge basis made up of past experiences, skills, and perspectives of involved people. However managing data, information, and in general the knowledge generated during the life cycle of a project is a difficult task and an inappropriate way of doing that may be a cause of failure. In particular, communication about risk is often very poor, even if the interactive process of exchanging information and opinions among all the concerned parties is a critical condition in the risk management process to effectively support decision-making [18]. Projects are often organised and managed in ways that create information and communication disconnects. Decisions about risk are made independently from one another according to the different nature of possible risky events (e.g. business, technical, operation, and country-specific) and the interactions among them are not taken into account. Participants in a project do not share a comprehensive understanding of the risks that may affect it and a life cycle view of uncertainty is usually uncommon. This brings compartmentalisation of risks because they are identified, assessed, and controlled by using only one perspective [19]. A structured communication of the objectives, instruments, and findings of the risk management process as well as of the required actions as a result of its output is strongly needed, being organisational and individual learning increasingly important when dealing with risk [20].

Communication among project parties generates awareness of risk and supports knowledge creation about both drivers and effects of uncertainty and approaches to cope with it.

A variety of practices exists to deepen the understanding of causes and consequences of uncertainty [4,21-23]. However, their application is still limited because several organisations do not systematically track past data and performance for this purpose. When there is a substantial lack of explicit information an important source of knowledge is represented by the implicit information held by the so called "experts". The term expert refers to those people to whom special knowledge about specific issues is attributed and from whom it is possible to obtain information that is useful for risk investigation. The process of extracting information from experts is named elicitation, which is defined as formulating a person's knowledge and beliefs about one or more uncertain quantities into a probability distribution for these quantities [24]. Elicitation of implicit expert knowledge is a core component of qualitative risk assessment, by means for instance of Delphi analysis or SWOT analysis, where it is used to define probability distributions for the occurrence and the impact of risky events.

Another relevant issue in knowledge creation about risk is related to the guidelines on how to approach it. As mentioned, literature offers a wide range of frameworks to identify risk sources, evaluate their probabilities and impacts in both a qualitative and a quantitative way, and set up risk response strategies. Also, there are some attempts to categorise these practices according to the nature of the data they rely on, the phase of the risk management process, the kind of project, or the purpose of the analysis [1,25-27]. However, existing contributions usually focus on just one single aspect and there is a lack of taxonomies that simultaneously look at all the relevant dimensions that should be taken into account when choosing an appropriate means of treating risk. In addition, the terminology used to address risk management practices is somewhat confused. The most common words that can be found in literature are tool, technique, and method but there is no widely accepted definition of these concepts and of the relationships among them in the field of risk management. Sometimes a same practice is referred to with different terms. For instance, while Delphi is generally classified as a technique [1,26], the Failure Mode and Effects Analysis (FMEA) is defined as either a tool [4] or a method [25]. However, determining the exact nature of risk instruments and creating a hierarchy among them help to recognise their scope and range of application and allow a more appropriate use at various risk management levels.

How to select the correct practices and capture their actual potentialities is of paramount importance to enhance the knowledge that is necessary to manage in an effective and efficient manner the risk and the associated information throughout the development of a project. Such understanding facilitates a clear view of the critical conditions of a project, thus fostering performance improvement and enhancing trust within the project team [28].

The developed framework focuses on the need for a comprehensive perspective on the factors affecting risk investigation and proposes a taxonomy based on the most significant elements characterising the scenario in which project risk is approached. The aim is assisting in the choice of the appropriate practices according to the level and the purpose of the risk management effort. Since the distinction among the different terms to address risk management practices is not the purpose of this work, they are all referred to as "techniques".

3. Dimensions for selecting techniques to support project risk management

There are multiple aspects that can be considered when facing the decision about the appropriate techniques to be applied for the purpose of risk identification, assessment, or control. They will be widely explained in the following sections.

3.1. A review of classification criteria

A commonly used criterion suggests looking at the nature of information that is available in a project. Qualitative and quantitative techniques are two fundamental groups applied to risk management. In the qualitative techniques risk assessment is connected with the determination of qualitative scales for evaluating the frequencies of occurrence of risky events and their impacts. They do not operate on numerical data but present results in the form of descriptions and recommendations basically according to opinions and risk tolerance boundaries collected from experts. The qualitative techniques are adopted to prioritise the identified risks for subsequent further action, such as quantitative risk analysis or response planning [1]. Moreover, they are used for determining highly risky areas in a short time, cheaply, and easily. At the other hand of the spectrum, quantitative techniques to support project risk management numerically analyse the effects of risks on overall project objectives in order to elaborate future trends [1,29]. They are applied to give an accurate image of risk that facilitates the cost and benefit analysis during the selection of reduction measures. However, the implementation of quantitative techniques is generally more expensive and requires greater experience than the application of qualitative techniques [30].

Another criterion is choosing techniques to support risk management according to the degree of knowledge about risk and the goal of the analysis. Kmec [27] discusses approaches to risk identification for the following situations: the majority of risks are known, the risks have been prioritised, the risk list is short, risks are classified according to some criteria, risks are broken down to build a hierarchy, relationships among risk are investigated, and risk evolution is studied overtime. Also, techniques for risk management differ according to whether the main aim is monitoring economic and financial outcomes, checking quality variance, tracking time delays or estimating the probability of the overall success or failure of a project.

In addition, risk management practices can be distinguished based on how the investigation is performed. Gidel and Zonghero [31] focus on selected techniques and suggest when they are suitable depending whether an analogical, heuristic, or analytic approach is applied to risk identification. With an analogical approach the study of risk mainly relies on the experience coming from the management of previous and similar projects. The heuristic approach uses the project team creativity or expertise through for instance brainstorming sessions. Finally, the analytic approach is typically based on FMEA and Fault Tree Analysis and aims to decompose a system to identify risky events for each sub-system together with their causes and effects.

Also, the nature, size, and phase of the life cycle of a project as well as the kind of associated consequences determine which techniques to support risk management should be used. Some authors highlight that, although risk management should assist in the entire life cycle of a project, it is particularly crucial in the planning stage and its scope and depth increase as the project moves towards the execution phase, while they decrease in the termination phase [13,32]. As a matter of fact, the earlier the risks are identified, the more realistic the project plan and the expectation of results and the more effective the contingency plans both during the development of the project and beyond [1,33].

Other works focus on the strong correlation between the risk profile of a project and its organisation: for instance, different procurement schemes require different risk practices [22].

Furthermore, every single step of managing risks, whether identifying or assessing them, developing response plans, or monitoring their execution, implies a different level of information and detail, thus it requires the application of different techniques. Literature reports numerous classifications of techniques according to the phase of risk management for which they are most suitable [1,34,35].

Finally, the project risk management capabilities of an organisation improve as its risk culture increases. A scarce awareness towards risk drives occasional applications of informal risk techniques to specific projects and problems are dealt with only when they show up. Recognising the relevance of risk, instead, is the condition for proactively managing uncertainty [33,36,37]. As a consequence, techniques supporting risk management require different levels of corporate risk maturity in order to yield the expected benefits and this constitutes a criterion according to which risk techniques may be classified [25].

3.2. Three dimensions to characterise project risk management techniques

Based on a careful analysis of the characteristics of the techniques supporting risk management proposed in literature and applied in business practice, the authors believe that among the discussed criteria

- the phase of the risk management process;
- the phase of the life cycle of a project;
- the corporate maturity towards risk;

are the three dimensions that encompass the most relevant aspects for understanding and choosing among project risk management techniques. In fact, the focus is on “risks” that occur in “projects” which are in turn run by “companies”. Moreover, such dimensions adequately reflect the crucial concept that risk practices can only be selected once a problem is structured and well understood and the application of these instruments depends on the circumstances of the problem, hence on the need to fully comprehend it.

Every specific risky event in a project has its own escalation process characterised by one or more sources or causes, an occurrence, and one or more consequences [35]. Each of these phases requires its own approach to be studied. Sources of risk are analysed by concentrating on their identification, description, and classification (e.g. internal and external causes), the occurrence is defined by the probability and the impact of the risky event, and the consequences are described in terms of time, cost, and quality variance against the expected performance.

Additionally, no practice is perfectly tailored to deal with every risk occurring in the course of a project [22]. Each of the risks faced during a project has its own specificity depending on its position within the project life cycle. For example, throughout the feasibility study, when the main issue is making appropriate strategic choices, the probabilities of occurrence of risks are difficult to be defined because of the still scarce level of information associated with that phase. By contrast, in the following phases risks are mainly related to the consequences of decisions made in the previous steps of the project and their sources, manifestation, and effects can be characterised in a more accurate way. Also, in the late phases of a project a risk may be the effect of other risks that manifested themselves in previous phases.

Besides the phases of the risk management process and the life cycle of a project, a third pillar constitutes the foundation of a sound selection of techniques supporting risk treatment: the reference context of the organisation that develops a project. In particular, this work is interested in the maturity towards risk, that is basically achieved through risk awareness, the consideration that the risk management activity is on the same level as cost, time, and scope management tasks, commitment to high quality of data, systematic implementation of instrument to deal with risk, development of responses to risk, and assessment of the obtained results [38]. The extent to which a company possesses these features represents that cultural bedrock that enables the application of specific techniques to prevent, accept, mitigate or exploit risky events and their effects. In particular, a high level of risk awareness, together with appropriate availability of knowledge, make possible to obtain that objective information allowing the quantification of risk.

A selection of support techniques based on the above dimensions represents a strength inside the risk management process because it stimulates the achievement of improved outcomes in terms of time, cost, and quality performance [39].

3.3. Phases of the risk management process

According to Hillson [40], risk management is about finding an answer to six simple questions such as "What do we want to achieve?", "What might affect us?", "Which of the things that might affect us are most important?", "What should we do about them?", "Did our actions work?", and "What has changed in the new scenario?". These questions represent the main issues of the risk management process, which is generally recognised as the process concerned with conducting the following phases: risk management planning, risk identification, risk analysis, risk response, and risk monitoring and control [1].

In risk management planning the objectives and the approach to carry out risk treatment tasks are decided together with assigning resources and time to these activities, with the aim of allowing a smooth conduction of the subsequent phases. Risk identification defines the risks to which the project is exposed and describes their causes and characteristics. The goal of the risk analysis phase, sometimes named risk assessment, is giving an importance priority to the identified risks to enable managerial actions and establishing the overall level of risk exposure of the project. In particular, qualitative risk analysis is focused on determining the probabilities of occurrence of risky events and the associated impacts on project outcomes, the time periods when the risks could affect the project, when it is possible to influence them, and the relationships between risks and cost, schedule, scope, and quality constraints. Quantitative risk analysis operates on those risks that substantially impact the project and numerically evaluates their effects. Risk response starts from the previously identified risks and their significance to develop actions to increase opportunities and decrease threats. Resources and activities are inserted into the budget, schedule, and project management plans. The final phase, risk monitoring and control, is the on-going identification and management of new risks that become known during a project, the tracking of already identified risks, the monitoring of residual risks, the implementation of planned responses as well as the review of their effectiveness, the development of additional actions, if needed, and the formalisation of lessons learned about risk [1,35].

The importance of the dimension of the risk management process phases for selecting techniques to support the treatment of risk is witnessed by the many works discussing instruments for each phase existing in literature. Some of them have been already presented in Section 3.1.

3.4. Phases of the project life cycle

In a similar way as when the risk management process is approached, undertaking a project means tackling some basic questions: "Who are the parties ultimately involved?", "What do the parties want to achieve?", "What is it the parties are interested in?", "How is it to be done?", "What resources are required?", and "When does it have to be done?". These questions are answered during the life cycle of a project, which is defined as a systematic way of conceptualising the generic structures of projects into a number of phases that assure better management control [1,13,41].

The project life cycle is domain specific and, because of the complexity and diversity of projects, its breakdown into phases is different based on several factors such as the size (e.g. small or large-scale projects) and the type (e.g. engineering and construction projects or new product development projects) of the project. Four general phases can be associated to the kinds of projects that are considered by this work: conceptualisation, planning, execution, and termination [1,13].The conceptualisation phase regards identifying an opportunity or a need, clarifying the purpose of the project by defining the relevant performance objectives and their importance, formalising the concept of the project, and evaluating its feasibility. The planning phase includes undertaking the basic design, developing performance criteria, formulating a base plan together with targets and milestones, and allocating internal and external resources to achieve the plan. With the execution step of a project action begins: the main tasks here are coordinating and controlling the performing of planned activities, monitoring progress, and changing targets, milestones, and resource allocation as required. Finally, the termination phase involves commissioning and handover, reviewing the lessons learned during the project, and assuring the necessary support to the product of the project until it is discarded or disposed.

It is widely recognised that a structured view of the project life cycle provides a proper frame for understanding major sources of uncertainty, as well as their occurrence timing and impacts, during all its phases [13]. Also, the project life cycle is a natural setting for distinguishing among approaches to risk management. As the life cycle evolves, different information becomes available about the aspects and components of both a project and its environment, such as stakeholders, scope, time, and cost as well as corresponding assumptions and constraints. Therefore, there are more risks at the beginning of a project, while they decrease as the project progresses towards its termination. As a consequence, the greatest opportunity to risk reduction resides in the early project stages. In general, during the conceptualisation phase, decision makers should focus on different sources of uncertainty, such as technological, cultural, social, and economical ones, to make sure about the feasibility of the project [42]. The identified uncertainties should be then taken into account during the planning phase of the project. The risk management process should monitor the changes as well as the new risks emerging in the execution phase and manage the appropriate actions to reduce or eliminate them [1]. Finally, the typical risks in the termination phase are related to the proper maintenance, improvement, and changing needs in light of evolving societal, demographic, operational, or economic conditions.

Since the sources of uncertainty change during the project life cycle, it is vital to understand how the risk management process has to vary accordingly. This consideration supports the need to enable project managers to focus on specific sources of uncertainty in each stage of the project by means of appropriate practices to identify, assess, and treat such uncertainty in order to optimise its impacts. In addition, a project life cycle-oriented view of risk management techniques helps to avoid compartmentalisation in approaching risk, which occurs when each participant looks at risks with a single, specific perspective and based on his own goals, irrespective of the other project parties [19].

3.5. Corporate maturity towards risk

The concept of maturity indicates an evolution from an initial state to a more advanced one through multiple intermediate states corresponding to different levels of awareness towards risk and capability to deal with it. The degree of maturity towards risk of an organisation depends on its risk culture, which is stimulated by the available informational context and the type and size of the organisation itself. All these factors also impact on the maturity of the project management process, that may go from basic project management, to the systematic planning and control of a single project, to the integrated planning and control of multiple projects, to the continuous improvement of the project management process [43], which in turn influences how risk management is applied.

Hillson [37] proposes a risk maturity model made up of four stages: Naïve, Novice, Normalised, and Natural. Naïve means that an organisation has not yet captured the need for managing risks and no structured approach is in place for this purpose. Novice defines an organisation that recognises the benefits of managing risk and is actually implementing some form of risk governance but it lacks a formalised process to perform this task. Normalised is the degree of maturity characterised by a formalised risk process included in routine business activities whose benefits, however, are not consistently achieved in every project. Finally, the Natural maturity level denotes an organisation that is completely aware of risk and proactively manages opportunities and threats through consistent risk information. A similar organisation will benefit from improved corporate planning, more transparent relationships with stakeholders, and better global performance [44].

Moving from one level to the upper one in this maturity scale implies that an organisation is willing to perform a more thorough and systemic analysis of the escalation processes of project risks. In order to do that, not only different but also more sophisticated and detailed techniques have to be applied [33,38]. Based on this, it can be stated that the more mature is an organisation towards risk, the more the phases of the risk management process it will implement. Companies with a low maturity degree only limit themselves to risk identification or qualitative risk analysis, while organizations with a higher level of maturity deal with all the stages of the risk management process, including collecting past data to carry out quantitative analysis. Thus, the maturity of a company towards risk and its response to possible consequences are strictly related to the development of the risk management phases.

4. Classifying techniques supporting project risk management

The three defined dimensions characterising the choice of project risk management techniques are here applied to a selection of practices that can be commonly found in both literature and practice.

First, the focus techniques are briefly described and their strengths and weaknesses highlighted (Table 1).

No.	Technique	Description	Strengths	Weaknesses
1	**Brainstorming** [1,13]	An effective way to generate lots of ideas on a specific issue and then determine which idea–or ideas–is/are the best possible solution. Ideas about project risk are generated under the leadership of a facilitator.	• Improves problem analysis by providing more possible solutions and unusual approaches to a problem. • Increases the chances of obtaining an excellent idea. • Involvement of individuals with a variety of backgrounds. • Utilises the thoughts of others. • Attempts to view situations from an unfamiliar perspective.	• Prone to the negative effects of personality excesses. • Difficult to create a criticism-free atmosphere. • Not much structured. • The smaller problems that can have severe consequences on the project success are not identified. • Reduced participation due to dominant personalities. • Inhibited participation due to inequalities in expertise [13].
2	**Cause and effect diagram** **or** **Cause Consequence Analysis (CCA)** [1]	It identifies the set of unwanted effects and goes backwards to trace the causal chain. It is also known as Ishikawa or fishbone diagram and is useful for identifying causes of risks.	• Helps to determine the root causes of a problem or of a quality characteristic in a structured way. • Increases knowledge of a process by helping everyone to learn more about the relevant factors and how they relate to each other.	• Not particularly useful for extremely complex problems where many causes and problems are interrelated.

No.	Technique	Description	Strengths	Weaknesses
3	**Change Analysis (ChA)** [18]	It is used to systematically investigate the possible risks and to identify the appropriate risk management strategies and measures in changing situations.	• Predictive and proactive risk analysis technique. • Can be used as a root cause analysis. [18]	• Relies on the comparisons between two or more systems or activities. • Does not traditionally involve the quantification of risk. • Depends very much on expert judgements. • Limited to the analysis of system changes. [18]
4	**Checklist** [1,13,20]	It is a detailed aide-memoire for the identification of potential risks. It can be developed based on historical information and knowledge that have been accumulated from previous similar projects.	• Systematically assesses the experience accumulated by an industry. • Can be prepared by a single analyst or a small group. • Uses high-level or detailed analysis [18]. • Simple to use at the basic level. • Useful as a memory jogger. • A guide to the existing risk and opportunity knowledge.	• Limited to previous experience only. • Traditionally it only provides qualitative information [18]. • Individual technique. • Useful only for the early stages of the selection of an idea. • Risk drivers are assumed to be independent. • Can become intimidating. • Length may discourage a more selective analysis of a subset of risk drivers.
5	**Decision Tree Analysis** [32]	It is usually structured using a decision tree diagram that describes a situation and the implications of each of the	• Many application possibilities in different areas. • Enables a detailed	• Must be careful when assigning probabilities. • Individual

No.	Technique	Description	Strengths	Weaknesses
		available choices and possible scenarios. It incorporates the cost of each available choice, the probabilities of each possible scenario, and the rewards of each logical path.	insight into the decision making process. • Appropriate for solving complex problems. • Often supported by statistics. • Can be computer assisted.	technique.
6	**Delphi** [1]	The purpose is to elicit information and judgments from participants to facilitate problem-solving, planning, and decision-making. A facilitator uses a questionnaire to solicit ideas about the important project risks and the experts participate anonymously.	• Group technique. • Mainly used as a forecasting technique. • Helps to reduce bias. • Keeps any person from having undue influence on the outcome. • Elimination of direct social contact. • Provision of feedbacks. • Opportunity to revise opinions.	• Very complex. • The quality of results depends on the competencies of experts and on the content of the questionnaire. • Time consuming and expensive. • No opportunity for verbal clarification or comment. • Conflicts not resolved.
7	**Event and Causal Factor Charting (ECFCh)** [18]	It consists of a graphical description of the sequence of events and conditions associated with an accident. The chart provides a logical progression of events.	• An effective technique for understanding the sequence of contributing events [18].	• Does not necessarily ensure that the root causes have been identified. • Can overwork simple problems that may not require an extensive investigation [18].
8	**Event Tree Analysis (ETA)** [18]	It is an analysis technique that models the range of possible outcomes of one or a category of initiating events.	• Highly effective in determining how various initiating events can result in accidents.	• Usually limited to one initiating event; multiple event trees may be needed. • Dependencies

No.	Technique	Description	Strengths	Weaknesses
			• Shares similar strengths with Fault Tree Analysis [18].	among system elements can be overlooked [18].
9	**Expected Monetary Value (EMV)** [1]	The EMV analysis is a statistical concept that calculates the average outcome when the future includes scenarios that may or may not happen.	• The EMV of opportunities is generally expressed as a positive value, while that of risks as a negative value.	• Requires a great availability of historical data.
10	**Expert Judgement** [1]	Technique based on the experts' opinion. It is useful for the evaluation of the failure rate and the success chances of the overall project.	• Uses experiences on past projects to assess factors about a new project. • Adapt to exceptional circumstances.	• The estimation can be biased. • No better results than those provided by the expertise of estimators. • May be repeated multiple times in order to get more accurate information.
11	**Fault Tree Analysis (FTA)** [45]	An approach that starts from a particular event, known as the top event, in an attempt to identify all the possible event sequences giving rise to it.	• Highly effective in determining combinations of events and failures. • Systematic, logical, and detailed system approach. • Applicable to any kind of complicated system or activity. • Quantification is possible [18].	• Usually employed to examine only one specific event at a time; multiple fault trees may be developed. • The levels and the organisation of the tree vary from analyst to analyst. • Quantification requires a high level of expertise [18].
12	**Failure Mode and Effects Analysis (FMEA)** [46]	An analysis technique used in high-risk organizations to identify failure modes in systems/processes and work out response strategies.	• Effective for collecting the information that is needed. • Widely used/ understood, provides a great understanding	• Examination of human errors is limited. It is focused on technical failures and operational errors may be overlooked. • Complex

No.	Technique	Description	Strengths	Weaknesses
			of a system. • Systematic and comprehensive [18].	interactions resulting from more than one failure are often omitted [18]. • Not appropriate for selecting single ideas. • Very complex.
13	**Failure Mode and Effects Criticality Analysis (FMECA)** [46]	An analysis technique used in high-risk organizations to identify and assess failure modes in systems/processes and work out response strategies.	Like FMEA	Like FMEA
14	**Fuzzy Logic** [47]	Useful approach to address the problems associated with imprecision, uncertainty, and subjectivity of data.	• Permits different kinds of data to be manipulated simultaneously using a standardised methodology and a common scale for expressing the significance of impacts.	• Offers no significant benefits in the case of simple projects. • Characterized by mathematical complexity.
15	**Hazard and Operability (HAZOP)** [48]	It is a hazard identification technique that uses a structured and systematic team review of a system or process to identify the possible deviations from normal operations and their causes and consequences. It uses a standard list of guidewords (e.g. "more," "less," "no") combined with process conditions to systematically consider all the possible deviations from the normal conditions. For each deviation, possible causes and consequences are identified as well as whether additional safeguards should be recommended.	• Uses the experience of operating personnel. • Systematic and comprehensive. • Effective for technical faults and human errors. • Employs a team approach requiring the interaction of several disciplines or organisations [18].	• Depends very much on expert judgements. • Optimised especially for sequential operations or procedures. • Requires the development of procedural descriptions that are often not available in detail. • Documentation is lengthy. • One of the most time consuming and expensive techniques [18].

No.	Technique	Description	Strengths	Weaknesses
16	**Hazard Review (HR)** [18]	The Hazard Review, also known as Hazard Survey or Safety Review, is mainly a qualitative review of an activity or system to identify the hazards and to gain qualitative understanding of their significance.	• Makes use of the existing experience taken from a wide range of sources. • Can be performed by a single analyst at a low cost [18].	• A lack of structure makes it difficult to audit. • Limited to previous experience and thus with a limited value for novel installations. • Does not produce a list of failure cases for a quantitative risk assessment [18].
17	**Human Reliability Assessment (HRA)** [49]	It is especially used for a detailed evaluation of human operations in procedural tasks. It is a special form of FTA and ETA, designed for modelling and analysing the range of possible accidents that may happen while performing a procedure.	• Provides useful information about the cost and value of human resources. • Helps an organisation to make the best utilisation of human resources.	• Focused on specific human reliability issues. • The evaluation of human assets is based on the assumption that the employees are going to remain with the organisation for a specified period. However, this assumption is wrong because employee mobility is very high.
18	**Incident Reporting (IR)** [50]	A structured mode for accident, incident, and near miss signalling collection.	• IR forms identify the barriers that prevent adverse situations. • IR schemes provide a means of encouraging staff participation in safety improvement.	• It can be difficult both to set up and to maintain.
19	**Interviews** [1]	The list of risks is produced by interviewing project managers or experts on the applications of the project. The risks are identified and	• Simple to use at the basic level. • Systematically assesses the experiences	• Limited to previous experience only. • Gives few insights into the nature of the hazards, may miss

No.	Technique	Description	Strengths	Weaknesses
		defined and a risk management capability score can be determined from a five-point scale.	accumulated by an industry. • Can be prepared by either a single analyst or a small group.	some potential problems. • Individual risk drivers may be described in insufficient detail to avoid ambiguity. • Can be limiting.
20	**Monte Carlo** [1]	A type of spreadsheet simulation that randomly and continuously generates values for uncertain variables to simulate a model.	• Allows to work in terms of real units. • Allows models to be firmly rooted in the plans of a project. • Makes the relationship between the output of models and real-world decisions relatively straightforward.	• No statistically sound basis to specify distributions. • No basis for estimating the most likely values. • No basis to create custom tailored distributions when real world data are missing.
21	**Pareto Analysis (PA)** **or** **ABC analysis** [51]	It is a technique that is used to identify and prioritise the most significant items, for example causes and contributing factors or effects of accidents. This technique employs the Pareto rule (or 80-20 rule),which says that about 80 percent of the effects are generated by about 20 percent of the causes.	• Provides quantitative results [18]. • Many application possibilities in different areas, from the activity or operations level to the system level, such as ranking activities or system accidents and their causes. • Can also be used to evaluate changes in risks after modifications in a system or activity. • Simple to use. • Individual or group technique.	• Focuses only on the past. • Produces considerable variability in the levels of risk assessment resolution. • Dependent on availability and applicability of data [18]. • Must be careful when setting importance criteria.

No.	Technique	Description	Strengths	Weaknesses
22	**Preliminary Hazard Analysis (PHA)** [52]	It is used to identify hazards, assess the severity of potential accidents that may happen, and identify measures for reducing or eliminating the risks associated with the hazards.	• Used as a proactive technique because it identifies the weaknesses of a system at the early stages of its life, thus saving time and money [18]. • May be applied to any kind of risk analysis and to any activity or system.	• Requires additional analysis to understand more in depth and evaluate hazards and potential accidents. • Relies heavily on the knowledge of subject matter experts [18].
23	**Risk Breakdown Matrix (RBM)** [23]	An activity and threat matrix where the value of risk associated with each activity and the most frequent overall risks are evaluated.	• Many application possibilities in different areas. • Individual or group technique. • Very detailed.	• Must be careful when setting scoring criteria. • Enables a more detailed analysis of vital factors. • Very complex, requires training.
24	**Risk Breakdown Structure (RBS)** [53]	It is a source-oriented grouping of project risks that defines the total risk exposure of a project. Each descending level represents an increasingly detailed definition of sources of risk to the project.	• Help the project/risk manager to better understand recurring risks and concentrations of risks which would lead to issues that affect the status of the project.	• The level of detail depends on the available information.
25	**Risk Mapping, Risk Matrix, Probability and Impact Matrix** [1,13]	It is a qualitative technique that can be used to evaluate and prioritise a group of risks which could significantly impact on a project.	• Allows to brainstorm the most likely project risks and to apply simple formulas to them. • Communicative. • Aids the creation of a shared understanding of the importance of various risks to the project. • Simple.	• Shortcomings result from a checklist approach (see Checklist). • Ratings have no absolute meaning. • Danger of prematurely defining high and low risks with no further considerations.

No.	Technique	Description	Strengths	Weaknesses
26	**Risk Probability and Impact Assessment, Risk Ranking/** **Risk Index** [1]	It investigates the likelihood that each specific risk will occur and the potential effects on the objectives of a project, such as time, cost, scope, or quality.	• Provides a high-level assessment [18]. • Identifies both negative effects for threats and positive effects for opportunities.	• Results can be difficult to link to absolute risks. • Appropriate ranking tools may not exist. • Does not account for unique situations [18].
27	**Sensitivity analysis** [1,13]	It helps to determine which risks have the most potential impact on a project.	• Useful for comparing the relative importance of variables that have a high degree of uncertainty to those that are more stable.	• Requires a great availability of historical data.
28	**Strengths, Weaknesses, Opportunities, and Threats (SWOT)** [54]	The SWOT analysis provides a good framework for reviewing strategies, positions and business directions of a company or an idea.	• Individual or group technique. • Very broad areas of application. • Easy to use.	• Not very applicable to general idea selection. • Mainly used in the business field.
29	**SWIFT Analysis** [18]	It is a more structured form of the "What-if Analysis" technique and it is used to identify hazards based on brainstorming and checklists.	• Possible problems and combinations of conditions that can be problematic are described. • Possible risk-reducing measures are identified.	• Requires a great variety of competencies of the analysis team.
30	**What-if Analysis** [18]	It is a brainstorming technique that uses a systematic, but broad and not very structured, questioning procedures to generate descriptive information.	• Highly effective to identify system hazards. • A simplistic approach that offers great value for minimal investment [18].	• Loose structure and reliance on judgements, likely to miss some potential problems. • Difficult to audit for thoroughness.

No.	Technique	Description	Strengths	Weaknesses
				• The danger in this technique lies in the unasked questions [18].
31	**"5 Whys" Technique** [18]	It is a qualitative brainstorming technique that attempts to identify root causes of accidents by asking "why" these events did occur or conditions did exist, in order to help to get to the true causes of problems.	• Used as an effective technique for identifying root causes of accidents and determining causal factors.	• Mainly based on brainstorming that is often time consuming. • The brainstorming process is very difficult to duplicate and the results may not be reproducible or consistent. • It does not ensure that all the root causes can be identified.

Table 1. Project risk management techniques

The selected project risk management techniques are now classified according to the three proposed dimensions (Table 2). It is worth remarking that the techniques have been matched with the dimensions based on their most frequent applications as documented by literature and on the authors' experience. Different categorisations may be possible according to the peculiar characteristics of specific project settings.

During the entire project life cycle and in every stage of the risk management process, the nature and the quantity of available information influence the choice of the techniques that should be applied. In the conceptualisation phase decision-makers have a high degree of freedom in defining project goals and how to achieve them. However, owing to the lack of project specifications on the ways to meet the set objectives in that stage of the project, all the necessary information for a complete investigation of risk is not always available. Then, we are in an uncertain scenario characterised by a limited amount of information or in a context where the source of information is subjective. Therefore, it is necessary to build a systematic framework that can be used by decision-makers to obtain subjective judgements from experts in a clear and straightforward manner. This can be accomplished by applying "extractors" of information like Interviews or the so called "group techniques" such as Brainstorming, Delphi, and Expert Judgment. At the same time, it is also necessary to train the experts so that they can make good judgements. Moreover, this context may just allow to define the strengths and weaknesses of the project and the decision-makers may stop their risk investigation at the identification phase by means of a SWOT analysis. However, if we are in the case of repetitive projects, the greater availability of information could allow the use of detailed tables, such as FMEA [25], and makes possible to define occurrence probabilities and economic and/or time impacts for every alternative event. In this situation, decision-makers could move on to a quantitative analysis of risks through the use of FMECA tables, Decision Trees, and Event Tree Analysis. As a consequence, the quantity and kind of information in the conceptualisation phase usually allow risk identification and they seldom enable also risk analysis. Coming to the planning phase, the ways and means to achieve the project objectives become clearer thanks to a considerable increase in the available information, which allows a complete investigation of risks. All the techniques for risk management can be used in this project stage based on the phases of identification, analysis, and response to risk and on the type of information available. In general, the degree of knowledge and the ability to influence the course of a project are inversely proportional to each other as the project develops overtime. Therefore, in the execution phase there will be a high level of knowledge about project constraints but a low ability to influence events because all the most important project and risk management choices have been already made in the previous phases. The result is that in this phase the time and economic performance resulting from the project choices and the actions undertaken to either mitigate or exploit risk can be mainly controlled and monitored. Therefore, in the execution phase the outputs obtained from the techniques applied in risk identification, analysis, or response will be revised and the results of the implementation of designed actions will be monitored by means of careful and sensible human action. In addition, in this project stage the risk management techniques used in the planning phase can be applied again to unveil new risks that have not emerged before. The termination phase is not considered by the classification in Table 2 because the risk management effort is more relevant in the previous stages of the project life cycle. Also, the risk management planning phase is not included being less operational in nature than the subsequent phases and more focused on the strategy to deal with risk and the project goals.

Finally, the level of maturity is very linked with the level of communication in the organisation and the availability of data/information about the project. The higher the maturity towards risk management of the project team the more common the use of various techniques, especially the quantitative ones, during the entire risk management process. For example, the Monte Carlo simulation technique, that can be applied in the phase of quantitative risk analysis, is basically used by companies with a high level of maturity towards data and information management and hence project risk. The last column of Table 2 refers to the maturity levels proposed by Hillson [37]: the Naïve stage is not taken into account because it is not characterised by the use of any risk management technique. Also, the following notation has been used in Table 2: I = “risk Identification”, QlA = “Qualitative risk Analysis”, QtA = “Quantitative risk Analysis”, and R = “risk Response”.

		Dimensions		
No.	**Technique**	**Risk Management Phase**	**Project Life Cycle Phase**	**Level of Corporate Maturity**
1	**Brainstorming**	**I** [1,32,55-58], **QlA** [29]	Conceptualisation [25], Planning, Execution	Novice [25], Normalised, Natural
2	**Cause and –effect diagram or Cause Consequence Analysis (CCA)**	**I** [1,22], **QlA**[11]	Planning, Execution	Normalised, Natural
3	**Change Analysis (ChA)**	**I** [59], **QlA**, **R** [18]	Planning, Execution	Normalised [18], Natural
4	**Checklist**	**I** [1,15,32,56], **QlA** [11]	Conceptualisation, Planning [25]	Novice [18], Normalised, Natural
5	**Decision Tree Analysis**	**QtA**[1,11,26,32], **R** [55]	Conceptualisation, Planning	Normalised, Natural
6	**Delphi**	**I** [1,22], **QlA** [29,60]	Conceptualisation [25], Planning	Novice [25], Normalised, Natural
7	**Event and Causal Factor Charting (ECFCh)**	**I** [18]	Planning	Normalised [18], Natural
8	**Event Tree Analysis (ETA)**	**I** [61], **QlA** [11], **QtA** [11,18,61]	Conceptualisation, Planning	Normalised, Natural [18]
9	**Expected Monetary Value**	**QtA**[1,11,32], **R** [55]	Planning, Execution	Natural
10	**Expert Judgement**	**I, QlA, QtA**[1], **R** [55]	Conceptualisation, Planning	Normalised, Natural
11	**Fault Tree Analysis (FTA)**	**I** [22,45], **QlA** [11], **QtA** [18,11]	Conceptualisation [25], Planning	Normalised, Natural [18]
12	**Failure Mode and Effects Analysis (FMEA)**	**I, R**[46,62]	Conceptualisation [25], Planning	Normalised [18]
13	**Failure Mode and Effects**	**I, QlA, QtA, R**	Conceptualisation [25],	Normalised [18],

		Dimensions		
No.	**Technique**	**Risk Management Phase**	**Project Life Cycle Phase**	**Level of Corporate Maturity**
	Criticality Analysis (FMECA)	[30,46,50,62]	Planning, Execution	Natural
14	**Fuzzy Logic**	**QtA** [11,47,63]	Planning	Natural
15	**Hazard and Operability (HAZOP)**	**I** [29,32,48], **R** [18]	Planning	Normalised [18], Natural
16	**Hazard Review (HR)**	**I** [18]	Planning	Novice, Normalised [18], Natural
17	**Human Reliability Assessment (HRA)**	**I, QlA,QtA, R**[18,49]	Planning, Execution	Normalised, Natural
18	**Incident Reporting**	**I**[50], **QtA**	Planning	Normalised, Natural
19	**Interviews**	**I** [1,22,58], **QlA** **[**57**],QtA**[1]**, R**[15]	Conceptualisation, Planning, Execution	Novice, Normalised, Natural
20	**Monte Carlo**	**QtA** [1,11,26,29,32]	Planning	Natural
21	**Pareto Analysis (PA)** **or ABC analysis**	**QtA** [18,51]	Planning	Natural
22	**Preliminary Hazard Analysis (PHA)**	**I** [52], **QlA**[52], **P**[52]	Planning	Novice, Normalised [18], Natural
23	**Risk Breakdown Matrix (RBM)**	**I,QlA,QtA** [23]	Planning	Normalised, Natural
24	**Risk Breakdown Structure (RBS)**	**I** [35]	Conceptualisation, Planning	Normalised, Natural
25	**Risk Mapping, Risk Matrix, Probability and Impact Matrix**	**I, QlA** [1,11,64,65]	Planning	Normalised, Natural

		Dimensions		
No.	**Technique**	**Risk Management Phase**	**Project Life Cycle Phase**	**Level of Corporate Maturity**
26	**Risk Probability and Impact Assessment, Risk Ranking/ Risk Index**	**QlA** [1], **QtA**	Planning	Normalised [18], Natural
27	**Sensitivity analysis**	**QtA** [1,11,26,32], **R**	Planning, Execution	Natural
28	**Strengths, Weaknesses, Opportunities, and Threats (SWOT)**	**I** [54,58], **QlA** [29], **R**	Conceptualisation, Planning	Normalised, Natural
29	**SWIFT Analysis**	**I, R**[18]	Planning	Normalised, Natural
30	**What-if Analysis**	**I, R**[18]	Conceptualisation, Planning	Normalised [18], Natural
31	**"5 Whys" Technique**	**I** [18]	Planning	Natural

Table 2. Classification of project risk management techniques

Table 2 allows to characterise each technique based on the risk management phases, the project life cycle phases, and the degree of corporate maturity towards risk for which it is most suitable. However, it does not provide a global view of how all the analysed techniques fit into the dimensions. In order to overcome this limitation, two bi-dimensional charts are built. On the one hand, Figure 1 places the techniques on a Cartesian plane according to the phases of the risk management process (x-axis) and phases of the project life cycle (y-axis) for which they can be used. On the other hand, Figure 2 compares the same techniques but against the risk management phases (x-axis) and the corporate maturity towards risk (y-axis).

These charts are intended to be a valuable mean to communicate and to stimulate knowledge creation about risk. They may be used by an organisation to select a set of techniques, discuss when they are appropriate, and decide which of them could be used, how, and in which part of the project and risk management processes. Also, such representations allow to make further considerations about the appropriateness of each technique. Figure 1 highlights that in the Planning phase of a project there are a lot of techniques that can be used. In fact, in this stage more time can be spent on strategic issues such as risk managing than in the Conceptualisation stage, which has usually a quite limited duration, and in the Execution stage, which is mainly focused on the achievement of the project objectives from an operational point of view. Figure 2 graphically proves the relationship between the maturity towards risk and the phases of the risk management process that are carried out by a company. By considering the maturity model proposed by Hillson [37], a Novice level of maturity usually implies performing just risk identification. A Normalised maturity also involves a qualitative risk analysis and, in some limited cases, also risk response and monitoring and control. Finally, a Natural maturity is associated with undertaking the complete risk management process, from identification to monitoring and control, including the quantitative risk analysis. Therefore, the quantitative analysis of risk distinguishes companies with a Natural maturity level from companies having a Normalised maturity level. Additionally, in the Natural maturity level there is a complete integration between the project management and the risk management processes that allows a regular revision of the outputs of the applied risk techniques.

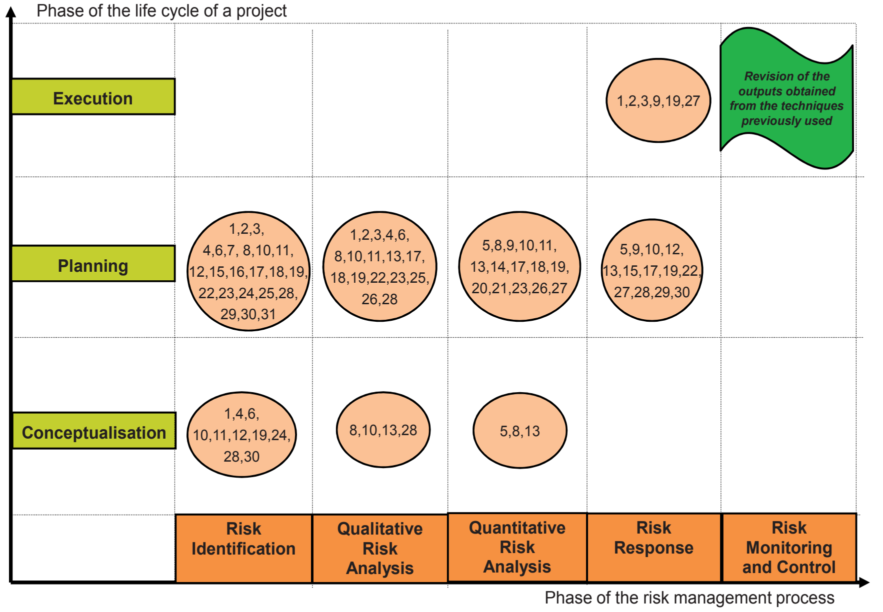


Figure 1. Risk technique mapping: risk management and project life-cycle phases

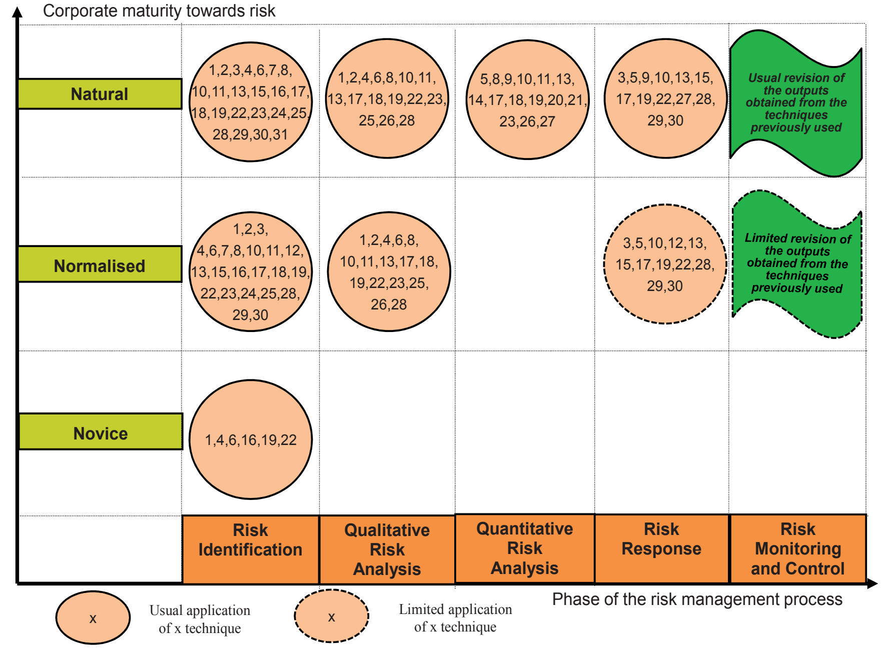


Figure 2. Risk technique mapping: risk management phases and corporate maturity levels

5. Discussion

Communication, information, and hence knowledge are the cardinal points for an attitude towards project risk management that goes beyond an informal approach limited to qualitative investigation. A systematic acquisition and organisation of information is a necessary step in order to move from a subjective knowledge about risk, that has to be elicited from experts, to an objective and easily accessible knowledge forming the condition for a quantitative risk analysis. The framework proposed in this chapter aims to help such transition by generating knowledge about the potentiality of application of common risk techniques.

Some advantages can be identified. First of all, the developed taxonomy helps to understand how the project environment relates to risk techniques. Also, the suggested scheme provides guidelines about the most relevant dimensions that should be taken into account simultaneously in a risk management process, thus making it more comprehensive, even if it can never be complete because of the limited amount of available resources and the bounded rationality of human beings [66]. This generates knowledge based on the degree of maturity towards risk of the organisation running the project and such knowledge in turn increases the level of corporate awareness towards the instruments to tackle risk. Furthermore, the proposed framework benefits from being quite general, so that it can be easily adapted to reflect the requirements of different industries. Finally, it is suitable to both small-scale and large-scale projects.

Tangible and intangible benefits can be derived from the application of the framework. Tangible advantages are associated with decision-making and include an improved understanding of projects, giving as a consequence a better control over resources, the provision of a structured support to develop and implement monitoring strategies, and a better use of means to identify and assess risk with an inherent positive impact on the evaluation of contingencies. Among intangible benefits, facilitation of a rational risk taking and improvement of communication can be mentioned. The developed framework also encourages a more proactive approach to risk as a result of a well planned management process. All these characteristics ultimately emphasise the integration among project and risk management.

However, the criteria and the classification of the techniques to support risk management have been derived exclusively from the available literature and from the authors' experience. Empirically testing the outcomes of this study by applying them to real projects would be of great value to validate and refine the framework.

Therefore, future research efforts will be directed towards the implementation of the framework in multiple project settings in representative industries. Enhancing the taxonomy by introducing further dimensions, such as the complexity level of a project and the degree of innovation of its product, will be considered. The degree of innovation of the product of a project is particularly interesting because it may be connected with the phases of the project life cycle. In fact, the more innovative is the outcome, the more the risk management process will be concentrated in the planning phase. Conversely, the less innovative the product the more the focus on risk in the execution phase. Additional evolutions will be concerned with a systematic analysis of the concepts of method, technique, and tool together with the study of the relationships among them, and with extending the framework to include new practices to support risk management. Finally, a further research line could deal with the integration of the proposed framework into a global project management process with the aim of overcoming the traditional separation between running a project and identifying, assessing, and controlling the associated risks.

6. Conclusion

The extreme importance of information and associated knowledge to ensure an effective management of risk demands paying greater attention both to the understanding of the effects of randomness in projects and to the learning of available means to capture this variability. The present work focuses on the second issue and introduces a framework to classify techniques supporting project risk management based on their purpose and the context for which they are most suitable. The main aim is incrementing communication and knowledge enabling a quantification of risk. The scheme is general and can be applied to very diverse projects in numerous industries.

Author details

Sabrina Grimaldi, Carlo Rafele and Anna Corinna Cagliano
Department of Management and Production Engineering, Politecnico di Torino, Torino, Italy

7. References

[1] Project Management Institute. A guide to the Project Management Body of Knowledge (PMBOK® Guide). 4th ed. Newtown Square, Pennsylvania: PMI Publications; 2008.
[2] Kotnour T. Organizational learning practices in the project management environment. Intern J Qual & Reliab Manag. 2000; 17 (4/5): 393-406.
[3] Van Wyk R, Bowen P, Akintoye A. Project risk management practice: The case of a South African utility company. Intern J Proj Manag. 2008; 26(2): 149-63.
[4] Carbone TA, Tippett DD. Project risk management using the project risk FMEA. Eng Manag J. 2004; 16(4): 28-35.
[5] Kutsch E, Hall M. Intervening conditions on the management of project risk: dealing with uncertainty in information technology projects. Intern J Proj Manag. 2005; 23 (8): 591-59.
[6] Guofeng W, Min W, Weiwei Z. Study on the existing problems and countermeasures of project risk management in China. Energy Procedia. 2011; 13: 2726-33.
[7] Helton JC, Johnson JD, Oberkampf WL. An exploration of alternative approaches to the representation of uncertainty in model predictions. Reliab Eng and Syst Saf. 2004; 85 (1-3): 39-71.
[8] Oberkampf WL, Helton JC, Joslyn CA, Wojtkiewicz SF, Ferson S. Challenge problems: uncertainty in system response given uncertain parameters. Reliab Eng and Syst Saf. 2004; 85 (1-3): 11-9.
[9] O'Hagan T. Dicing with the unknown. Signif. 2004; 1(3): pp. 132-3.
[10] Hillson D, Simon P. Practical project risk management. The ATOM methodology. Vienna, Virginia: Management Concepts; 2007.
[11] Del Cano A, De La Cruz M.P. Integrated methodology for project risk management. J Constr Eng and Manag. 2002; 128 (6): 473-85.
[12] Pipattanapiwong J, Watanabe T. Risk management for an international project: a case study of infrastructure construction project financed by the Asian Development Bank (ADB) in Thailand. In: Proceedings of the 18th Annual Forum on Construction Management Related Issues; 2000 December; Tokyo, Japan. p. 207-16.
[13] Chapman C, Ward S. Project risk management. Processes, techniques and insights. 2nd ed. Chichester, England: John Wiley & Sons; 2003.
[14] Seyedhoseini SM, Hatefi MA. Two-Pillar Risk Management (TPRM): a generic project risk management process. Trans E: Ind Eng. 2009; 16(2): 138-48.
[15] Association for Project Management. Project risk analysis and management guide. 2nd ed. High Wycombe, England: APM Publishing; 2004.

[16] The Institution of Civil Engineers & The Faculty and Institute of Actuaries. RAMP. Risk analysis and management for projects. 2nd ed. London: Thomas Telford Publishing; 2005.
[17] Dey PK, Kinch J, Ogunlana SO. Managing risk in software development projects: a case study. Ind Manag & Data Syst. 2007; 107(2): 284-303.
[18] Mullai A. Risk management system–risk assessment frameworks and techniques. Turku, Finland: DAGOB Publication series 5:2006; 2006.
[19] Walewski J, Gibson GE. International project risk assessment: methods, procedures, and critical factors. Report of the Center Construction Industry Studies, The University of Texas at Austin. 2003 September; [cited 2012 April 2]. Available from: http://www.ce.utexas.edu/org/ccis/a_ccis_report_31.pdf.
[20] Artto KA, Kähkönen K, Pitkänen PJ, editors. Unknown soldier revisited: a story of risk management. Helsinki: Project Management Association Finland; 2000.
[21] Ahmed A, Kayis B, Amornsawadwatana S. A review of technicians for risk management in projects. Benchmarking Intern J. 2007; 14(1): 22-36.
[22] Dey PK, Ogunlana SO. Selection and application of risk management tools and techniques for build-operate-transfer projects. Ind Manag & Data Syst. 2004; 104 (4): 334-346.
[23] Hillson DA, Grimaldi S, Rafele C. Managing project risks using a cross risk breakdown matrix. Risk Manag. 2006; 8: 61–76.
[24] Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. J Am Stat Assoc. 2005; 100(470):680-701.
[25] Grubisic V, Ferreira V, Ogliari A, Gidel T. Recommendations for risk identification method selection according to product design and project management maturity, product innovation degree and project team. In: Culley SJ, Hicks BJ, McAloone TC, Howard, TJ, Cantamessa M, editors. Proceedings of the 18th International Conference on Engineering Design (ICED 11); 2011 August 15-18; Copenhagen.
[26] Hillson D. Extending the risk process to manage opportunities. Intern J Proj Manag. 2002; 20(3): 235-40.
[27] Kmec P. Temporal hierarchy in enterprise risk identification. Manag Decis. 2011; 49(9): 1489-509.
[28] Klakegg OJ, Marianeschi MM, Risan, Ø. Building trust – the power of analysing risk using resource groups in managing business by projects. In: Artto KA, Kähkönen K, Koskinen K, editors. Managing business by projects. Helsinki: Project Management Association Finland and NORDNET; 1999. p. 909-20.
[29] Berg H-P. Risk management: procedures, methods and experiences. RT&A [serial on the Internet]. 2010 June; [cited 2012 March 30]; 79-95. Available from: http://gnedenko-forum.org/Journal/2010/022010/RTA_2_2010-09.pdf.
[30] Rot A. IT risk assessment: quantitative and qualitative approach. In: Proceedings of the World Congress on Engineering and Computer Science (WCECS 2008); 2008 October 22 - 24; San Francisco, USA.
[31] Gidel T, Zonghero W. Management de projet 1 - Introduction et fondamentaux. Paris: Lavoisier: 2006.

[32] Lyons T, Skitmore M. Project risk management in the Queensland engineering construction industry: a survey. Intern J Proj Manag. 2004; 22(1):51-61.

[33] Hopkinson M. The project risk maturity model. Measuring and improving risk management capability. Surrey, England: Gower Publishing Limited; 2011.

[34] Hassan A. Risk management practices of Islamic banks of Brunei Darussalam. J Risk Finance. 2009; 10(1): 23-37.

[35] Hillson D. Effective opportunity management for projects. Exploiting positive risk. New York: Marcel Dekker; 2004.

[36] De Bakker K, Stewart WM, Sheremeta PW. Risk management planning-how much is good enough? In: Proceedings of the Fifth European Project Management Conference, PMI Europe 2002; 2002 June 19-20; Cannes, France.

[37] Hillson DA. Towards a risk maturity model. Intern J Proj & Bus Risk Manag. 1997; 1(1): 35-45.

[38] Hulett DT. Key characteristics of a mature risk management process. In: Proceedings of the Fourth European Project Management Conference, PMI Europe 2001; 2001 June 6-7; London.

[39] Grimaldi S, Rafele C. A tool selection and support methodology for project risk management. In: Proceedings of the 22nd IPMA World Congress "Project Management to Run"; 2008 November 9-11; Rome.

[40] Hillson D. Risk management isn't hard. Proj Manag Today. 2009 May; 25.

[41] Labuschagne C, Brent AC. Sustainable project life cycle management: the need to integrate life cycles in the manufacturing sector. Intern J Proj Manag. 2005; 23(2): 159-68.

[42] Cohen MW, Palmer, GR. Project risk identification and management. AACE Intern Trans. 2004.

[43] Kwak YH, Ibbs CV. Project Management Process Maturity $(PM)^2$ model. J Manage Eng. 2002; 18(3): 150-55.

[44] Subramaniam N, Collier P, Phang M, Burke G. The effects of perceived business uncertainty, external consultants and risk management on organizational outcomes. J Account & Organ Chang. 2011; 7(2): 132-57.

[45] Chapman RJ. Simple tools and techniques for enterprise risk management. Chichester, England: Wiley Finance; 2006.

[46] Bouti A, Kadi DA. A state of the art review: of FMEA/FMECA. Intern J Reliab Qual Saf Eng. 1994; 1(4): 515-43.

[47] Bellagamba L. Estimating risk adjusted cost or schedule using fuzzy logic. In: Proceedings of the 9th Annual Symposium of the International Council on Systems Engineering; 1999 June 9; Brighton, England.

[48] Kletz T. HAZOP and HAZAN: identifying and assessing process industry hazards. London: Taylor & Francis; 1999.

[49] Lyons M, Adams S, Woloshynowych M, Vincent C. Error reduction in medicine. Final report to the Nuffield Trust. Department of Surgical Oncology & Technology, Imperial College, London. 2005; [cited 2012 March 1]. Available from: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/Error_reduction_in_med%20Dec%202004.pdf.

[50] Cinotti R. La gestione del rischio nelle organizzazioni sanitarie. Rome: Il Pensiero Scientifico Editore; 2004.
[51] Rebernik M, Bradač B. Idea evaluation methods and techniques. Institute for Entrepreneurship and Small Business Management, University of Maribor, Slovenia. 2008; [cited 2012 February 27]. Available from: http://www.creative-trainer.eu/fileadmin/template/download/module_idea_evaluation_final.pdf.
[52] Tarakji G. Tools of Risk Management. Engineering 801: Engineering Management. 2003 April 23; [cited 2012 March 14]. Available from: http://userwww.sfsu.edu/~engr801/projects/risk.pdf.
[53] Hillson DA. The Risk Breakdown Structure (RBS) as an aid to effective risk management. In: Proceedings of the Fifth European Project Management Conference, PMI Europe 2002; 2002 June 19-20; Cannes, France.
[54] Emblemsvåg J, Kjølstad LE. Strategic risk analysis-a field version. Manag Decis. 2002; 40(9): 842 - 52.
[55] Dey PK. Decision support system for risk management: a case study. Manag Decis. 2001; 39(8): 634 - 49.
[56] Raz T, Michael E. Use and benefits of tools for project risk management. Intern J Proj Manag. 2001; 19(1): 9-17.
[57] Institute of Management Accountants. Enterprise risk management: tools and techniques for effective implementation. 2007; [cited 2012 March 20]. Available from: http://mgt.ncsu.edu/erm/documents/IMAToolsTechniquesMay07.pdf.
[58] Gupta PK. Risk management in Indian companies: EWRM concerns and issues. J Risk Finance. 2011; 12(2): 121-39.
[59] Spath P. Analyse process changes to identify system failures. Hosp Peer Rev. 2000 May; 25(5):65-7.
[60] Bandyopadhyay K, Mykytyn PP, Mykytyn K. A framework for integrated risk management in information technology. Manag Decis. 1999; 37(5): 437- 45.
[61] Rao Tummala VM, Leung YH. A risk management model to assess safety and reliability risks. Intern J Qual & Reliab Manag. 1996; 13(8): 53-62.
[62] Sinha PR, Whitman LE, Malzahn D. Methodology to mitigate supplier risk in an aerospace supply chain. Supply Chain Manag: Intern J. 2004; 9(2): 154 - 68.
[63] Konstandinidou M, Nivolianitou Z, Kiranoudis C, Markatos N. A fuzzy modeling application of CREAM methodology for human reliability analysis. Reliab Eng and Syst. 2006; 91(6): 706–16.
[64] Scandizzo S. Risk mapping and key performance indicators in operational risk management. Economic Notes by Banca Monte dei Paschi di Siena. 2005; 34(2): 231-56.
[65] Sheffi Y, Rice JB. A supply chain view of the resilient enterprise. MIT Sloan Manag Rev. 2005; 47(1): 41-52.
[66] Green SD. Towards an integrated script for risk and value management. Intern J Proj Manag. 2001; 7(1): 52-8.

Chapter 5

Long Memory in the Volatility of Local Currency Bond Markets: Evidence from Hong Kong, Mexico and South Africa

Pako Thupayagale

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50888

1. Introduction

Investors can potentially improve the risk-adjusted performance of their portfolios by investing internationally and thereby take advantage of the associated return and diversification benefits. The potential gains provided by emerging markets have attracted significant investor attention which, in turn, has led to substantial capital inflows to these economies. Local currency-denominated sovereign bonds have been the fastest growing market in emerging market space in the past few years. More recently, the global quest for yield in a context of accommodative monetary policies in the advanced economies has created a positive external environment for emerging market bonds. In addition, the secondary market for emerging market local currency debt has also been supported by high interest and amortisation payments. Emerging market bonds are also benefiting from a track record of strong risk-adjusted returns and low correlations with other asset classes. Such characteristics are attractive from a portfolio optimisation perspective. These attributes have also seen considerable attention devoted to analysis of the various risk-return attributes of these markets. In particular, recent empirical literature has focused on the characterisation of the volatility profile of emerging market bond returns. Indeed, the accurate estimation of volatility plays a central role in many applications in finance, including optimal portfolio selection (e.g., diversification strategies), valuation of derivatives (e.g., option pricing) and risk management (e.g., value-at-risk calculation). These applications have motivated an extensive empirical literature on volatility modelling.

While recent empirical literature has focused on the characterisation of the volatility profile of a variety of asset classes; in particular, the long memory properties of these assets, surprisingly little attention has been devoted to the analysis of fixed income markets, especially in emerging markets. The empirical literature on long memory dynamics of fixed income volatility and its implications for portfolio and risk management appears to be limited. Most of the extant literature appears concentrated in the advanced economies. For example, several authors have examined the various aspects of long memory behaviour in interest rates and yield spreads. [1-3] The purpose of this chapter is to augment this line of analysis concerning the long memory attributes of fixed income volatility in emerging market local currency debt market in light of investor interest in the potential alpha generation of these markets, amid wider capital inflows into emerging market bonds as investors search for yield.

This study will focus on government bond markets from Hong Kong, Mexico and South Africa. According to the most recent survey from the Emerging Markets Trading Association (EMTA) these three local currency debt markets are among the most vibrant and actively traded in emerging markets. [4] As a result of their (comparatively high) liquidity, developed institutional frameworks and credible monetary policies, these markets are therefore of interest to investors. Indeed, local currency bond markets have emerged as an important asset class in many emerging markets; a point which becomes salient in the current low-yield environment, where investors targeting high returns and diversification benefits have channelled capital to emerging bond markets such as these.

As a result of regulatory initiatives and various reforms, emerging market local currency sovereign debt markets have has grown rapidly in size and sophistication. According to [4], emerging market debt trading volumes were at USD1.8 trillion in the third quarter of 2011. This represents a 3 percent increase from the USD1.7 trillion reported for the second quarter of 2011. Turnover in local market instruments was at USD1.3 trillion in the third quarter of 2011, (i.e., 76 percent of total reported volume). Mexican and Hong Kong securities were the first and second most frequently traded emerging market debt in the third quarter of 2011 at USD282 billion and USD176 billion, respectively compared to USD136 billion and USD201 billion a year earlier. The next most frequently traded local markets debt were those from Brazil (USD160 billion), South African (USD113 billion) and Turkey (USD69 billion).[1]

Domestic institutional investors are typically the largest investors in local currency bond markets. For example, pension funds tend to have long-term liabilities which are typically funded by investments in long-term investment grade securities that provide a prudent risk-return profile. As a result, the examination of long memory in volatility would appear to be of interest to institutional and other long-term investors. Furthermore, it has been shown that it is important to model the long memory volatility structure when pricing derivative contracts with long maturity. [5] In addition, in order to assess future returns from both active and passive investment strategies or the need for policy intervention, especially over long horizons, it is important to forecast volatility. These applications have motivated an extensive empirical literature on modelling long memory dynamics in asset return data. Analysis of the long-term volatility dynamics of emerging market local currency bonds appear limited. Therefore, this

[1] Analysis is on Hong Kong, Mexico and South Africa due to the availability of data.

study attempts to help fill this gap by addressing a range of issues relating to the estimation and forecasting of fixed income return volatility especially over long horizons.

Against this background, this paper has three objectives. First, evidence of long memory in volatility within leading emerging fixed income markets is investigated. In particular, the existence of long memory behaviour in the volatility of returns from Hong Kong, Mexico and South Africa are examined which appears to have little or no previous research establishing the existence of its long memory properties. In order to estimate the long memory parameter *d*, this study makes use of methods based on wavelets, which have been recently used to capture the fractal structure of high frequency data.[6] Second, the existence of long memory dynamics in bond volatility data (i.e., a fractal structure in the data) will be further investigated in order to test if the extraction of this long-run component can be exploited for purposes of generating improved volatility forecasts especially over long horizons. The long memory property is examined in order to determine if it helps deliver more accurate forecasts over a long(er) horizon. Third, an important and topical area of research concerns the calculation of value-at-risk (VaR) in financial markets. This methodology is widely used by financial institutions and regulatory agencies to measure, monitor and manage market risk. This analysis compares whether long memory volatility estimates can help deliver more accurate VaR estimates relative to standard models (i.e., GARCH and RiskMetrics).

In total, the findings of this investigation will provide a range of volatility estimates and forecasts which could potentially inform portfolio management strategies and guide policymaking. In particular, while most empirical studies focus on the United States and other developed markets, recent research has begun to look at emerging markets, however, limited evidence exists with respect to these markets. This analysis contributes to the empirical literature by focusing on various aspects of long memory behaviour in local currency debt markets. The findings from this research complement those in previous studies and may provide an interesting comparison to existing studies.

The rest of the chapter is structured as follows. Section II presents a description of long memory in time series. Section III introduces the data. Section IV presents the empirical methodology and associated results. In particular, this starts with a presentation of the standard GARCH model which is often used to present initial evidence of long memory behaviour. Then the wavelet method and the estimator employed is introduced, along with the relevant findings. This includes a discussion on wavelet analysis, the discrete wavelet transform, the estimator employed and the relevant findings. Section V provides the forecast evaluation techniques used and the out-of-sample forecast results. Section VI considers the evaluation of value-at-risk in the context of the Basle adequacy criteria. Section VII concludes and identifies topics for further research.

2. Long memory in time series

Interest in long memory (or long range dependent) processes can be traced to the examination of data in the physical sciences. Formal models with long memory initially pertained to hydrological studies investigating how to regularise the flow of the Nile river in view of its nonperiodic (flooding) cycles. [7] This feature was described as the "Joseph effect" alluding to the biblical reference in which seven years of plenty where to be followed by seven years of famine. [8] In this sense, long memory process concern observations in the remote past that are highly correlated with observations in the distant future. The implications of long memory in financial markets was related to the use of Hurst's 'rescaled range' statistic to detect long memory behaviour in asset return data. [9] It was observed that if security prices display long memory then the arrival of new market information cannot be arbitraged way, which in turn means that martingale models for security prices cannot be derived through arbitrage. As such, long memory processes can be characterised as having fractal dimensions, in the form of non-linear behaviour marked by distinct but nonperiodic cyclical patterns and long-term dependence between distant observations. [10]

A variety of measures have been used to detect long memory in time series. For example, in the time domain, long memory is associated with a hyperbolically decaying autocovariance function. Meanwhile, in the frequency domain, the presence of long memory is indicated by a spectral density function that approaches infinity near the zero frequency; in other words, such series display power at low frequencies. [11] Finally, a pattern of self-similarity in the aggregated sequences of a time series is an indicator of long memory (this refers to the property of a self-similar process, in which, different time aggregates display the same autocorrelation structure). These notions have led several authors to develop stochastic models that capture long memory behaviour, such as the fractionally-integrated I(d) time series models. [12-13] In particular, fractional integration theory asserts that the fractional difference parameter which indicates the order of integration, is not an integer value (0 or 1) but a fractional value. Fractionally integrated processes are distinct from both stationary and unit-root processes in that they are persistent (i.e., they reflect long memory) but are also mean reverting and as a consequence provide a flexible alternative to standard I(1) and I(0) processes. [14] Specifically, the long memory parameter is given by $d \in (0, 0.5)$ while when $d > 0.5$ the series is nonstationary and when $d \in (-0.5, 0)$ the series is antipersistent.

Since, non-zero values of the fractional differencing parameter imply dependence between distant observations, considerable attention has been directed to the analysis of fractional dynamics in financial time series data. Indeed, long memory behaviour has been reported in the returns of various asset classes. [15] Against this background, a rapidly expanding set of models has been developed to capture long memory dynamics is asset return data.

3. Data description

The data analysed in this study are obtained from the global bond index (GBI) series for emerging markets (EM) compiled by JP Morgan. In particular, the fixed income data used comprise of daily total returns for Hong Kong, Mexico, and South Africa from December 31, 2001 to April 9, 2012, representing 2571 observations.

More formally, the change in the local bond index $\hat{B}$ can be expressed as:

$$\hat{B}/\hat{B}_{t-1} - 1 = \Re_t + \gamma_t \tag{1}$$

where $\Re$ is the local currency return and γ is the currency return. $\Re_t = \left(yield_t - yield_{t-1}\right) \times DV01_t + coupon\ return_t$ and $\gamma_t = \left(1 + \Re_t\right)\left(FX_t / FX_{t-1} - 1\right)$ and the coupon return is the return derived from the interest payment made on the fixed income product. Therefore, when an investor buys a local market index, equation (1) suggests that fixed income returns can be decomposed into its predictable coupon return, FX changes, and changes in local yields. Furthermore, in order to compute return volatility, this analysis focuses on squared daily returns, as a proxy for the volatility of the selected emerging markets. In addition, the volatility series is standardised prior to further analysis.

4. Empirical methodology and results

4.1. Preliminary observations

Table 1 presents the time series properties of the data using some basic methods. The results of the Augmented Dickey-Fuller (ADF) unit root test offer evidence in favour of stationary fixed income returns. While this test may be deficient in terms of its ability to capture an order of integration that may not be an integer, the finding of stationary bond returns is consistent with those of many previous studies. [15] However, based on the standard normality and Lagrange Multiplier ARCH tests, fixed income return data exhibit non-normality and ARCH effects. [16-17] These non-white noise characteristics of the data motivate estimation of GARCH(1,1) model using the assumption of the Student *t* distribution.

	Mexico	South Africa	Hong Kong
Mean	0.040496	0.040533	0.017929
Standard deviation	0.339043	0.381478	0.186277
Skewness	0.538071	-0.220239	-0.008431
Kurtosis	20.74158	8.738281	6.603305
Normality test	33829**	3546**	1390**
ARCH (5) test	64.97**	92.71**	72.38**
ARCH (10) test	42.23**	53.92**	47.92**
ADF unit root test:			
Constant	-46.72**	-33.53**	-61.55**
Constant & Trend	-46.81**	-32.88**	-61.57**

Note:
1/ '**' and '*' indicate statistical significant at the 1% and 5% levels, respectively.
2/ Normality test follows a Chi-squared distribution
3/ ARCH (x) test follows an F-statistic with parameters (x, n-x)

Table 1. Description of the Data

4.2. GARCH(1,1) model

The GARCH (1,1) specification comprises a return (or mean) and a variance equation. In particular, the returns generating process can be described by:

$$r_t = \mu + \varepsilon_t \quad \text{where} \quad \varepsilon_t \left| \Phi_{t-1} \sim N\left(0, h_t\right) \right. \tag{2}$$

where r_t denotes the returns process, which may include autoregressive and moving average components, and ε_t is the error term, which is assumed to be normally distributed with zero mean and variance h_t, given the information set Φ_{t-1}. The conditional variance is modelled as:

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{3}$$

For h_t to be well-defined, ω, α, and β are constrained to be non-negative. In addition, the unconditional variance is given by $\sigma^2 = \omega / 1 - \alpha - \beta$ and for a finite unconditional variance to exist $\alpha+\beta < 1$. Furthermore, in the GARCH model the effect of a shock on volatility decays exponentially over time and the speed of decay is measured by the extent of volatility persistence (which is reflected in the magnitude and significance of the summation of the α and β parameters).

The GARCH (1,1) model estimates are reported in Table 2. The results confirm the previous findings on the importance of GARCH effects by showing that the GARCH and ARCH terms are all statistically significant. The parameters of the conditional variance equations are all positive and statistically significant. Furthermore, they satisfy the positivity constraint for the GARCH(1,1) model. Furthermore evidence of persistence in variance as measured by the GARCH model is reflected in the magnitude and significance of the ARCH and GARCH terms (indeed, as this sum approaches unity the greater the degree of persistence). Therefore, in order to have an indication of long memory in fixed income return volatility the level of volatility persistence (i.e., $\alpha+\beta$) is assessed.

The results indicate that volatility in these markets is very persistent, with the level of volatility persistence being 0.9775 for Mexico, 0.9782 for South Africa and 0.9912 for Hong Kong, which underscores the highly persistent nature of shocks to volatility, which also in turn is suggestive of a long memory component to volatility behaviour in these fixed income markets. The models' appropriateness has also been checked by applying the Box-Pierce Q statistic test to standardised and squared standardised residuals. Basic diagnostics indicate that the GARCH models are well-specified.

4.3. Wavelet analysis

To estimate the long memory parameter d of emerging market local currency debt of Hong Kong, Mexico and South Africa this study considers methods based on the discrete wavelet transform. Wavelet analysis, plays an important role in the characterisation of time series, by detecting scaling structures in data. More precisely, since wavelets are localised both in time and frequency, they provide a means to collect information on both the frequency and time characteristics of a time series. In fact, wavelets are already widely used as detectors of patterns in areas as diverse as digital signal processing and exploration geophysics. In the empirical finance literature they have recently been used to determine time-dependence in asset return data by comparing the scale decompositions of observations that exhibit significant autocorrelation between observations widely separated in time. [18] In this manner, the scaling properties of daily returns of emerging market government bonds are analysed in order to capture temporal dependencies in the volatility process.

	Mexico	South Africa	Hong Kong
ω	0.0029 [0.0006]**	0.0031 [0.0007]**	0.0003 [0.0001]*
α	0.1245 [0.0166]**	0.0756 [0.0118]**	0.0510 [0.0089]**
β	0.8530 [0.0170]**	0.9026 [0.0123]**	0.9402 [0.0095]**
Q(5) 1/	1.2197	4.0842	1.0838
Q(5) 2/	1.8322	5.4248	1.3785
Sign bias test	1.1775	2.1147	2.6833
Negative size bias test	1.3357	1.4692	2.2911
Positive size bias test	3.1661**	3.3072**	5.6812**
Joint test	7.7588**	8.3955	11.3702**

Notes:
1/ The Ljung-Box Q test applied to standardized residuals.
2/ The Ljung-Box Q test applied to squared standardised residuals
3/ The numbers in () and [] refer to lag lengths and standard deviations
4/ '**' and '*' indicate significance at the 1% and 5% levels

Table 2. GARCH Estimates

A wavelet is defined as a wave-like function with an amplitude that oscillates around zero and has a finite or quickly decreasing time support. These functions are well suited to locally approximating variables in time or space as they have the ability to be manipulated by being either 'stretched' or 'squeezed' so as to mimic the series under investigation. [19] The power of wavelet analysis is that it makes it possible to decompose a time series into its high- and low-frequency components, which are localised in time. Wavelets, also allow the selection of an appropriate trade-off between resolution in the time and frequency domains, while traditional Fourier analysis stresses resolution in the frequency domain at the expense of the time domain. [20] Wavelets therefore provide a convenient and efficient method to analyse complex signals. [21] Wavelet theory is applicable to several subjects. They are especially useful where a signal (e.g. long memory) lasts for a finite time or shows markedly different behaviour in different time periods. These methods, have emerged as a useful tool in the empirical finance literature where long-run and short-run relationships can be distinguished. [22]

4.4. The discrete wavelet transform

A discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discreetly sampled, as is often the case in econometric analysis. A wavelet transform is a scaling function used to transform a signal into a father (Φ) and mother (μ) wavelet, where the former, are representations of a signal's trend component (i.e., the approximation coefficients) and the latter represent the deviations from the trend component (i.e., the detail coefficients). The discrete wavelet series approximation to a continuous signal *f(t)* is given by

$$f(t) \approx \sum_k a_{j,k}\Phi_{J,k}(t) + \sum_k d_{j,k}\mu_{J,k}(t) + \sum_k d_{j-1,k}\mu_{J-1,k}(t) + \cdots + \sum_k d_{1,k}\mu_{1,k}(t) \qquad (4)$$

where j is the number of multi-resolution scales, and k ranges from 1 to the number of coefficients in the corresponding scale and the coefficients $a_{J,k}, d_{J,k}, \ldots, d_{1,k}$ are the wavelet transform coefficients. Applications of wavelet analysis with respect to time series analysis make use of a DWT. The DWT maps the vector $\mathbf{f} = (f_1, f_2, \ldots, f_n)'$ to a vector of wavelet coefficients, which contains $a_{j,k}$ and $d_{j,k}$, $j = 1, 2, \ldots, J$, which are the approximation and detail coefficients, respectively.

In the empirical literature, Haar and Daubechies wavelets represent typical wavelets and have been used in the characterisation of the time series properties of asset return data. The Haar wavelet is the simplest wavelet and provides a basis for studying more complex wavelets.[2] Since a wavelet is used to decompose a given function into different scale components, it follows that each scale component can then be studied with a resolution that matches its scale. [23]

The data used in this study are discreetly sampled, accordingly, the discrete wavelet transform is used, which permits the generation of the approximation coefficients, $(a_{j,k})$, which capture the trend of a time series and the detail coefficients, $(d_{j,k})$, reflecting the deviations from the trend at each scale. Because the original function can be represented in terms of a wavelet expansion, data operations can be performed using the corresponding wavelet coefficients. This leads to a continuum of time-scale representations of the signal, all with different resolutions. Hence, multi-resolution analysis, which allows the computation of the coefficients corresponding to the wavelet transform of the observed time series.

The analysis of fractionally integrated processes through the use of wavelets is based on the result that the detail coefficients of a zero mean long memory process are asymptotically normally distributed with variance $\sigma^2 2^{-2(J-j)d}$, where σ^2 is a finite constant, j is the scaling parameter and d measures long memory in the relevant volatility series. [24] To estimate d using wavelet theory, the logarithmic variance transformation regression estimator is widely used. This procedure is based on the exploitation of the variance of the detail coefficients at each scale, which generates a statistically consistent estimator of the long memory parameter. The estimator of the parameter of the fractional integration, d, is based on the following least squares regression:

---

[2] The Haar transform assumes a discrete signal and decomposes the signal into two sub-signals of half its length reflecting the trend process and fluctuations from the trend process.

$$\ln Var\left(d_{j,k}\right) = \ln\sigma^2 + d\ln 2^{-2(J-j)} + \varepsilon \quad (5)$$

where $Var\left(d_{j,k}\right)$ is the variance of the detail coefficients associated with the value of the scaling parameter j, where, j =1, . . ., J, and ε is a random error term. Since the variance of the detail coefficients decomposes the variance of the initial time series over different scales, this permits an analysis of the dynamics of the series at each scale.

To estimate d, this study uses multi-resolution analysis via the Haar wavelet to generate the respective detail coefficients of each volatility series at each dimension of scaling parameter. From here, the variance of the detail coefficients at each scale are computed and then the regression specified in equation (5) is performed, where the slope coefficient provides an estimate of d. To check for the robustness of the results, and, therefore, avoid spurious conclusions of long memory dynamics the Daubechies 4 (D4) wavelet is also examined.

The regression results of this analysis are presented in Table 3 and 4. The slope coefficient of the regression given in equation (3) provides an estimate, d. Table 3 shows that, when the Haar wavelet is used, this study is able to find evidence of long memory across the three volatility measures used. The long memory parameter, d, ranges from 0.2363 (Mexico) to 0.3423 (Hong Kong). Furthermore, the evidence obtained is significantly different from zero for all the fixed income markets. These results indicate that volatility realisations have a predictable component insofar as distant observations in the volatility series are associated with each other, albeit over long lags. The significant size of d obtained from this model illustrates the importance of modelling long memory in fixed income data. Furthermore, the result of $d \in (0, 0.5)$ from these models is in contrast to the findings from the unit root tests that led to a conclusion of $d = 0$.

Volatility series	Identifier	Parameter Estimate	Standard Error	R^2
Hong Kong	Intercept Slope (d)	1.4167** 0.3423	0.1288 0.0310	 0.9396
Mexico	Intercept Slope (d)	1.0344* 0.2363**	0.2958 0.0281	 0.9164
South Africa	Intercept Slope (d)	1.0151* 0.2679**	0.1989 0.0664	 0.9172

Notes: To estimate the long memory parameter d, the following regression is performed on the respective volatility series: $\ln Var\left(d_{i,k}\right) = \eta + d\ln 2^{-2(J-j)} + \varepsilon$, where $Var\left(d_{i,k}\right)$ is the variance of the detail coefficients corresponding to the value of the scaling parameter j = 1,. . . , J, and ε is the error term. "**" and "*" indicate statistical significance at the 1% and 5% levels, respectively.

Table 3. Estimates of the Long Memory Parameter using the Haar Wavelet

In econometric analysis, it is important to perform diagnostic checks in order to assess the validity of the initial estimates of d. Therefore, to avoid spurious evidence of long memory (due to the choice of wavelet employed) in the volatility process of the time series, equation

(4) is re-estimated using the Daubechies 4 (D4) wavelet. These results are presented in Table 4. The results are broadly similar in magnitude to those obtained using the Haar wavelet. The noticeable exception relates to the case of South Africa where the long memory parameter falls from 0.2679 (when the Haar wavelet is used) to 0.1784 (when the D4 wavelet is used). This notwithstanding, the results are all statistically significant. In sum, the results of this analysis suggest that bond return volatility in emerging markets is characterised by stochastic processes which have a long memory component.

Volatility series	Identifier	Parameter Estimate	Standard Error	R^2
Hong Kong	Intercept Slope (*d*)	1.0822** 0.3577**	0.0105 0.0887	 0.8953
Mexico	Intercept Slope (*d*)	1.5824** 0. 2611*	0.1996 0.1083	 0.9076
South Africa	Intercept Slope (*d*)	1.6585** 0.1784**	0.2210 0.1575	 0.9412

Notes: To estimate the long memory parameter *d*, the following regression is performed on the respective volatility series: $\ln Var\left(d_{i,k}\right) = \eta + d \ln 2^{-2(J-j)} + \varepsilon$, where $Var\left(d_{i,k}\right)$ is the variance of the detail coefficients corresponding to the value of the scaling parameter $j = 1, \ldots, J$, and ε is the error term. '**' and '*' indicate statistical significance at the 1% and 5% levels, respectively.

Table 4. Estimates of the Long Memory Parameter using the Daubechies 4 Wavelet

The analysis indicates robust evidence of long memory behaviour in the return volatility of emerging market debt. Further, wavelet methods provide a robust fit for the data as evidence by the R^2 readings presented in the final columns of both Table 3 and 4. If fixed income data exhibit long memory, then it displays significant autocorrelation between distant observations. This, in turn, implies that the series realisations may have a predictable component; and, perhaps, past trends in the data can be used to predict future volatility. Therefore, attention now turns to an exploration of the forecast performance of models with long memory relative to the standard volatility models.

5. Volatility forecasting

The evidence accumulated so far suggests that fixed income return volatility in emerging markets follows a long memory process. This, in turn, implies the existence of fractional dynamics in the data which may be exploited to potentially construct improved volatility forecasts, especially over longer forecasting horizons. In order to evaluate the forecasting performance of long memory models (especially over long(er) horizons) versus short memory models (i.e., the GARCH model), the respective data sets are simply split in half and then each model is estimated for all series covering the first part of the sample and then these estimates are used to forecast volatility over the sample period covered by the second half of the data. In this manner, out-of-sample forecast accuracy is evaluated. In addition to calculating the daily forecasts, this study also calculates monthly forecasts using the well-known property that volatility forecasts are additive, such that the sum of five daily volatility forecasts produces the weekly forecasts. And, the summation of weekly forecasts produces monthly forecasts.

In addition to the GARCH and long memory model the RiskMetrics model is also considered for comparative purposes. The RiskMetrics model was popularised by the investment bank JP Morgan and is widely used by financial institutions to model and forecast volatility, especially in the context of the Basle Committee adequacy criteria. This model is essentially an exponentially weighted moving average (EWMA). Under the EWMA, the fitted variance from the model, h_t, which provides the multi-step ahead volatility forecast, is a weighted function of the immediately preceding volatility forecast and actual volatility is given below:

$$h_t = \lambda h_{t-1} + (1-\lambda)\hat{h}_{t-1} \tag{6}$$

where $0 \le \lambda \le 1$ is the smoothing parameter, such that when $\lambda = 0$ the model reduces to a random walk process and when $\lambda = 1$ the model is equivalent to the prior period forecast of volatility. The value of λ is determined empirically by the value that minimizes the 'in-sample' sum of squared prediction errors. In this study λ is set to 0.94 following standard market practice, which is also consistent with previous research which indicates that this value produces accurate forecasts. [25]

5.1. Standard forecast evaluation

Two standard symmetric measures are used to evaluate forecast accuracy, namely, the mean absolute error (MAE) and the root mean square error (RMSE). They are defined below:

$$MAE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}\left|h_t^f - r_t^2\right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{\tau}\sum_{t=T+1}^{T+\tau}\left(h_t^f - r_t^2\right)^2} \tag{8}$$

where τ is the number of forecast data points and r_t^2 is the proxy for volatility . Both the MAE and RMSE assume the underlying loss function to be symmetric. Furthermore, under these evaluation criteria the model which minimises the loss function is preferred.

Table 5 reports out-of-sample performance of the estimated models based on the MAE and RMSE forecast error statistics. At the daily level, the results are not unexpected. That is, the GARCH model dominates forecast accuracy for South Africa on the basis of both the MAE and RMSE. For Mexico, the GARCH model dominates forecast performance on the basis of the MAE while the RiskMetrics models delivers the most accuracy when the RMSE is used as a criterion. For Hong Kong the GARCH process is preferred on the basis of the MAE, which surprisingly, the long memory model delivers the best performance when the RMSE is used as a reference. However, in some cases the forecast accuracy of all the models are close; for instance at the daily level the forecast MAE statistics for the GARCH, RiskMetrics and FIGARCH models are virtually indistinguishable. More generally, the findings of GARCH superiority at the daily level are consistent with a wide empirical literature attesting to the forecast superiority of the GACH model at forecasting volatility over daily frequencies or short horizons.

Model	GARCH		RiskMetrics		FIGARCH	
Forecast Error Statistic	MAE	RMSE	MAE	RMSE	MAE	RMSE
Hong Kong	1.84e-05*	2.82e-04	3.36e-05	8.09e-04	2.73e-04	1.92e-05*
Mexico	1.91e-04	3.18e-04*	4.49e-05*	5.72e-04	6.22e-05	6.73e-04
South Africa	2.63e-05*	1.83e-04*	2.82e-04	1.95e-04	2.88e-05	2.31e-04

Notes: '*' indicates the preferred model.

Table 5. Daily Forecast Results

At the monthly level (i.e., at a longer horizon) the GARCH model also delivers the most accurate results. This finding is surprising. Long memory implies that widely separated observations are associated with each other which in turn suggests that volatility realizations are connected over long lags. The results shows that at even comparatively longer horizons the GARCH model still delivers the most accurate volatility forecasts. Indeed, Table 6 shows that the forecast MAE statistics for Mexico and South Africa are 3.13e-03 and 3.92e-03, respectively, which are smaller than those from long memory models. The same results holds true for the forecast RMSE error statistics. For Hong Kong fixed returns This result appears to suggest that long memory models while theoretically appealing are not particularly helpful in deriving accurate volatility forecast especially over long horizons.

Model	GARCH		RiskMetrics		FIGARCH	
Forecast Error Statistic	MAE	RMSE	MAE	RMSE	MAE	RMSE
Hong Kong	4.27e-04	2.26e-03*	2.35e-03*	4.09e-03	4.39e-04	4.17e-03
Mexico	3.13e-03*	4.66e-03	5.72e-03	4.31e-03*	4.58e-03	6.30e-03
South Africa	3.92e-03*	4.23e-03*	4.89e-03	4.82e-03	4.27e-03	4.31e-03

Notes: '*' indicates the preferred model.

Table 6. Monthly Forecast Results

6. Value-At-Risk evaluation

VaR is a widely used measure to capture the exposure of a portfolio to market risk. The VaR of a position describes the expected maximum loss over a target horizon within a given confidence interval due to an adverse movement in the relevant fixed income yield (or price). VaR is now widely used as an internal risk management tool by financial institutions and as a regulatory measure of risk exposure. [26] In addition, the VaR method is the cornerstone of the 1996 market risk amendment to the Basle Accord (Bank of International Settlements, (BIS), 1996). The Basle Accord prescribes the VaR method in order that financial institutions can meet the capital requirements to cover the market risk they incur in the process of their daily business operations. Under this framework, operational evaluation takes the form of backtesting volatility forecasts and exception reporting.

In particular, the Basle Accord stipulates that for the purpose of calculating regulatory market risk capital it is required that VaR estimates be calculated at the 99 percent probability level using daily data over a minimum sample period of at least one business year (equivalent to 250 trading days) and that these estimates be updated at least every quarter (i.e., 60 trading days). Against this background, the well-known delta-normal specification is employed:

$$VaR = N_{\alpha} h^{f} 3V \tag{9}$$

where N_{α} is the appropriate standard normal deviate, h^{f} is the volatility forecast, the number three represents the minimum regulatory Basle multiplicative factor and V is the initial portfolio value. While Basle Accord prescribes a 99 percent probability the 97.5 percent and 95 percent confidence level is also examined for greater comprehensive and consistency with previous studies. The validity of such VaR calculations are assessed or 'backtested' by comparing actual daily trading (net) losses with the estimated VaR and noting the number of 'exceptions', in the sense of days when the VaR estimate was insufficient to cover actual trading losses. Regulatory scrutiny is therefore triggered where such exceptions occur frequently, and in practice this leads to a range of penalties for the financial institution concerned. [27]

In line with the rolling window approach to VaR evaluation mandated by the Basle Committee rules initial volatility forecasts and VaR measures are constructed over intervals of 60 trading days, with the initial estimation sample then rolled forward and the models updated every 60 observations before the next set of volatility forecasts are produced. The first 3-years of data (representing 752 observations) are used for initial model parameter estimation, leaving 1819 observations for volatility forecasting and the construction and evaluation of VaR measures. Specifically, this provides 30 sub-samples of 60 trading days length over which VaR is assessed. This assessment is conducted through appraisal of the out-of-sample VaR failure rates associated with VaR measures constructed using the forecast values of those volatility measures. The assessment of VaR performance is conducted through appraisal of the 'out-of-sample' VaR failure rates associated with VaR measures constructed using the forecast values derived from the GARCH, RiskMetrics and long memory model. The focus on the 'out-of-sample' failure rates is motivated by the requirements of risk managers, who obtain VaR estimates in real time and must use parameters obtained from an already observed sample in order to evaluate the risks associated with current and future random movements in risk factors. As a result, credible test of VaR construction methods under alternative volatility forecasting models is their performance outside the sample used to estimate the underlying parameters.

Table 7 reports the out-of-sample VaR failure rates. The results are very diverse and highlight that in many of the markets considered the forecasting model that minimises the percentage number of daily VaR exceedances is sensitive to the specification of the probability level. When the Basle Committee rules are applied (i.e., the 99 percent probability level) the results indicate that the GARCH and RiskMetrics that provide the exceedance-minimising method for the fixed income markets considered. At the 99 percent probability level the long memory model is the generally the weakest performer. However, in the case of Hong Kong and South Africa the long memory model is second best model in terms of delivering accurate VaR measures. In addition, it is important to note that in many cases is level of accuracy between the various models is close as reflected by the closeness of the VaR failure rates. At the 97.5 and 95 percent probability levels model performance is more varied with all models demonstrating varied degrees of accuracy. As a generalization, these results are mixed but the evidence suggests that at the Basle prudential level the simpler models help in providing improved VaR estimates that minimise occasions when the minimum capital requirement identified by the VaR methodology would have fallen short of actual trading losses.

	Hong Kong			Mexico			South Africa		
model	99%	97.5%	95%	99%	97.5%	95%	99%	97.5%	95%
RM	0.0178*	0.0347	0.0224*	0.0154	0.0326	0.0378	0.0224	0.0311	0.0218*
GARCH	0.0192	0.0256*	0.0536	0.0152*	0.0312	0.0356*	0.0192*	0.0286*	0.0261
FIGARCH	0.0185	0.0391	0.0493	0.0179	0.0297*	0.0521	0.0222	0.0303	0.0323

Notes:
1. VaR is value-at-risk
2. RM is the RiskMetrics model, GARCH is the generalized autoregressive conditional heteroskedasticity and FIGARCH is the fractionally integrated GARCH.
3. Model failure rates are the number of exceptions divided by the number of observations.
4. '*' (asterisks) denote the preferred model.

Table 7. VaR Failure Rates – Out-of-Sample

7. Conclusions

Recent empirical evidence concerning the nature of volatility dynamics in fixed income markets suggests the existence of a long memory component. Since volatility in fixed income returns is an important aspect of portfolio management it is essential to accurately characterise the time series properties of fixed income volatility especially in the context of emerging markets where local currency-denominated sovereign bonds have been the fastest growing market in recent years. Accordingly, the objective of this analysis was to examine the existence of long memory behaviour in the volatility structure of total return indices for the local currency bond markets of Hong Kong, Mexico and South Africa. Against this background, the long memory parameter is estimated using methods based on wavelets, which have gained prominence in recent years. Furthermore, this study has compared and evaluated the performance of a long memory model versus a standard volatility models (the ubiquitous GARCH and RiskMetrics processes) in order to evaluate their power in delivering accurate volatility forecasts over long(er) horizons in an out-of-sample setting. This endeavour is motivated by recognition of the importance of accurate volatility forecasts in a wide range of applications, including tactical and strategic decision making and the limited empirical evidence available to date for emerging fixed income markets. Then, the performance of the standard GARCH, RiskMetrics and FIGARCH models are evaluated in the context of value-at-risk (VaR) estimation given the Basle regulatory framework.

The main findings of this research are threefold. First, evidence of long memory is conclusively demonstrated in emerging market local currency sovereign debt markets. In addition, to counteract the possibility of finding spurious evidence of long memory a variety of wavelet forms are considered. The findings from these tests are complementary and therefore suggest that the finding of long memory is not spurious. Second, the presence of a long memory structure in the volatility of these fixed income markets suggests volatility observations in the recent past and the remote past are associated with each other. Since the series realisations are not independent over time then past volatility may potentially be exploited to predict future volatility, especially over long horizons. Accordingly, the out-of-sample forecasting performance of the long memory model and the standard GARCH and RiskMetrics models are compared. While, none of the estimated models consistently outperforms the others, a key generalisation can be made. In particular, on the basis of the forecast MAE and RMSE statistics it is shown that the information content of long memory models does not consistently generate improved volatility forecasts, especially over long horizons, relative to the standard GARCH model. Indeed, the GARCH model generally provides the most accurate forecasts at the monthly horizon. With respect to VaR estimation, the results show that both the standard GARCH and RiskMetrics models generally deliver more accurate VaR measures relative to the long memory process.

These findings have three important implications. First, the exploitation of long memory models based on wavelet analysis may not have great relevance in the context of emerging market debt in terms of delivering superior forecast performance. Second, the existence of a long memory structure in volatility is not an essential condition for the derivation of accurate volatility forecasts at any time horizon, especially over a long horizon. Indeed, this research suggests that long memory models appear to be of limited practical forecast value, especially over long horizons, for Hong Kong, Mexico and South Africa. Put differently, the computational complexity of long memory modelling is not commensurate with the benefits (in terms of forecast power). Third, the results of the VaR estimation may provide guidance on more effective prudential standards for operational risk measurement and, as result, may help ensure adequate capitalisation and reduce the probability of financial distress. The results highlight the importance of using out-of-sample forecasting techniques and the stipulated probability level for the identification of methods that minimise the occurrence of VaR exceptions. Standard models – RiskMetrics and GARCH – that are already widely used by market participants are generally shown to outperform the more computationally intensive wavelet-derived FIGARCH model in estimating VaR across the probability levels considered.

In sum, this research has evaluated the long memory properties of return volatility in fixed income markets. This paper also complements the literature on long memory models and the forecast performance of these models that has attracted interest in other asset classes. In addition, the results of this study may potentially be used to inform portfolio and risk analysis. In particular, it is shown that in the context of VaR estimation existing models based on the GARCH and/or RiskMetrics process are more accurate (and simpler) than their long memory counterpart. Some caveats to these results exist, however. First, squared returns provide a noisy proxy for the 'true' volatility. In the case of this analysis, data constraints limited alternative options. However, future research may find that the application of realised variance may produce more accurate forecasts. Second, future research may also consider exploring the relevance of other long memory models, for example, models with asymmetric effects given that market volatility is often reported as being 'directional', i.e., volatility is higher in a down – than an upmarket.

Author details

Pako Thupayagale
Bank of Botswana, Botswana

8. References

[1] McCarthy J, Pantolon C, Li HC. Investigating long memory in yield spreads. J Fixed Income. 2009; 19(1) : 73-81.

[2] Schotman PC, Tschernig R, Budek J. Long memory and the term structure of risk. J Financ Econ. 2008; 6: 459-97.

[3] McCarthy J, DiSario R, Saraoglu H, Li H. Tests of long-range dependence in interest rates using wavelets. Q Rev Econ & Financ. 2004; 44: 180-9.

[4] Emerging market trading association. Third quarter 2011 debt trading volume survey supplementary analysis. 2011. EMTA.

[5] Bollerslev T, Mikkelsen HO. Modelling and pricing long memory in stock market volatility. J Econometrics. 1996; 73: 151-84.

[6] Gencay R, Selcuk F, Whitcher B. Scaling properties of foreign exchange volatility. Phys A, 2001; 249-66.

[7] Hurst HE. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers. 1951; 116: 770-99.

[8] Mandelbrot BB, Wallis J. Noah, Joseph and operational hydrology. Water Resources Research. 1968; 4: 909-18.

[9] Mandelbrot BB. When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models. Rev Econ Stat. 1971; 53: 225-236.

[10] Mandelbrot BB. Fractals: form, chance and dimensions. New York: Free Press. 1977.

[11] Lo A. A long-term memory in stock market prices. Econometrica. 1991; 59(5): 1279-1313.

[12] Granger CWJ, Joyeux R. An introduction to long-memory time series models and fractional differencing. J Time Series Analysis. 1980; 1(1), 15-29.

[13] Hosking JRM. Fractional differencing. Biometrika. 1981; 1(1), 165-76.

[14] Baillie RT. Long memory processes and fractional integration in econometrics. J Econometrics. 1996; 73(1): 5-59.

[15] Poon SH. A practical guide to forecasting financial market volatility. John Wiley & Sons Ltd: Chichester; 2005.

[16] Jarque CM, Bera AK. A test of normality of observations and regression residuals. Int Stat Rev. 1987; 55:163-72.

[17] Engle RF. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica. 1982; 50: 987-1007.

[18] DiSario R, Li H, McCarthy J, Saraoglu H. Long memory in the volatility of an emerging equity market: the case of Turkey. J International Financial Markets Instittions & Money. 2008; 18: 305-12.

[19] Jensen M. Using wavelets to obtain a consistent ordinary least squares estimator of the fractional differencing parameter. J Forecasting. 1999; 18: 17-32.

[20] Gencay R, Selcuk F, Whitcher B. An introduction to wavelets and other filtering methods in finance and economics. Academic Press: San Diego; 2002.

[21] Ramsey JB. Wavelets in economics and finance: past and future. Stud Nonlinear Dyn E. 2002; 6(3): 1-27.

[22] Aussem A, Campbell J, Murtagh F. Wavelet-based feature extraction and decomposition strategies for financial forecasting. J Computational Intelligence Financ. 1998; March/April: 5-12.

[23] Mallat, SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989; 11: 674-693.

[24] Jensen M. An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. J Econ Dyn Control. 2000; 24: 361-387.

[25] Fleming J, Kirby C, Ostdiek B. The economic value of volatility timing. J Financ. 2001; 56: 329-455.

[26] Jorion P. Value at risk. 3rd ed. New York: McGraw-Hill; 2007.

[27] Saunders A, Cornett MM. Financial institutions management: a risk management approach. 4th ed. New York: McGraw-Hill; 2003.

Chapter 6

Risk Management in Collaborative Systems

Marius Ioan Podean and Dan Benta

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50844

1. Introduction

Many domains reached a point in which the knowledge required for skillful, professional practice can no longer be acquired in a decade, factor that generates increased specialization [2]. This increased specialization makes collaboration crucial because complex problems require more knowledge than any particular person possesses. The relevant information required to solve complex problems is normally distributed among different persons or stakeholders. In order to create insight, new ideas or new artifacts it is considered a prerequisite to bring different and often controversial points of view together, and create a shared understanding among stakeholders. It is generally considered that insight moments for creative individuals are the result of working in isolation, but it has been proven that the role of interaction and collaboration is critical [19].

Collaboration is a very dynamic process that combines functionality that supports communication, management and involves content handling. During the execution of a project team members are not always collaborating and their work alternates with cooperation, when a greater emphasis is placed on a value-chain model of producing results. Project management is a tool used to provide a team the capabilities required to produce the benefits defined by vision [24]. Risk management is a critical element in defining the relationship between risks, uncertainty and objectives thus contributing to the chances of success in the execution of a project [9]. In the present context of information society both project and risk management should reconsider collaboration by thoroughly understanding its mechanisms and adapting its tools in order to fully harness it. In this study we will present some of the main aspects regarding collaboration that are relevant in order to build an efficient project and risk management strategy. This will be followed by some approaches from software development that we consider relevant for the context and present our risk management approach for similar projects.

2. Collaboration

The term collaboration is used often when one refers to quite different aspects of working together, like cooperation or even communication [33]. Focusing on national programs that involve agencies, David Osher [30] identifies collaboration as being the most sophisticated level of relationship because it requires efforts to unite people and organizations in order to achieve common goals that could not be achieved by any single individual or organization acting alone.

On the other hand, in [15] collaboration is regarded from a more project management oriented point of view focusing on the elements required to achieve this level of relationship. Collaboration is identified as a process in which entities share information, resources and responsibilities to jointly plan, implement, and evaluate a program of activities to achieve a common goal. Following this author, collaboration is a process of shared creation that involves mutual engagement of participants to solve a problem together and implies sharing risks, resources, responsibilities, and rewards. Sharing risks, resources, responsibilities, and rewards can also give the group to an outside observer the image of a joint identity.

This approach implies that collaboration is more than what will later see that is identified as cooperation by adding joint identity and novelty to the goal. Michael Schrange in his Shared Minds [37] focuses especially on this novelty of the group's goal. Schrange considers collaboration as a process of shared creation: two or more individuals with complementary skills interacting to create a shared understanding that none had previously possessed or could have come to on their own.

Collaboration is often confused with cooperation. Because for many people the two terms are indistinguishable [15], in the following we will take a closer look at what collaboration is and how can it be attained. One general accepted model that describes what is collaboration and what are its main components is the 3C Collaboration Model [15]. This model states that collaboration is attainable by implementing three main processes: communication (networking), coordination and cooperation.

Communication [22] is the starting process in each collaborative process. It is a general belief that efficient or so to say "ideal" communication will provide better common understanding or agreement but when people communicate accurately they realize more precisely the differences that exist on their perspectives of the concepts in use [18]. Different types of agreement tend to mask the differences in perception that accurate communication would uncover.

The second process required by the 3C model is coordination. Coordination refers to [22] the management of people, their activities and resources. Coordination allows team members to manage conflicts and activities in order to increase the efficiency of communication and cooperation efforts. Networking or communication is used as a foundation [45] but involves also altering activities for mutual benefit and for a common goal. It also increases resource usage efficiency and the ability to meet the targets.

Cooperation refers to the interaction among group members in order to produce, manipulate, and organize information, or build and refine cooperation objects like documents, spreadsheets etc. [22] This process requires a shared workspace that should provide the required tools in order to manage these artifacts, tools like version and access control and authorization. The shared workspace is very important because it allows group members to count on group memory and it provides also some basic awareness mechanisms.

Cooperation, coordination and collaboration are often used interchangeably, but they should be describing different stages in the transformation of the relationship between groups and organizations [30].

The backbone of collaboration is not the process of relationship but the strict following of a specific result [18]. We consider that an extra element must be added to the 3C model of collaboration, namely "Creativity". In order to support creativity in collaborative systems we consider that *a)* integrated light-weight project and risk management, and *b)* flexible content management tools have to be implemented [33].

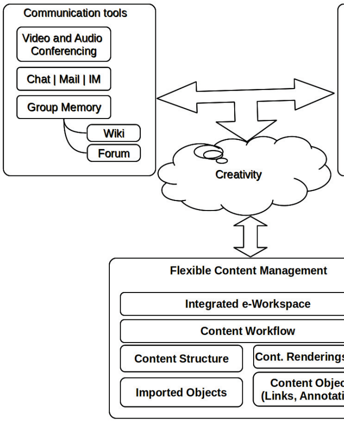


Figure 1. The 4C Collaboration model [33].

In order to discuss about collaboration, a key prerequisite must be satisfied, namely having a joint/ compatible goal or problem to solve (it is not enough that parties have their own individual goals) [15]. A vision specifies the scope and extent of these benefits but does not provide the means to attain them. Project management is a tool used to provide a team the capabilities required to produce the benefits defined by vision [24]. Vision delineates a *strategy* and project management sets the *tactics* by detailing the steps required to put it in practice. Teams can successfully implement their vision if they can bind it with the tactics.

Hilson [24] identified that a "zero risk" zone not also that it does not exist, but it is not even desirable because the available benefits are determined by the degree of risk it is confronted. Risk is defined as [35] "any uncertain event or set of circumstances that, should it occur, would have an effect on one or more objectives". Thus an uncertainty that does not effect the objectives is not a risk, but it can even be an opportunity. Risk management it is an important component in defining the relationship between risks, uncertainty and objectives thus contributing to the chances of success in the execution of a project.

Project management is about making complexity manageable and it is important to collaboration because it provides teams an organized way of keeping in touch with their goals. As defined by The Project Management Institute (PMI) [35], project management is "the application of knowledge, skills, tools, and techniques to project activities to meet the project requirements".

Project risk management provides approaches by which uncertainty can be understood, assessed, and managed within projects. One of the key prerequisites of success in implementing risk management is to have a proactive approach. This process must supplement project management and must follow a holistic approach. Risk management will not provide great benefits to the teams using it if their practices fall under two limitations: focus on tactics and focus on threats [24].

3. Creativity

The rationality behind collaboration is creativity. Creativity, and especially scientific creativity, is a process of achieving an outcome that is recognized as innovative by the relevant community. As defined in [16], this process does not occur in one person's head, but in the interaction between that person's thoughts and a sociocultural context.

Creativity can refer to the work of artists, but can also refer to every-day problem-solving abilities. This type of creativity is essentially, equally significant because enables people to become more productive and make better results. Support for divergent and convergent thinking, development of shared objectives and reflexivity[20] are identified as key requirements for creativity.

Divergent thinking represents the ability to return in response to a challenge or undertaking an extensive set of alternatives, responses, ideas or opinions [20]. On the other hand, convergent thinking represents the ability to narrow down to one set of alternatives and is comparable with consensus building. If a group manages to build trust, uninhibited exchange of ideas and mutual support, each member can use the mind of the other as an extension of his own. At this stage, each person acts as a critic for the other and ideas are reworked into components of a shared vision. Cognitive conflicts within groups that use brainstorming effectively can alleviate divergent thinking. This stimulates thoughtful consideration of new and innovative ideas during brainstorming in a collaborative interaction[20]. Colleagues disagreement can result in evolutionary thinking, viewed as a series of small alterations that cause substantial changes in a concept.

Minority dissent [20] stimulates a reappraisal of a situation manifested in the search for further information and the use of thoughts and strategies about the issue. Some studies have shown that arguments can stimulate creative solutions even when wrong. Dissent stimulates both divergent and convergent thinking; the minority's consistency in advocating his opposition during convergent thinking may cause questions about the majority position, resulting therefore in making better decisions.

Clarity of goals is a necessary requirement for creativity flow [16] but team members have to take in consideration how they express these goals. Having clear objectives helps convergent thinking filter with greater precision. Developing shared objectives is a necessary condition for creativity because it requires group members to share their domain-specific knowledge and generates less resistance to change.

Upgrading from 3C to collaboration starts with the development of shared goals. Shared goals are more that common or compatible goals because they provide the team a joint identity by eliminating the value-chain model of building a common understanding or result. A component that supports the definition of a shared vision and a description of this vision in terms of shared goals is essential. This must allow users to (re)adjust their goals according to the changes in the environment [31].

Obtaining immediate feedback is essential in having complete participation in the task at hand [17, p. 54]. In the context of a group, this refers to the extent to which members collectively reflect on the group's objectives. This process is known as reflexivity and consists of three elements: reflection, planning and action or adaptation [20].

In order to allow creativity to develop in a group, requirements like support for reflexivity must be implemented [20]. Reflexivity is built on processes like:

1. reflection: consists in evaluating, filtering and selection of ideas that will be further taken in consideration.
2. planning: is one the main subprocesses of project management and its main relevance for creativity is that it guides member's attention towards action and means of accomplishing their goals. Good planning reveals hidden aspects of the task at hand and is a good support in motivating team members.
3. action or adaptation: adaptation is actually a simple process of risk management because it represents a continuous adaptation of tasks and objectives in concordance with the changes in the environment. Integrated risk management tools are a must in order to identify, mitigate and define action plans that cover decisions at both strategic and tactics level and include not only threats but opportunities also [24].

In order to make planning more accessible and easy to access process maps are identified as a solution [40]. The process map should be derived from the project management and risk mitigation plan. It's main purpose it to *i)* provide an clear understanding of the hole workflow involved in reaching a goal, and *ii)* to supply a open delineation of the responsibility distribution. This is particularly useful in avoiding responsibility overlapping or uncovered responsibility areas. A mechanism that can manage the sequence of tasks and their execution for a particular process is required in order to enable the coordination process.

Reflection [20] is based on critical thinking, which is a form of thinking that is focused, disciplined, consistent and constrained. Critical thinking can be associated with convergent thinking because it implies evaluating what divergent thinking offers, filters it using acceptability criteria and selects the ideas that will be also taken in consideration.

Planning [20] creates conceptual readiness for relevant opportunities and guides group member' attention towards actions and means to achieve goals. Planning generates high reflexivity if during the process factors like potential problems, hierarchical ordering and short/ long term planning are taken in consideration.

Action or adaptation [20] refers to the continuous renegotiation of group's reality during interaction between group members, and members and the environment. Adaptation consists in goal-directed behaviors that are relevant to achieving the desired changes in group objectives, strategies and processes identified by the group during the stage of reflection. Risk management is used to identify, mitigate and define action plans for the full range of uncertainty, including both risks and opportunities.

Serendipity represents the possibility of making unexpected and fortunate discoveries and consists in *a)* discovering new pieces of information that can lead thinking to new ideas, and *b)* the sagacity needed to make the connections between different pieces of information [1]. Computers can help develop serendipity by *i)* revealing stimulating connections, and *ii)* supporting either the growth or the sharing of an idea.

Creative activities are the result of the relationship between an individual [2] and the realm of his or her work, and other human beings. The resolution of complex problems is based on different aspects of knowledge that reside in the minds of individual stakeholders as tacit knowledge [2]. Full agreement may not always be achievable but what is decisive is to make informed compromises. This can be achieved by using the so called symmetry of ignorance [2]. Symmetry of ignorance is also highly valuable in identifying risks that may affect the team's goal negatively. Stakeholders bring perspectives to the collaborative process, perspectives that are of considerable importance in framing the problem. This concept requires externalizations or boundary objects that capture distinct domains of human knowledge.

At the core of intelligent human performance, is not the individual human mind in isolation, but the interaction [2] of the mind with *(i)* tools and artifacts, and *(ii)* groups of minds in interaction with each other. In the first case, the interaction with externalization objects is straightforward, the knowledge an individual needs is distributed between his head and the world - an address book, a file system, etc. In the second one, the need for such objects becomes more critical because the distribution of knowledge has to be available to all members implicitly.

Externalization objects are essential to collaboration [2] because they *a*) create and store mental effort records, evidence that is outside the memory; and *b*) represent artifacts that provide information and form the basis for critique and debate. Very valuable assets for a group or organization are not only the results but also the way people think, the way they get to god results. It is a significant challenge to try to capture the thinking process in tools that are remarkably easy and intuitive to use.[31]

4. Agile team management

Concerns about principles and practices of agile development approaches are well examined in many works as [28] where authors offer several answers to area. Agile is a people-focused approach [27] and considered effective in terms of team work and collaboration of small teams. People factors that can affect the performance of the team from an agile perspective are analyzed. The ability of adapting to changes efficiently and cost-effectively influenced in a radical way the path to success and survival in business environments [38]. As business processes become more complex, interconnected, interdependent, and interrelated than ever before, traditional approaches [4, 6] are no longer used and are characterized as reflecting linear and sequential processes. Even seemingly minor changes can produce unanticipated effects, as systems become more complex and their components more interdependent. Agile development methodologies were developed in response to this complexity, with a focus on rapid iterative delivery, flexibility and working code [4, 5]. However, for this migration to be successful, organizations must channel the bulk of their change efforts to the development team. The reason is that, the development team serves as a "fulcrum" on which the organization's efforts are applied and through which these efforts are turned into software products for customers [4].

Although agile development methodologies are viewed as an improvement to the traditional methodologies several key considerations need to be taken into account when deciding to migrate to Agile. In order for this migration to be successful we must take into account several variables such as the organizational culture, resistance to change, accepting the new roles by the members, the use of self-organizing teams, the insurance of face to face communication on behalf of the organization and the motivation of the members composing the team.

Traditional teams are characterized as command and control teams while agile ones are characterized as self-organizing teams. From this view, agile teams have a number of key advantages over traditional ones. For this reason, agile teams succeed to deliver products at better quality in less amount of time. A main problem in adopting agile is represented by human resources and their resistance to change. For example the manager might be afraid of losing his power over the members of the group, or becoming “role-less”. Although the roles might change the manager is still needed not for planning and controlling ability, but for the important job of interfacing on the team’s behalf with the rest of the organization [4]. Project managers pass from planning, controlling, directing, and managing, to new roles like building trust, facilitating and supporting team decisions, expanding team capabilities, anticipating and influencing change. They become facilitators, liaisons and network builders, boundary managers, “resource allocators”, team champions and advocates, and in most cases, still have responsibility to watch the project budget. The roles of team members also change, as the group takes on increasing ownership of work processes and the agile practices. They become decision makers, conflict managers, innovators and conveners of standup meetings. These new and unique roles of the project manager and team members engender a more co-operative team with rapid results [4].

The agile methods focus on each team member and offers motivations to achieve a common goal of the team and the project. Building projects around motivated individuals, giving them the environment and support they need, and trusting them to get the job done; is central to the success of an agile team [42]. Different agile methods have different approaches for encouraging motivated individuals in development teams. For example, Extreme Programming, XP [7] has the concept of “collective ownership”, while Scrum [41] has “scrum teams”.

Communication represents a crucial factor on agile methods and treated in a different manner as traditional methods do. Face-to-face communication is preferred rather than written documents and this new mentality increases the quality of coordination and communication. This is in line with one of its core principles that states that, the most efficient and effective method of conveying information to and within a development team is face-to-face conversation [42].

It is well known that organizational culture influences the project trajectory and its implementation environment. Therefore, the only way to succeed in such a change is through a conscious programme of continual reinforcement of the desired behaviors until they become a natural and automatic part of the daily business of the organization [4].

To sum up, although agile development might be an improvement on the traditional development methodologies, an organization looking to migrate to agile needs to take into consideration several aspects regarding organizational culture and other variables that may prevent performance during this transition.

5. Proactive risk management

Project risk management is an essential and determinant step towards successful projects. A detailed analysis and a precise definition of risk can lead to achievement of objectives. PMI states [35] that "the objectives of Project Risk Management are to increase the probability and impact of positive events, and reduce the likelihood and consequences of negative events in the project".

One of the key prerequisites of success in implementing risk management is to have a proactive approach [11]. Risk management must supplement project management and must follow a holistic approach. Project management can be seen as an attempt to control the uncertain environment, through the use of structured and disciplined techniques such as estimating, planning, cost control, task allocation, earned value analysis, monitoring, and review meetings. Each of these project management elements has a role in defining or controlling inherent variability in projects. Project risk management provides approaches by which uncertainty can be understood, assessed, and managed [12, 13].

According to PMI [35], risk management can be implemented using the following sub-processes: plan risk management, identify risks, perform quality and quantitative risk analysis, plan risk responses and monitor and control. Based on the description provided by PMI, we will outline the main aspects involved by these processes.

Planning risk management [35] consists in agreeing on how to conduct risk management activities. This implies that the process should start at the beginning of the project and evaluate the extent and type of risk management with the project's importance and scope. This is essential in order to avoid performing extensive risk analysis on small projects that might not require such a considerable extent. Planning risk management uses as inputs documents like the project scope statement, cost plan, and schedule and will produce documentation that describes risk categories (a clear understanding on the types of risks - technical, external, organizational or project management related) and risk probability and impact (in order to have a common understanding of the terms used in identifying the risks a glossary of terms must be defined).

Identifying risks is an initiative that requires the identification and documentation of the risks that may affect the project. This sub-process should be regarded as an iterative one because new risks may evolve or emerge as the project progresses. Identifying risk requires that all the relevant stakeholders should be involved. Identifying risks is the most complex sub-process, and uses techniques like documentation review, information gathering (tools like brainstorming, interviewing, root cause analysis and Delphi technique[1] can be used), checklist analysis, assumptions analysis, diagramming techniques or SWOT analysis[2]. This subprocess should have as output the list of identified risks and possible responses.

Performing quality risk analysis consists in prioritizing risks by combining their likelihood and impact [32]. Performing quantitative risk analysis requires a numerical analysis of the effects of identified risks on the project objectives. Planing risk responses requires that options

[1] The Delphi technique consists in reaching consensus of experts by requesting ideas about relevant project risks anonymously using questionnaires. The results are summarized and recirculated to the experts for further analysis. This approach reduces bias in the data and restricts a person from having excessive weight on the outcome.

[2] Other techniques for risk assessment can be consulted in [29, p. 41-46]

are developed in order to increase opportunities and reduce threats. In order to determine the appropriate alternatives, strategies like avoidance (changing the project management plan in such a way that the identified risk does not affect it anymore), transfer (requires shifting all or part of the negative effect along with the responsibility to a third party), mitigation (consists in reducing the chances and/or impact of a risk in order to be in satisfactory threshold limits) and compliance. Monitoring and controlling is an ongoing process and consists in implementing risk responses, tracking identified risks and identify new ones.

As we have seen in Section Creativity, identified as an essential requirement for creativity was reflexivity support. Reflexivity consists in the group's focus on its objectives and the actions identified as required in order to reach them. Adaptation, as one subcomponent of reflexivity, consists in goal-directed behaviors that are relevant to achieving the desired changes in group objectives, strategies and processes identified by the group during the stage of reflection. Risk management can be used as a tool by team members in order to respond to changes in the environment.

The Practice Standard for Project Risk Management [36] identifies the following factors as critical for a successful risk management:

1. individual commitment and responsibility (risk management is the responsibility of all team members)
2. open and honest communication. All team members must participate at risk management and avoid actions or attitudes that can hinder communication because it can lead to ineffective risk management
3. risk effort scaled to project (the costs of risk management should be appropriate to its potential value)
4. integration with project management. Risk management can be performed only in strict correlation with project management.

Risk management will not produce substantial benefits to the teams using it if their practices fall under two limitations: focus on tactics and focus on threats [24]. Focusing only on project related issues (like risks related to project processes or performance) or technical functionality and not considering strategic sources of risk that menace the vision a decoupling between objectives and project deliverables will appear. On the other hand, focusing only on the negative side of uncertainty (threats, risks) will lead to ignorance towards opportunities, situations that should be integrated in the risk management plan in order to maximize benefits. This shift requires minor changes in the typical process, but the greatest challenge is to produce the change in people's approach which is centered on the threat aspect of uncertainty.

In [35] the following strategies for handling opportunities are identified:

a) exploitation (eliminate the uncertainty factor and ensure that the opportunity definitely happens);

b) sharing (requires allocation of some or all of the ownership of the opportunity to a third party so that all parties gain from their actions);

c) enhancing (identifying the key driving factors and enhancing their probability and/or positive impact); and

d) accepting (take advantage of an opportunity if it comes, but not actually pursuing it).

An integrated risk management [24] approach is required in order to cover both strategy and tactics, and opportunities and threats. Such an approach can provide increased benefits not available to a limited scope risk management. According to the aforementioned author, these benefits consist in:

- project deliverables become more concordant with the group's needs and vision by reducing the strategy - tactics gap,
- taking in consideration that projects as a tool are used to *a)* produce results that meet specific needs, and *b)* provide benefits, a focus on these aspects is necessary in order to avoid simply producing a series of deliverables,
- avoid having a poor or to late response to opportunities by enabling a proactive management of opportunities at both strategic and tactical level,
- support the best possible decision approach in an uncertain environment by providing relevant information to decision-makers,
- reduce waste and stress and increase productivity and effectiveness by managing uncertainty in advance with planed responses to known risks,
- increase the chances of achieving both strategic and tactical objectives by minimizing threats and maximizing opportunities,
- safe risk-taking leads increased rewards and that requires an appropriate level of risk to be taken with full awareness of the degree of uncertainty and its potential effects on objectives, and
- develop a risk-mature culture that is aware that risks exists at all levels and they can and should be managed proactively.

As we have seen in this section, risk management is a key factor for all teams, regardless of size and work domain, to help them make more informed decisions and take a proactive approach in regard to risks and opportunities.

5.1. Risk mitigation

A project driven organization represents a main approach for most medium and large sized institutions to successful deliver projects. Projects are more often used to transfer all personnel focus and involvement in day to day work, tasks and activities. Thus, concerns as respecting predefined terms become their main objectives. A proper project risk management approach deals with a set of techniques by which inevitable uncertainty can be understood, assessed, and managed within projects. Its main objective is to deliver projects in predefined costs, time and quality respecting customer specifications.

A main problem in project management and project risk management is represented by optimism bias and by human behavior to underweight outcomes that are merely probable and to overweight outcomes with a low probability to happen. This problem was clearly identified and presented by Daniel Kahneman and Amos Tversky [25, 26] in prospect theory with origins in decision making under risk. Prospect theory is based on some well-observed deviations from rationality and helped Kahneman win the 2002 Nobel Prize in Economics.

As a first experiment, they used a group to choose between two options. For first problem (Figure 2), the group had to choose between option A (to win 2500 with a probability of 33% or 2400 with a probability of 66% or there were a 1% chance to win nothing) and option B (to win 2400 certainly). For this first problem, 18% of respondents chose option A, while 82% of respondents chose option B.

*Problem #1: Choose between A or B	
A Win 2500 with a probability of 33% Win 2400 with a probability of 66% Win 0 with a probability of 1%	B Win 2400 with a probability of 100%

Figure 2. Problem 1 from prospect theory

Their choice was influenced by the 1% probability to win nothing even if in option A they had 66% chance to win more than option B and 33% chance to win equally to option B (a 99% chance for option A to win more or equally to option B). Events with low probability (1% chance to win nothing) influence their decision.

According to results (18% for option A and 82% for option B), utility function means: 33% utility to win 2500 plus 66% utility to win 2400 is lower than 100% utility to win 2400.

$$33\%u(2500)+66\%u(2400) < 100\%u(2400)$$

$$33\%u(2500) < 100\%u(2400)-66\%u(2400)$$

$$\mathbf{33\%u(2500) < 34\%u(2400)}$$

For second problem (Figure 3), the group had to choose between option C (to win 2500 with a probability of 33% or there were a 67% chance to win nothing) and option D (to win 2400 with a probability of 34% or there were a 66% chance to win nothing). For this first problem, 83% of respondents chose option C, while 17% of respondents chose option D.

*Problem #2: Choose between C or D	
C Win 2500 with a probability of 33% Win 0 with a probability of 67%	D Win 2400 with a probability of 34% Win 0 with a probability of 66%

Figure 3. Problem 2 from prospect theory

Their choice was influenced by the +100 win and 1% probability difference. The 1% probability difference is now ignored and they chose higher win. Probability difference (1% as difference between 67% and 66%) influence their decision.

According to results (83% for option C and 17% for option D), utility function means: 33% utility to win 2500 is higher than 34% utility to win 2400.

$$\mathbf{33\%u(2500) > 34\%u(2400)}$$

This is exactly the reverse inequality from first problem. This is a clear example how human behavior tends to underweight outcomes that are merely probable and to overweight outcomes with a low probability to happen. As can be observed, second problem is derived from first problem (Figure 4) by elimination of Win 2400 with a probability of 66% option.

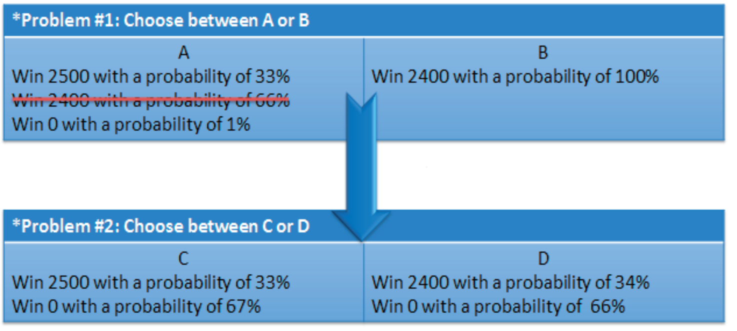


Figure 4. Problem 2 is derived from Problem 1 (from prospect theory)

This theory is main root for Reference Class Forecasting, which for a particular project, aims to identify a relevant reference class of similar projects from the past, to establishing a probability distribution for the selected reference class and to compare the specific project with the reference class distribution for establishing the most likely outcome for the specific project [21]. Another concern specifies that "to eliminate intentional or unintentional planning mistakes leading to time and cost overrun, current project management theories prescribe the eradication of any bias leading to overly optimistic forecasts. In an effort to hamper optimism bias, normative project management theory and practice introduce further tools and processes to eradicate the causes of optimism bias" [44].

According to [23], risks are defined as uncertainties which, if they occur, would affect achievement of the objectives negatively (threats). Similar, opportunities are uncertainties which, if they occur would affect positively the project. Examples include the possibility that planned productivity targets might not be met, interest or exchange rates might fluctuate, the chance that client expectations may be misunderstood, or whether a contractor might deliver earlier than planned. These uncertainties should be proactively managed through the risk management process. Changes in the business environment, project evolution and percent of completion, will continue to affect the risk situation inside projects. As each project is unique and has its own trajectory, project risk management should be a proactive approach for efficient and effective decision-making.

Projects continue to fail because of a proper project risk management lack and the management problem is getting bigger. A large number of organizations and institutes spend a lot of time and resources to develop and improve standards and methodologies for project risk management. Among these, the most common are:

- AS/NZS 4360:2004, The Australian and New Zealand Standard on risk management, [14] developed by Broadleaf Capital International PTY LTD;
- IRM Standard, jointly developed by The Institute of Risk Management (IRM) [43], The Association of Insurance and Risk Managers (AIRMIC) and by ALARM The National

Forum for Risk Management in the Public Sector (widespread in the United Kingdom) – adopted by the Federation of European Risk Management Associates (FERMA);

- PMBOK Guide developed Project Management Institute [35] (widespread in the U.S.A);
- PRAM (Project Risk Analysis and Management) Guide developed by Association for Project Management[3] (widespread in United Kingdom).

Adapted after PMBOK Guide, developed by Project Management Institute in a Practice Standard for Project Risk Management [35, 36], and by improving planning tools and techniques, a framework for project risk management can be derived, as presented in Figure 5.

A risk identification step where techniques and special tools are used is needed because a risk cannot be managed unless it is first identified. After identification, a risk assessment and evaluation step follows. Each risk is evaluated based on the probability of its occurrence and its impact on project objectives (in qualitative risk analysis) and a numerical estimate of the overall project risk is provided (in quantitative risk analysis). The risk analysis should be performed not only once but in iterative cycles along major. Monitoring is also recommended in continuous intervals. Monitoring is mandatory to permanently track and control identified risks and to promptly respond in case of new risks occurrence that were not previously identified. Risk management does not aim to eliminate risks, but focus to actively manage risks in a business context [39].

Back to risk mitigation step, to certain risks may apply a set of mitigations. Often, mitigations address processes of the organizations. For fruitful results, this process may require long time from implementation. The strategies to manage the risks and the method of mitigation and the priorities are mostly decided by the project manager. Each risk is identified in previous steps and each one is characterized by probability of occurrence and impact. The focus should be on risks that yield a high impact and are identified with high probability of occurrence. Figure 5 presents risks after and before mitigations. Another approach is to address risk after mitigation costs ranking. Deriving a root cause of several risks represents another method to mitigate risks; improvements may be applied there. In this way, the impact of multiple risks that are related to root cause may be decreased.

There are several strategies that can be applied in risk mitigations step. First of all, a risk can be avoided, by not performing tasks that imply risks. Despite this method is not always fruitful from economic perspective, those tasks may be replaced if suitable. A second method aims to reduce the negative effect and impact of the risk in terms of monetary costs, delay or low quality. In some cases the risk can be transferred to another party. Not all risk can be mitigated or make sense to mitigate, in some cases risks should be accepted (in an inconvenient cost/benefit analysis). After successful mitigation, risks are represented based on probability vs. impact. After mitigation, in case the probability decrease the risks are represented on a lower as a parallel shift with y-axis; in the case the impact decrease, risks shift left as Figure 6 shows.

Mitigation actions should target those risks associated with high leverage towards minimizing the residual risk afterwards or acceptable mitigation costs. The risk analysis methodology provides means of visualizing mitigation costs and expected residual risks with and without mitigation. After mitigation actions and responsibilities for the measures are agreed and

[3] http://www.apm.org.uk/

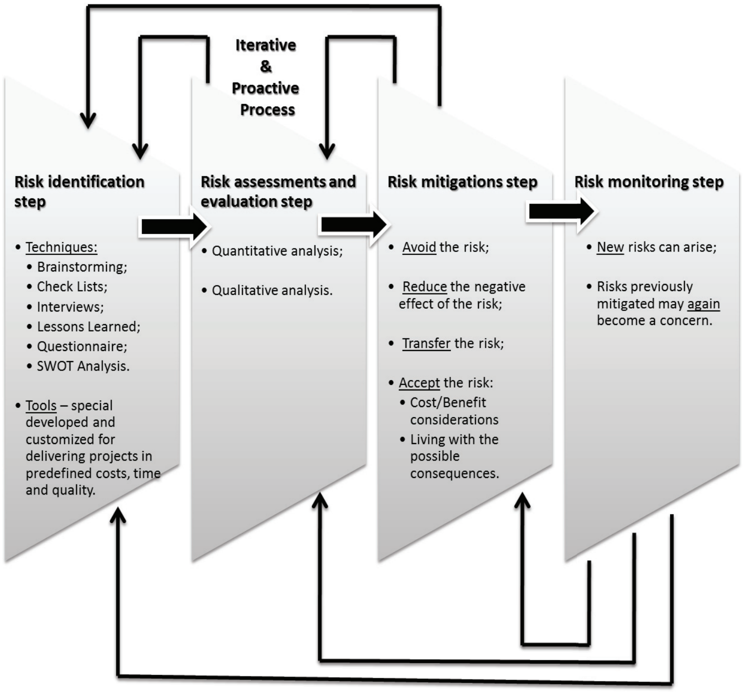


Figure 5. Risk management framework [12]

decided upon, the implementation of these actions need detailed planning and controlling. According to the findings in the risk mitigation phase the responsible action owners should be mandated by the project manager. Hence, the project manager needs to prioritize mitigation options and corresponding costs and efforts to decide whether measures will be taken or not. Performing agreed mitigation actions will contribute to reduce the risk status and create a new risk situation in the project. However, the implementation and controlling of the mitigations will be driven by the project team.

High risks are represented with red, medium risks with yellow and low risks with green in both situations: after and before risk mitigations. Our focus is on high and medium risks.

Despite the enormous sums of money being spent on project planning, risks arise. A risk management oriented culture is preferred to manage risks (cost, time or quality related) in conditions that most of the big projects fall victim to significant cost overrun.

5.2. Opportunity management

In project environment, risks are considered as uncertain events that if occurs has a negative effect on a project objective. Contrasting, the positive and desired effect is considered opportunity. A proactive risk management plan should be applied to successful manage uncertainties. Threats risks and opportunities should be treated as single unit and

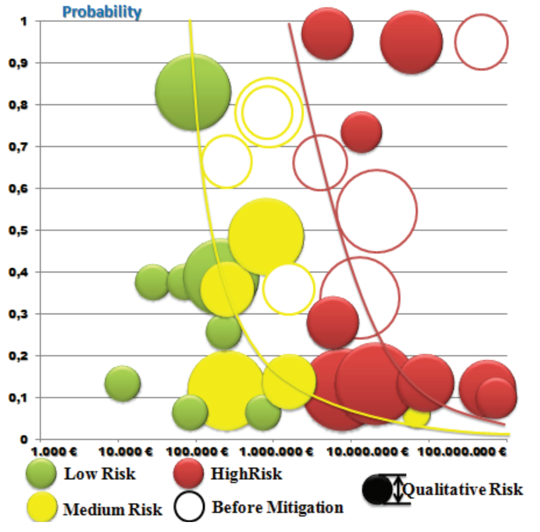


Figure 6. Risk representation Probability vs. Impact (After and Before Mitigations) [12]

individually defined in a single plan. Traditional project risk management approaches offer less attention on project opportunities. For example, check lists are considered a great technique to identify risks and captures previous experiences but most of the time this technique includes only threats and misses opportunities. Another technique is represented by Lessons Learned Reports that prevents making the same mistakes or missing the same opportunities twice. It also offers solutions and ideas where the opportunities can be unlocked. SWOT analysis addresses both threats and opportunities. As risks, opportunities are linked to one project objective (time, cost, quality, scope etc.). Some of the authors consider that companies can overlook opportunities that provide significant possibilities for organizational innovation and new competitive advantage by focusing on the downside of risk [8]. Both, risk and opportunities require attentions and should be treated. By managing risks, threats can be and also an opportunity perspective is available. Such opportunities may provide significant results in terms of innovation and project delivery at high-quality. In many situations, the opportunity is easily missed. Some strategies for identifying opportunities are listed and presented in [8]: learning from the past, customer sensitivity, learning from others, scanning, scenario planning, identifying the market gaps and change the game, idealized design and competing in advance and market sensitivity.

6. Conclusion

The rationality behind collaboration is creativity. Collaboration uses communication, coordination and cooperation as a backbone but it is seen as something more than that. It is about creating shared understanding that no member could achieve on his own and in order to generate insight, new ideas or new artifacts it is a necessity to bring together different and often controversial points of view.

Collaboration can be achieved if defined in a goal oriented framework and should use knowledge from project and risk management to increase the chances or obtaining the envisioned deliverables. These tools can lead to successful achievement of objectives, but they are not sufficient.

During the execution of a project team members are not always collaborating and their work alternates with cooperation, when a greater emphasis is placed on a value-chain model of producing results. Focus permanently switches from the flexible content approach to the management tools according to task's specific. In order to fully support the process of collaboration, the aspects that precede it or come in-between the collaboration sessions must be fully supported so that they will not represent a problem that can hinder collaboration. Both collaboration and cooperation require efficient project and risk management and this can only be achieved if collaboration prerequisites are sufficiently understood and integrated in the overall strategy.

Author details

Marius Ioan Podean and Dan Benta
Babes-Bolyai University of Cluj-Napoca, Romania

7. References

[1] André, P., Schraefel, M. C., Teevan, J. & Dumais, S. T. [2009]. Discovery is never by chance: designing for (un)serendipity, *Proceeding of the seventh ACM conference on Creativity and cognition*, ACM, New York, NY, USA, pp. 305–314.
URL: *http://doi.acm.org/10.1145/1640233.1640279*

[2] Arias, E., Eden, H., Fischer, G., Gorman, A. & Scharff, E. [2000]. Transcending the individual human mind - creating shared understanding through collaborative design, *ACM Trans. Comput.-Hum. Interact.* 7(1): 84–113.

[3] Arikpo, I. I. & Osofisan, A. O. [2010]. Migrating from Traditional Software Development Processes to Agile Software Development: The Role of Organizational Culture, *COMPUTING AND INFORMATION SYSTEMS JOURNAL* 14(2).

[4] Arikpo, I. I. & Osofisan, A. O. [2011]. Your Development Team: A Fulcrum for Successful Migration to Agile, *Computing and Information Systems Journal* 15(1).

[5] Augustine, S., Payne, B., Sencindiver, F. & Woodcock, S. [2005]. Agile project management: steering from the edges, *Communications of the ACM* 48(12): 85–89.

[6] Baskerville, R., Ramesh, B., Levine, L., Pries-Heje, J. & Slaughter, S. [2003]. Is Internet-Speed Software Development Different?, *IEEE Software* 20(6): 70–77.
URL: *http://doi.ieeecomputersociety.org/10.1109/MS.2003.1241369*

[7] Beck, K. [2000]. *Extreme Programming Explained: Embrace Change*, Addison Wesley.

[8] Bekefi, B. T., Epstein, M. J. & Yuthas, K. [2008]. Creating growth from effective opportunity risk management, *CMA Management* 205(February): 72–79.

[9] Benta, D. & Podean, M. I. [2011]. Risk Management approaches for successful projects, *Proceedings of the 2nd Symposium on Business Informatics in central and Eastern Europe*, Osterreichische Computer Gesellschaft, pp. 39–49.

[10] Benta, D., Podean, M. I., Jecan, S. & Mircean, C. [2011]. Simple steps for Risk Management in small and medium projects, *Proceedings of The 7th International Conference on Management of Technological Changes*, Alexandroupolis, Greece.

[11] Benta, D., Podean, M. I. & Mircean, C. [2011]. On best practices for risk management in complex projects, *Informatica Economica. INFOREC Association* .

[12] Benta, D., Podean, M. I., Rusu, L. & Mircean, C. [2011]. From Nobel Prize to Reference Class Forecasting for Successful Projects, *Proceedings of The 7th International Conference on Management of Technological Changes* , Alexandroupolis, Greece.

[13] Benta, D., Rusu, L. & Podean, M. I. [2011]. Successful Implemented theories for Reference Class Forecasting in industrial field, *in* D. Marca, B. Shishkov & M. van Sinderen (eds), *ICE-B 2011 - Proceedings of the International Conference on e-Business, Seville, Spain, July 18-21*, INSTICC Press.

[14] Broadleaf Capital International PTY LTD [2007]. The Australian and New Zealand Standard on risk management (AS/NZS 4360:2004). Last Accessed: April 2012.
URL: *http://www.ucop.edu/riskmgt/erm/documents/asnzs4360_2004_tut_notes.pdf*

[15] Camarinha-Matos, L. M. & Afsarmanesh, H. [2008]. *Collaborative Networks: Reference Modeling*, Springer US, chapter Collaboration forms, pp. 51–66.

[16] Csikszentmihalyi, M. [1997]. *Creativity: Flow and the Psychology of Discovery and Invention*, Harper Perennial, New York, USA.

[17] Csikszentmihalyi, M. [2008]. *Flow: The Psychology of Optimal Experience*, Harper Perennial Modern Classics, New York.

[18] Denise, L. [1999]. Collaboration vs. C-Three (Cooperation, Coordination, and Communication), *Innovating* 7(3). Last Accessed: October 2010.
URL: *http://www.zsr.org/results-framework-resources-2/collaborationvstheothercwords.pdf*

[19] Engelbart, D. C. [1995]. Toward augmenting the human intellect and boosting our collective IQ, *Commun. ACM* 38(8): 30–32.
URL: *http://doi.acm.org/10.1145/208344.208352*

[20] Farooq, U., Carroll, J. M. & Ganoe, C. H. [2005]. Supporting creativity in distributed scientific communities, *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, GROUP â€™05, ACM, New York, NY, USA, pp. 217–226.
URL: *http://doi.acm.org/10.1145/1099203.1099242*

[21] Flyvbjerg, B. [2007]. Eliminating Bias through Reference Class Forecasting and Good Governance, *Technical report*. Concept Report No 17 Chapter 6.

[22] Fuks, H., Raposo, A., Gerosa, M. A., Pimental, M. & Lucena, C. J. P. [2008]. *Encyclopedia of E-collaboration*, Information Science Reference - Imprint of: IGI Publishing, Hershey, PA, chapter The 3C Collaboration Model, pp. 637–644.

[23] Hillson, D. [2007]. When is a risk not a risk?, *Project Manager Today* pp. 15–16.

[24] Hilson, D. [2006]. Integrated risk management as a framework for organisational success, *Proceedings of the PMI Global Congress*, Seattle WA, USA.

[25] Kahneman, D. & Tversky, A. [1979a]. Prospect theory: An analysis of decision under risk, *SIAM Journal on Computing* 47(2): 263–291.

[26] Kahneman, D. & Tversky, A. [1979b]. *Studies in the Management Sciences: Forecasting*, Vol. 12, Amsterdam, North Holland, chapter Intuitive Prediction: Biases and Corrective Procedures.

[27] Lalsing, V., Kishnah, S. & Pudaruth, S. [2012]. People facetor in agile software sevelopment and project management.
URL: *http://airccse.org/journal/ijsea/papers/3112ijsea09.pdf*

[28] Lee, G. & Xia, W. [2010]. Toward Agile: An Integrated Analysis of Quantitative and Qualitative Field Data on Software Development Agility, *MIS Quarterly* 34(1): 87–114.

[29] Moeller, R. [2007]. *COSO Enterprise Risk Management: understanding the new integrated ERM framework* , John Wiley & Sons, Inc., New Jersey, USA.
[30] Osher, D. M. [2002]. Creating Comprehensive and Collaborative Systems, *Journal of Child & Family Studies* 11(1): 91–99.
[31] Podean, M. I., Benta, D. & Costin, R. A. [2011]. On supporting creative interaction in collaborative systems. A content oriented approach, *The First International Workshop on Sustainable Enterprise Software (SES2011) held in conjunction with The 13th IEEE Conference on Commerce and Enterprise Computing (CEC2011), Luxembourg, September 5-6*.
[32] Podean, M. I., Benta, D. & Mircean, C. [2010]. Overlapping Boundaries of the Project Time Management and Project Risk Management, *Informatica Economica* 14(4): 156–163.
[33] Podean, M. I., Benta, D. & Rusu, L. [2011]. About Creativity in Collaborative Systems. Why it matters and how it can be achieved, *in* D. Marca, B. Shishkov & M. van Sinderen (eds), *ICE-B 2011 - Proceedings of the International Conference on e-Business, Seville, Spain, July 18-21*, INSTICC Press, pp. 151–154.
[34] Podean, M. I., Nitchi, S. I. & Benta, D. [2011]. Content Management in the Context of Collaboration. The Third International Conference on Creative Content Technologies (CONTENT 2011), September 25-30, 2011, Rome, Italy.
[35] Project Management Institute [2008]. *A guide to the Project Management Body of Knowledge*, Fourth edn, Project Management Institute, Inc., Atlanta, USA.
[36] Project Management Institute [2009]. *Practice Standard for Project Risk Management*, Project Management Institute, Inc., Pennsylvania, USA.
[37] Schrage, M. [1990]. *Shared Minds: The New Technologies of Collaboration*, Random House.
[38] Shahrbanoo, M., Ali, M. & Mehran, M. [2012]. An Approach for Agile SOA Development using Agile Principals, *International Journal of Computer Science & Information Technology* 4(1).
[39] Spedding, L. S. & Rose, A. [2007]. *Business Risk Management Handbook: A sustainable approach*, CIMA Publishing.
[40] Straus, D. [2002]. *How to Make Collaboration Work: Powerful Ways to Build Consensus, Solve Problems, and Make Decisions*, 107-8, Berrett-Koehler Publishers, San Francisco.
[41] Sutherland, J. [2007]. *The Scrum Papers. Nut, Bolts, and Origins of an Agile Framework*, PatientKeeper, Inc., USA.
[42] The Agile Manifesto [n.d.]. The Manifesto for Agile Software Development. Last Accessed: February 2010.
URL: *http://Agilemanifesto.org*
[43] The Institute of Risk Management [2002]. A Risk Management Standard. Last accessed: April 2012.
URL: *http://www.theirm.org*
[44] Weyer, B. [2011]. Working Papers No. 59 in Perspectives on Optimism within the Context of Project Management: A Call for Multilevel Research, *Technical report*. Working Papers of the Institute of Management Berlin at the Berlin School of Economics and Law.
[45] Wolff, T. [2005]. Collaborative Solutions - True Collaboration as the Most Productive Form of Exchange. Last Accessed: October 2010.
URL: *http://www.tomwolff.com/collaborative-solutions-newsletter-summer-05.htm*

Chapter 7

On the Very Idea of Risk Management: Lessons from the Space Shuttle *Challenger*

Robert Elliott Allinson

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51666

1. Introduction

1.1. The case for risk taking or risky decision making

In this chapter, we will argue that the very concept of risk management must be called into question. The argument will take the form that the use of the phrase 'risk management' operates to cover over the ethical dimensions of what is at the bottom of the problem, namely, risky decision making. Risky decision making takes place whenever and wherever decisions are taken by those whose lives are not immediately threatened by the situation in which the risk to *other* people's lives is created by their decision. The concept of risk management implies that risk is already there, not created by the decision, but lies already inherent in the situation that the decision sets into motion. The risk that exists in the objective situation simply needs to be "managed". By changing the semantics of 'risk management' to 'risk taking' or 'risky decision making', the ethics of responsibility for risking other people's lives will come into focus. The argument of the chapter is that by heightening the ethical sensitivity of decision makers, these decision makers will be less likely to make decisions that will cause harm and/or death to those who are the principal actors in the situation created by the decision.

2. Definition of terms

We should first define our terms. The phrase 'risk management' will refer primarily to the risk of life and limb of human beings, present and future and to the life and health of the planet. Secondarily, it will refer to the taking of risk with the possessions, e.g., the wealth of human beings. For the most part, the discussion will refer to the primary sense of the term and only occasionally in special contexts, to the second. It should be pointed out, at the outset, that the phrase 'risk management' is already biased in favor of risk. The phrase implies that the taking of risk is either necessary and/or advantageous and if it carries any negative consequences, these consequences can be mitigated or eliminated by proper management. This bias toward risk is assumed in the acceptance of the use of the phrase 'risk management'. A major purpose of this investigation is to call into question this built-in bias toward risk.

In the strictest sense, the proper reference for the inquiry should be 'risky decision taking' or 'risky decision making' rather than 'risk management'. Such a refining of the subject of inquiry would make the concept of risk either neutral or questionable since the 'taking' of risk implies negative consequences whilst 'risk management' carries with it the hidden meaning that risk is already being protected against or absorbed by effective management policies. The phrase 'risk management' grants risk a protective coating such that the consequences of the risk are camouflaged. As a result, the following discussion will focus primarily on the entire question of 'taking risk' or making 'risky decisions' in the first place and only secondarily with the 'management of risk'.[1] It is important to make this distinction because the entire question of the ethics of risk must be considered in the first place.[2] *Where, to any degree, whatsoever, 'risk' is already acceptable, the question of the ethics of risk becomes diluted in value. When the taking of risk is itself under question, the ethics of risk takes on greater meaning.*

In the specific case from which we will draw most of our information and discussion, the launching of the U. S. space shuttle, *Challenger*, the action to be taken, was considered to be an 'acceptable risk'. The question of 'acceptable risk' was applied to the action to be taken, not to the decision to take the action. Had the notion of 'acceptable risk' been applied to the decision to launch, some ethical responsibility for the *decision* would have been present. As it turned out, the ethical responsibility for the decision to launch was conspicuous by its very absence.

The ethical issues involved in the decision to launch were further nullified by the choice of terminology utilized to classify the level of risk involved in the malfunction of the part that eventually did malfunction (the O-ring). The label utilized was criticality 1, the definition of the consequence of its occurrence being, loss of mission and life. When one of the four managers who overrode the unanimous decision of 14 engineers and managers not to launch was asked at the Presidential Commission hearings, whether the phrase 'loss of mission and life' had a negative connotation, the answer given by the manager, Larry Mulloy, was that such a description had no negative connotation and simply meant that you have a single point failure with no back-up and the failure of that single system is catastrophic.[3]

[1] For the sake of convenience, in general we will employ the term 'risk taking' as a short-hand for 'risky decision making' which will always stand for 'risking the consequences to the principal risk takers as a result of decision making'. The phrase 'principal risk takers' refers to those whose lives will be affected by the occurrence of the risk, the primary actors who will be directly involved in taking the risk in the risky situation; e.g., in the case of the *Challenger*, the astronauts and civilians aboard rather than the 'decision makers', the four middle managers.

[2] It is gratifying to learn that the U.S. government is currently teaching the terms and definitions for the proper understanding of risk management that the author of this chapter originated that are at the basis of the ideas in this chapter in its educational training courses for FEMA, the Federal Emergency Management Association.

[3] Richard C. Cook, *Challenger Revealed, An Insider's Account of How the Reagan Administration Caused the Greatest Tragedy of the Space Age*, New York: Avalon, 2006, p. 243. Without Richard Cook's early articles in the New York Times, it is possible that the entire *Challenger* investigation would not have occurred. It was a source of inspiration to be in

The reaction of Richard Cook, the budget analyst at the time, shows how the language of choice removes ethical considerations from one's consciousness:

How extraordinary: possible "loss of mission and life" doesn't have a negative connotation.[4]

Cook goes on to say that there was no negative connotation because it had been deemed an acceptable risk and, moreover, there had never been any criteria for defining an "acceptable risk". This was a result of using a failure modes effects analysis without any quantitative risk measures. In other words, the odds of the risk of a catastrophe occurring to the Challenger, were conjured out of thin air, not calculated, because it was a subjective engineering judgment that was not based upon any previous performance data.

Three conclusions present themselves. Firstly, there must exist considerable ethical blindness when loss of life has no negative connotation. Secondly, there must be considerable ethical blindness when loss of life is somehow considered an acceptable risk. Thirdly, there must be considerable epistemological blindness when a notion of "acceptable risk" can be in use without being based on any objective measurements. Cook, one of the few if not the only source on the Challenger who goes into detail on this point, points to one study, conducted in 1984, which concluded that the chance of a Solid Rocket Booster explosion was one in thirty-five launches.[5] Nevertheless, the Marshall managers spoke of "acceptable risk". If you "lost" one astronaut, that was "data" in the risk equation.

In 1977, NASA commissioned a group called the Wiggins Group to study the possibility of flight failures and by examining data for all previous space launches, a likely failure rate of one in fifty-seven was derived. According to Cook, NASA complained that many of the launch vehicles the Wiggins group included were not similar enough to the *Challenger* shuttle to be part of the data base. So, Wiggins changed the probable failure rate to one in 100. Another study conducted by the Air Force placed the likely failure rate for the shuttle in a similar range to the Wiggins analysis. According to Cook, none of these studies were publicized and most of the newspaper reporters who covered NASA most likely had never heard of them. When NASA was forced to arrive at an official number, their chief engineer came up with the infamous and arbitrary estimate of one in 100,000. As Cook put it, 'At a rate of twenty-four launches per year, this meant that NASA expected the shuttle to fail catastrophically only once every 4,167 years.[6] In the language of risk utilized in this chapter introduced below, the possible incidence was therefore negligible. By presenting, not calculating, the *incidence* of risk as virtually non-existent, it was possible to immunize oneself against the realities of the *consequences* of the risk. One could then make a decision to risk other people's lives, because statistical probability had eliminated the problematic dimension of the risk factor, the consequences, from the equation. When a figure is used, such a one in 100,000, one might assume that this is a calculated risk since it is put in the

communication with Richard Cook at the time of my writing the *Challenger* chapters in my book, *Saving Human Lives: Lessons in Management Ethics.*

[4] Ibid.

[5] Ibid., p. 356.

[6] *Ibid*., pp. 126-7.

mathematical language of percentages and statistics. But, this was not a "calculated risk" at all, but fantasy parading as mathematics. A figure picked out of thin air had granted the decision makers moral immunity.

3. Incidence and consequence

An important distinction to keep in mind when considering the whole question of risk taking is the distinction between the likelihood of the occurrence of the risk under question versus the harmful consequence of the actual occurrence of risk under question. Clearly, when the likelihood of occurrence is high and the consequence of occurrence severe in terms of harm doing to risk takers, be it the public at large and/or the environment, this is an example of the kind of risk that should never be taken except under the most severely warranted circumstances. The example of self-defense, when confronted with a murderer who is armed and plans to murder you, comes to mind. Here, the likelihood that your defense will result in his escalating his response is very high, and the consequence of his escalation, in the hypothetical circumstance, could prove to be fatal. Nevertheless, in order to protect your own life, you must take the risk of self-defense under these circumstances.

There is no need to consider the case in which the incidence of occurrence is low and the consequence is also low.[7] Here, the example of starting to speak while inside a house during the winter in Finland with all the windows and doors closed and the sound of a bird chirping that interrupts one's speech may be taken as an example of low possibility of occurrence and low severity of consequence. Of course, one may imagine a scenario in which one's speech was to warn another of a fatal and impending danger, and this warning would be blocked, but any example can be played with to tamper with the point it is designed to illustrate. All we needed to do in this case would be to qualify the original example to state that the speech one was about to utter was an exclamation of how blue the sky was this morning. But, it is important to consider what point an example is designed to make since one can always find some way to find some fault with the example one is choosing to illustrate one's point.

We also need not consider the kind of risk that involves little consequence even if the possibility of its incidence is high. For example, when we carry a glass of milk across the floor, we may easily spill some milk. But, the consequence of the spill (excepting the scenario that we or the person to whom we are carrying the milk are starving) is of no great import. Our discussion need not include examples of high risks of incidence that involve harmless consequences.

The kind of risk with which we most frequently struggle is the risk in which the incidence of its occurrence is low, but the consequence of its occurrence is grave. The example of an airplane exploding in mid-air is a good example of this type of risk. We assume in the state of technology that currently exists commercial air travel is an advantage that we do not wish

---

[7] For the sake of convenience, whenever the term 'consequence' is used it is understood that what is meant is 'harmful consequence' to primary risk taker, general public, the environment, future generations or all of the above.

to surrender. We also know that in the general, unknown risk category, that a plane may explode in mid-air. This risk taking is minimized by careful and regular inspection of the mechanical parts of the airplane and a replacement of said parts and said plane on a needful basis. Other aspects of this risk are minimized by guarding against a drunken pilot, hijacking by suicidal terrorists, etc. In such cases, risk is minimized. It is more accurate to consider that the risk in these cases is *minimized* rather than *managed*, because its possibility of occurrence is reduced rather than the occurrence of its risk being managed. The latter understanding is how the term 'risk management' might well be construed. In fact, it is difficult to understand what the term 'management' means in the case of 'risk management'.

4. The ethics of Risk

We assume, as an ethical premise, that there should never be a risk taken to potential life and limb unless it is absolutely necessary. A good example of this would be the Hippocratic Oath taken by physicians which begins with the axiom: '*Primum non Nocere*', 'Do no Harm'. One takes risk, as with surgery, only when it is necessary to promote or safeguard health. In other words, risk is justifiable only when it is absolutely necessary *in the service of life*.

What about cases of advantage rather than absolute necessity? For example, let us again consider the case of commercial airplane travel. There is certainly risk involved and it would seem to be the case that the concept of risk management would come into play. Upon closer examination, however, when one considers the safety precautions that are taken through mechanical inspection, etc., one realizes that it is 'risk taking' that is modified, that is, one is reducing the risk involved, rather than "managing" an existing risk. *One minimizes the risk involved: one is not managing risk; one is minimizing risk.*

5. General unknown risk versus specific and foreknown risk

There is a confusion that is frequently made between general and unknown risk that is operative in the universe and any specific risk that is known in advance to exist. For example, whenever one gets out of one's bed in the morning, one may trip, fall, crack one's skull and have a concussion. This is the general and unknown risk that is operative in the universe. We should not construe risk in these terms as this kind of risk exists, for the most part, outside of human control and intervention.

Risks that are foreknown in advance to the principals involved in the risk-taking are the only kind of risks that our discussion should consider. A classic case in point of the contrast that exists between general, unknown risk and specifically foreknown risk is the case of the faulty and dangerous O-rings that were known in advance to exist (by managers and engineers though not by the principal risk takers) prior to the fatal flight of the space shuttle *Challenger.* The classic case of the *Challenger* disaster can be used to illustrate the fallacies of the concept of 'risk management' and the need to replace this concept with the more accurate, new concepts of 'risk taking' or 'risky decision making'. While other cases could also be chosen, the availability of overwhelming, documented evidence in the case of the

Challenger disaster makes it an ideal case study for the purposes of examining the concept of risk management.

One can argue, fallaciously, that whenever an astronaut goes into space, that astronaut is subject to that general, unknown, universal-style risk of space travel. This, however, is not comparable to an astronaut going into space equipped with the full knowledge of the existence of a real, specific and pre-existing mechanical fault that could be potentially fatal to her or his space craft. It is only with the latter kind of known risk, known fully to those taking the risk, that a discussion on risk management should focus. Such was not the case with the astronauts and civilian passengers of the U.S. space shuttle, *Challenger*.

In the case of the *Challenger* launch, the overwhelming evidence has revealed that the astronauts were completely unaware of the specific dangers that the O-rings posed. According the Malcolm McConnell, the science reporter, no one in the astronaut corps had been informed of any problem with the SRB field joints.[8] According to Charles Harris, Michael Pritchard and Michael Rabins, authors of *Engineering Risks, Concepts and Cases*, '… no one presented them [the astronauts] with the information about the O-ring behavior at low temperatures. Therefore, they did not give their consent to launch despite the O-ring risk, since they were unaware of the risk.'

Claus Jensen writes in *No Downlink, A Dramatic Narrative about the Challenger Accident and Our Time*, that when the Rogers commission summoned a group of space shuttle astronauts, 'During this session, the astronauts reiterated that they had never been told about the problems with the solid rocket booster.' And, in a private correspondence with the present author and Roger Boisjoly, the late senior scientist and the engineer who knew the most about the O-rings, Boisjoly wrote, 'I KNOW for a FACT that the astronauts on Challenger did NOT KNOW about the problem with the O-rings at temperatures below 50 degrees F.' (emphasis his)[9] According to Richard Lewis' book, *Challenger, The Final Voyage*, 'Along with the general public, the astronauts who were flying the shuttle were unaware of the escalating danger of joint seal failure. So were the congressional committees charged with overseeing the shuttle program. NASA never told them that the shuttle had a problem.' Later, in the same work, Lewis pointedly quotes from the Presidential Commission report:

> *Chairman Rogers raised the question of whether any astronaut office representative was aware [of the O-ring problem]' Weitz, [an astronaut's representative] answered: "We were not aware of any concern with the O-rings, let alone the effect of weather on the O-rings.*[10]

Despite the very clear declarations above, nowhere in the 575 pages of Diane Vaughan's book, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA*, is it ever mentioned that the astronauts and the civilians were not informed of the O-ring dangers.[11] By not ever mentioning this crucial point of information, her book leaves one with the impression either" (i) that the astronauts knew about the risk they were taking and

---

[8] Robert Elliott Allinson, *Saving Human Lives, Lessons in Management Ethics*, Dordrecht: Springer, 2005, p. 156.
[9] *Ibid*., p. 184.
[10] *Ibid*., pp. 184-5, 187.
[11] *Ibid*., p. 187.

thereby had given their informed consent, or, (ii) that their knowledge and thereby consent to take such a risk, was completely irrelevant.

All of the above clearly points to the fact that the objection that one always takes a general, unknown existential risk in life and is aware of this fact is irrelevant whenever there is a specifically foreknown risk that carries with it great and unfortunate consequence. These points make a further argument to support the case that the very notion of risk management needs to be replaced with the idea of risk taking when a known risk carries with it a high probability of life and death danger and there is nothing crucial to be gained by taking such a risk, such a risk is never "managed", such a risk is never *taken*. In such a case, any possibility of risk is *eliminated* by not being taken in the first place. *Risk management becomes risky management.* We can now turn to the examination of risky management.

6. Risky technology versus risky management

In the classic case of the space shuttle *Challenger* example, two years after the horrific event, when the official, U. S. government committee on Shuttle Criticality, Review and Hazard Analysis examined the risk that had been taken in launching the *Challenger*, the Chairman of that committee wrote in the very first paragraph of chapter four of their report, entitled 'Risk Assessment and Risk Management':

> *Almost lost in the strong public reaction to the Challenger failure was the inescapable fact that major advances in mankind's capability to explore and operate in space – indeed, even in routine, atmospheric flight – will only be accomplished in the face of risk.*

And, later, in the body of that same report, the Committee wrote: 'The risks of space flight must be accepted by those who are asked to participate in each flight …[12]

It is rather easy to spot the fallacy that is being made in this case. It is the fallacy of equating a general unknown risk (space flight in general) with a specifically foreknown fatal risk (the flawed design of the O-rings about which the senior engineer Roger Boisjoly had issued red flagged warnings).[13] If the committee were pressed on this matter and replied that they were aware that there was some risk in the use of the technology employed at the time (the hazardous O-rings), in the case of the *Challenger* disaster, what the Committee on Shuttle Criticality, Review and Hazard Analysis would then seem to be saying is that it was justifiable to employ the 'risky technology' that was employed at the time because of the understanding that in order to make progress in space exploration that one was required to take chances with very risky technology. Is this actually the case? Was it necessary to take cavalier chances with risky technology in order to make progress in the arena of space exploration? The implication is that the risk would have to be taken and the "management" of the risk would seem to amount to something on the order of crossing one's fingers, and hoping that nothing would happen.

---

[12] Preface by Alton Slay, Chairman, Committee on Shuttle Criticality, Review and Hazard Analysis, *Post-Challenger Evaluation of Space Shuttle Risk Assessment and Management*, Washington, D.C.: National Academy Press, 1988, p. v; p. 33.

[13] Robert Elliott Allinson, *Saving Human Lives, Lessons in Management Ethics*, Dordrecht: Springer, 2005, p. 138.

When one examines the *Challenger* case more closely, one discovers that there was no need to choose the technology that was chosen in the first place. The risky technology in this case in point was the engineering design of the O-rings. Of four designs submitted, the one chosen was the least safe (and the cheapest). This choice was a case of risky judgment on the part of the managers who chose this design that was an initiating cause of the *Challenger* disaster, and not the fact that they were obliged to employ a risky design in order to venture into space. The design itself was risky. Why? This design was chosen because financial cost factors, read profit, were taken as a priority over safety. This decision to place cost ahead of safety is an example of ethical misjudgment. Thus, we may consider the decision to choose this unsafe design a case of management malpractice. There was no necessity to choose this particular design. Alternative designs were available.

The O-ring design of giant rubber gaskets keeping combustible hot gases from leaking out, in actual fact, ranked fourth out of four submitted engineering designs and, according to an important article co-written by Trudy Bell, Senior Editor of the engineering journal, *IEEE Spectrum* and Karl Esch: the selection of this design was the chief cause of the *Challenger* disaster.[14] For the next space flight, this design was replaced with the safest design, demonstrating that the safest design could have been chosen in the first place and it was economically feasible to have done so.

There was also no need to choose a design for the space shuttle that did not include an abort system, that is explosive bolt hatches, space pressure suits and a parachute descent system. That it did not include such a safety system was a matter of policy, not impossibility since earlier spacecraft had been equipped with launch escape systems. Again, nowhere, in any place, in the 575 pages of Ms. Vaughan's, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA*, is it ever mentioned that there was no necessity to omit an escape system for the crew and passengers. Omission of this detail can readily be regarded as ethically unpardonable. It implies that there was no way to save the lives of the crew and passengers when in fact *they were indeed alive when the space craft broke apart* (the space shuttle did not explode, as was popularly reported – the spectacular image of the smoke was due to the chemical reaction of hydrogen colliding with oxygen). The astronauts were conscious as they hit the ocean floor at the tremendous impact (some of the crew had actually activated and used their emergency air packs) that caused their death. By omitting this incredibly important fact in her volume, Vaughan leaves one with the impression that the life and death risk that was taken was one that could not possibly have been prevented![15] That it could, indeed, have been prevented places an entirely different perspective on the kind of unnecessary, and therefore, incredibly unwarranted risk that the astronauts were forced to take. *The Challenger space shuttle astronauts never needed to take a risk with their lives. The risk did not exist: it was created by risky and unethical management decisions.* The false impression created by interpreting the plume of smoke to mean that the space shuttle *Challenger* had exploded blinds one to the fact that the crew compartment had separated from the orbiter

---

[14] Robert Elliott Allinson, 'Risk Management: demythologizing its belief foundations,' *International Journal of Risk Assessment and Management*, Volume 7, No. 3, 2007, p. 302.

[15] *Op. cit.*, pp. 188-189.

and thus the lives of all passengers could have been saved with a parachute descent system. There was no need to take a risk with their lives! The continuing belief that the shuttle exploded continues to play its role in veiling the real issue of the ethics of risky decision making.

The above examples make it clear that the problem of risk lies in the *choices* to be made by the risk takers, or, more precisely, the risk decision makers, not in the *management* of risk already taken. By focusing on the term 'management', one takes for granted that a risk must exist in the first place and needs only to be managed. It is not even clear in this case how the risk is managed unless crossing one's fingers counts as management. In the case of the O-rings, there was no need for this risk to exist in the first place.

One can look into the matter in more detail. Suffice it to say for the present discussion and analysis that there were two levels of risk taking that were matters of choice, not management. The first level was the choice of design. The choice of design could have been altered to have prevented the breaking apart of the space shuttle. The choice of the design could have been altered to prevent the death of the crew and passengers. It was design *choices* that determined the *cause* of the disaster and the *fatal consequences* of the disaster. The second was the choice to fly under weather conditions that heightened the risk involved. These were both management choices or decisions. The focus should be on risky management in the sense that there was risky choice and risky decision-making, not on the technology involved, because the risky technology only came into being in the first place because of these two very risky management choices: the choice of technology and the choice of launch timing.

There has been much erroneous discussion of the technical issues of the temperature and the O-rings and as a result, massive misconceptions have created complicated layers of confusion around the issue.[16] Diane Vaughan's book, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA*, has been responsible for generating much of this confusion. Note the first phrase in her sub-title: 'Risky Technology'. The clear implication is that the fault lay not in the decision to employ such technology but in the problem of having to rely upon risky technology.

According to Roger Boisjoly, the senior scientist at NASA and the one who knew the most about the O-rings, she routinely mixed up data relating to field joints and nozzle joints and did not have a clue about the difference between the two joints.[17] Diane Vaughan is dismissive of Professor Feynman's famous gesture of dipping a piece of an O-ring into a

---

[16] The mass of confusion surrounding the technical issues is, to the best of this author's knowledge, discussed in proper factual detail in the two chapters on the *Challenger* disaster in the author's *Saving Human Lives, Lessons in Management Ethics* (now in Kindle and Google Books). The author is indebted to extensive private correspondence with Roger Boisjoly for clearing up the confusion about the technical issues that made understanding the *Challenger* issues murky and seemingly resistant to plain understanding. To the best of the author's knowledge, such clear explanations that Boisjoly gives, sorting out the mistakes in previous analyses of the *Challenger* disaster such as in Diane Vaughan's confused studies on the issue only exist in print in the author's *Saving Human Lives*. *Cf.*, especially Chapter Seven, 'The Space Shuttle Challenger Disaster,' and Chapter Eight, 'Post-Challenger Investigations', pp. 107-197.

[17] *Saving Human Lives*, pp.192-3 *et passim*.

glass of ice water during the televised Commission hearings and of his astonishment over the concept of 'acceptable risk'. In contrast, Boisjoly argued that Feynman's scientific experiment proved exactly what Feynman was trying to prove: rubber gets stiff at freezing temperatures. In Boisjoly's words, 'As the temperature decreases the sealing performance of the O-ring gets worse. At freezing temperature or below, it will get much worse. IT'S REALLY THAT SIMPLE.'[18] Whose opinion should we accept, the opinion of a Nobel laureate physicist and the senior scientist who knew the most about the O-rings, or a sociologist with a Master's Degree who is evaluating the validity of a scientific experiment carried out by a Nobel laureate physicist the validity of which is confirmed by a rocket scientist?[19]

We can point to three elementary characteristics of warning signals to illustrate that the risk was foreknown and knowledge of it was transmitted: The three characteristics of warning signals. To focus on risky technology is to put the cart before the horse. Risky choice is the horse: it pulls the technology along behind it. Without the horse, the cart would not move. Technology has no power on its own. It is the servant of decision making 10 14-15.

7. That the risk was foreknown and knowledge of it was transmitted

In the case of the *Challenger* disaster, nothing should blind us to the point that the most senior engineer involved was keenly aware of the fatal risk involved and sent red-flagged warnings to this effect. That these warnings were not heeded is sometimes obscured with the "argument" that one cannot hold up actions to be taken on the basis of warnings since every possible action will always have risks and it is next to impossible to take note of every single warning that comes across one's desk. The existence of warnings that cannot supposedly be noticed is referred to under the hypothesis of "weak signals".[20] The hypothesis of "weak signals" is offered up as a rationale why warnings cannot always be well noted.

This "weak signals" hypothesis is easily refuted when one considers the *source*, the *content* and the *form* of the signals. The *source* of the signal in this case is the most senior scientist and the one who knows the most about the O-rings. It is not a crank call made by a tourist to the space center at the information desk. The source in this case was the project's senior engineer himself. It was a warning from the inside, by an insider, who knew the most about the technology about which he was issuing the warning.

The *content of the signal* warns of the danger being fatal to all aboard. There is no weakness in terms of the content of the message. In Boisjoly's famous memorandum of July 31, 1985, he warned before the fact that if there were a launch, 'The result would be a catastrophe of

[18] *Ibid*., p. 192. (emphasis his)

[19] *Ibid*., pp. 172-3. The author of this chapter was gratified when he personally received a hand-written letter written to him by the famous sociologist David Riesman from Harvard, author of *The Lonely Crowd*, who wrote to commend the merits of this author's first review of Diane Vaughan's book, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA*, which was to come out in *Society* . Later, he was prompted by this author's second review, which took a very different tack, and which appeared in *Business Ethics Quarterly* to read her book after which he wrote to this author to say that Diane Vaughan's book was simply, in his words, 'a bad book'.

[20] The hypothesis of "weak signals" to describe the warnings of the failure of the O-rings was put forward by Diane Vaughan in her book, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA*.

the highest order – loss of human life.' In his earlier warning of July 22, 1985, he warned of a horrifying flight failure unless a solution were implemented to prevent O-ring erosion.[21] One can readily see that there is no mincing of words to minimize the possibility that the danger might be understood to be less than absolutely extraordinary. The consequences in terms of life and death danger are spelled out in detail. The specific risk factor is named. One could not possibly ask for a stronger signal.

Finally, the *form* of the message is red-flagged to indicate the most serious of attention must be given to this message. There is no rational way that this warning can be construed as a "weak" signal. It is the strongest possible signal in three terms: its source, its content, and its form. Since the warning is multiple and not single, the form of multiple warnings should further make the case that this is not a single and therefore "weak" signal that might not be noticed. In the case of Boisjoly's warnings, there were two such red-flagged warnings.

It should be noted that this signal was not a one-time occurrence. *This signal was continuously sounded for eight years.* It was not a one-time message that could conceivably have been missed. There is a memorandum written in 1978 by John Q. Miller, Chief of the Solid Rocket Motor Branch to the Project Manager, George Hardy, when referring to the Thiokol field joint design that was chosen, in which he writes that this design was so hazardous that it could produce ' … hot leaks and resulting catastrophic failure.'[22] *It was known eight years prior to the Challenger launch that this design choice was a dangerously, faulty one which could end in a catastrophe!*

In terms of, in terms of the form of the "weak signals" hypothesis, it should be noted that when it came to the timing of the launch, 14 managers and engineers alike *unanimously* voted against it.[23] One could not conceive of a stronger signal than this. That this decision was overturned in a meeting of 4 managers (with no engineers present who were not managers) does not take away from the fact that these 4 managers were fully aware of the previous signal of 14 unanimous votes against the launch.[24] There is no possibility that the signals that were made were not the strongest possible signals to be made. To refer to the warnings as "weak signals" is to turn a red light into a green light.

8. Two straw men

It should be emphasized that the Committee on Shuttle Criticality, Review and Hazard Analysis did not focus on the life and death consequences to the principals involved in taking the risk of launching the Challenger, but rather focused on the general, abstract case of space flight as an opportunity for learning about the universe. In short, they examined a "straw man" and not the real case in front of them to examine.

[21] *Op. cit.*, p. 170.
[22] *Ibid.*, p. 151.
[23] *Ibid.*, pp. 174, 195.
[24] *Ibid.*, for an extended analysis of the unsound and unethical decision making process engaged in by the four middle managers.

In the case of the red-flagged warnings, regarding the launching of the *Challenger*, the consequences were not the consequences marked out by the Committee on Shuttle Criticality, Review and Hazard Analysis. The consequences marked by that Committee were outlined in terms of the risk that was to be accepted in terms of general space exploration. The implication is that those who accepted such a risk were space explorers and that they were fully aware of the risk that they were taking. Neither of these implications is accurate.

The first straw man was to examine the case as if all aboard were astronauts. This assertion was contrary to fact. Not every person aboard the shuttle was an astronaut. In addition to the five astronauts, one of whom was a female, Judith Resnik, there were two passengers: one, a junior and high school teacher, a 37 year old mother of two, Christa McAuliffe and the other, Greg Jarvis, an engineer from Hughes Aircraft (not a member of the Astronaut Corp) who had been given a ride in a space flight as a prize for winning a company competition. Mrs. McAuliffe was scheduled to deliver a nationwide "lesson from space" called the "Ultimate Field Trip" to the nation's school children. She was also supposed to receive a telephone call from President Reagan in mid-flight.

Part of the consequences, then, was the risk of the death of two civilians, who were there, not to operate the space craft, but one to act as a Teacher in Space and the other to claim his contest prize. Both civilians were given the deceptive, camouflage designation of "payload specialists" that implied some kind of crew responsibilities that did not exist. For a cover, Mr. Jarvis was given the task of conducting a fluid dynamics experiment and Mrs. McAuliffe was to videotape six science demonstrations of the effects of weightlessness on gravity, etc.[25] The Teacher-in-Space mission portion was to be featured in President Reagan's State of the Union address the evening of Tuesday, January 28, though the White House later denied it.

The second straw man argument is to assume that knowing the risk of space flight in general was equivalent to knowing the risk of a launch under weather conditions that were known to be unsafe with the O-ring technology in use. Again, neither the astronauts nor the two civilians aboard were made aware of the risk that they were taking. They may have been aware of the existential, general unknown risk of space exploration, but they were not aware of the specific, needless risk they were actually committing to take by being launched into space with a known, faulty technology.[26] Indeed, would not the inclusion of the schoolteacher Christa McAuliffe, who had been given the understanding that the launch was safe, create the vivid impression to all aboard, and to the public at large around the world that this was a very safe flight? The genuine risk being taken was perceived to be minimal around the world by the inclusion of non-necessary personnel on the shuttle.

Was Mrs. Christa McAuliffe made aware of the risk that she was taking? Grace Corrigan wrote:

---

[25] Richard C. Cook, *Challenger Revealed, An Insider's Account of How the Reagan Administration Caused the Greatest Tragedy of the Space Age*, New York: Avalon, 2006, pp. 177-8.
[26] *Op. cit*.., p. 156.

With respect to the Challenger launch, ' … Christa felt no anxiety about the flight. 'I don't see it as a dangerous thing to do.' She said, pausing for a moment. 'Well, I suppose it is, with all those rockets and fuel tanks. But if I saw it as a big risk, I'd feel differently.'[27]

It was not only the case that Christa felt no pre-flight fears. She was never even informed that there was any real problem about which she should have been informed. Corrigan relates Christa McAuliffe's account of what the President and the pilot from a previous launch told her in the White House:

> *'They were told about the dangers of the space program. She said that one could be intimidated thinking of all that he had said until you realize that NASA employed the most sophisticated safety features, and they would never take any chances with their equipment, much less an astronaut's life.'*

When interviewed by *Space News Roundup*, she said that, 'When the Challenger had the problem back in the summer with the heat sensors on the engine … and … one of Boston's papers called me and asked me what I thought was wrong, … I said, 'I have no idea. What has NASA said?'[28]

It is obvious that Christa McAuliffe was not informed of any O-ring faults and the consequent life and death risk she was engaged and committed to taking by her participation. Is this a case where the statement from the Committee on Shuttle Criticality, Review and Analysis, that 'The risks of space flight must be accepted by those who are asked to participate in each flight' is even relevant when no one has been informed about the specific risks to which this particular flight will be prone? The fallacy of conflating the general, unknown risks of space flight with the specifically known risks of this flight makes "risk management" here into an unethical practice. There is, strictly speaking, no risk management that is being practiced. There is simply a wanton *risk taking* with human life.

9. Subjective judgment versus performance data

The Nobel laureate physicist, Richard Feynman was shocked to learn that NASA management had claimed that the risk factor of a launch crash was 1 in 100,000 which they had arrived at through a subjective engineering judgment without relying upon any actual past performance data. If one calculated risk based upon actual past performance data, the risk was, according to Professor Feynman, 1 in 100.[29] While management, in defending its decision to launch, pointed to the risk involved as being 1 in 100,000, there was no examination of how these figures were generated. If one took the actual performance data of rocket engines in the past, as the Nobel laureate physicist Roger Feynman did, the risk was far greater. When one does this, one can more clearly consider the case of the possibility of incidence versus the actuality of consequence. Does one wish to risk the lives of the astronauts and the civilians when the chance of their death is 1 in 100?

---

[27] Grace George Corrigan, *A Journal for Christa, Christa McAuliffe*, Teacher in Space, Lincoln and London: University of Nebraska Press, 1993, pp. 115, 118.

[28] *Op. cit*., p. 185.

[29] *Ibid*., p. 183.

From the above example, one can generate the conclusion that whenever possible, when calculating risk, one should not calculate risk in any other way than from the conclusions generated from actual, real-life, past performance data. The lesson to be learned is that in risk assessment, past performance data, when available, must always be consulted. One should avoid guess work. Unless the risk estimates are based on past performance data as a data base, according to Professor Feynman, 'it's all tomfoolery'.[30] There is no reason why we should not learn from the lessons of the *Challenger* disaster to generalize sound conclusions concerning the method we should employ when engaging in risk assessment.

10. Safety margins

Major General Kutnya, Director of Space Systems and Command Control, USAF, and Presidential Commission member, argued that the O-ring evidence was analogous to evidence that an airliner wing was about to fall off. Professor Feynman pointed out with respect to Diane Vaughan's contention that there was a 'safety factor of three', that because in previous cases, the O-ring had burned only one third of the way through, that did not prove that there was a safety factor of three. If we merge the O-ring and the airplane wing examples, the argument that General Kutnya, an Air Force General and Professor Feynman, a Professor of Physics, give is that if the wings of an aircraft have burned one-third of the way through, that did not mean that they had a two-thirds safety margin as Diane Vaughan, a sociologist with a Masters Degree in Sociology, thinks. If a part that is designed to hold back inflammatory gases is weakened by one-third, then its capacity to hold those gases back is diminished by one-third. In such a weakened state, the margin between its holding up and its caving in to the pressure of the gases is seriously undermined. It is not that it possesses a two-thirds safety margin; it is that one-third of its capacity is diminished. It may not be capable of standing up to a heavy load. Its safety margin at that point may be zero.

In Professor Feynman's words:

> *If a bridge is designed to withstand a certain load … it may be designed for the materials used to actually stand up under three times the load … But if the expected load comes on to the new bridge and a crack appears in a beam, this is a failure of the design. The O-rings of the solid rocket boosters were not designed to erode. Erosion was a clue that something was wrong. Erosion was not something from which safety could be inferred.'*[31]

If we are to generalize from these arguments to future scenarios of risk assessment, we must be careful never to consider problems that develop as evidence that the design is still basically sound. Problems are danger signals, not signals that everything is fine. When safety is compromised, it does not signify that there is still a viable margin of safety. When safety is weakened, what we have left is not a state of safety conditions which are a little less than perfect; we have conditions which are not safe.

[30] *Ibid*., p. 183.
[31] *Ibid*., p. 183.

11. Safety back-ups

One must also be very careful when one considers safety back-ups. In engineering language, safety back-ups are referred to as redundancies. (It is important to take note of the difference in English language usage in the case of technical engineering terms and ordinary English language usage since, in the latter case, a redundancy is that which is unnecessary!) In the case of the *Challenger*, the back-up system was a secondary O-ring seal. The secondary seal was known also to be prone to failure. In effect, there was no secondary seal. Of course, if we are to consider the argument made above carefully, if the primary seal is unsafe, and the secondary seal is made from the same materials with the same design, how is it any safer? When examining any safety back-ups, one must be certain that the back-ups are not of the same faulty design as the technology that they are supposedly "backing-up". Since the O-ring was considered to be at Criticality 1 (no back-up), it is thus not suprising that the secondary seal was not considered to be a back-up by this designation.

In addition, it must not be forgotten that initially all 14 engineers and managers unanimously voted against launching the *Challenger*. Such a vote of no confidence would be proof enough that all 14 engineers and managers had no confidence in the secondary seal. If, when Professor Feynman in his famous, improvised experiment during the televised hearings dropped a piece of the rubber O-ring into a glass of ice water obtained from a waiter, and demonstrated that it had no resiliency left to it at a freezing temperature and therefore could not expand to contain superhot inflammable gases, how would a second piece of the same rubber material be of any use? If one piece of rubber would not seal, why would another piece of the same rubber not also be stiff?[32] There is no safety back-up if the materials of the back-up suffer from the same defect as the materials of the primary material they are supposedly backing-up.

12. The right to make decisions over others' life and death

As a final point in the question of risk taking, we must consider what gives any person or group of people the right to make decisions which will have life and death consequences for those who partake in the actions that will take place as a result of that decision-making. It has been argued that astronauts already knew the risks of space travel when they undertake such an adventurous role. Again, this is to confuse the general unknown risk of space travel with the specifically foreknown risk of launching with an unsafe part. It is also to ignore the fact that two civilians were aboard in addition to the five astronauts.

What if we consider the question of decision making over life and death when the one involved is oneself? Suppose, for example, the astronauts and the civilians had been told of the dangers. It is entirely possible that the astronauts would still have chosen to launch. Even if the decision were placed in the hands of the Captain of the astronauts on that launch and he was fully informed of the dangers, should he have the right to decide if the launch should proceed? In a war, a general does decide if troops should engage in battle. There, of

[32] *Ibid.*, p. 171.

course, there may be a set of circumstances when something very valuable (the lives of one's countrymen for example) must be protected. But, here, we are still considering the scenario where there is nothing great to be gained – such as protecting one's countrymen from destruction - by proceeding with the timing of this launch..

What if we leave the decision to launch to the astronauts themselves and make sure that they are fully informed of the dangers? (For the sake of this discussion, we except the case that there are two civilian passengers aboard). The crew, who perceive themselves as heroes, are likely to decide to vote to launch. When the author had the privilege to personally interview Kathyrn Cordell Thornton, the celebrated astronaut who was part of the 100 strong astronaut corps at the time of the *Challenger*, and then Director of the Center for Science, Mathematics and Engineering Education at the University of Virginia, and the author of this chapter was privileged to be accompanied in this interview by the distinguished business ethicist, Patricia Werhane, we discussed this very issue. The author of this chapter vividly recalls Thornton saying that, 'if these astronauts had refused to go up, there would be 100 others behind them waiting to take their place.' This is not surprising when one considers the peer expectations of that kind of group. There is a psychological expectation that they should be fearless. This is no different from football players going into the field with a head injury. The question now has become, does one have a right to take a risk with one's own life? One could say that this is a matter of individual choice. But, is it?

One can consider the choice of a circus performer who decides to risk her or his life by performing on a high wire with no safety net. Should this be a matter of individual choice? Suppose the circus performer has overestimated her or his abilities? Again, the possibility of the incidence may be low, but the consequences are fatal. The performer is placing a greater value on the spectacle value of a performance without safety than on the value of her or his life. Should such a decision be left in the hands of an individual performer? Or, should such a decision be overridden, if necessary, by the director of the circus?

Do we ever have a right to make life and death decisions that affect our own life and death? Or, is the value of life so precious that it should never be risked unless there is an absolute necessity? If we take the latter position, no one has the right to take such a risk with one's own life except under the conditions of absolute necessity; e.g., as in the case of self-defense. *The point here is that the ethics of the preciousness of life is valued over the concept of absolute choice.*

The case becomes more clear when we consider the circumstance of intended suicide. Assuming a situation in which the intended is not the victim of an incurable disease and is not making a forced choice (such as when a Nazi may say to a Jew, 'Kill yourself now or I will kill your child'), does anyone have a right to gratuitously take her or his own life? While it may be thought that this is still a matter of individual choice and one can point to the custom of ancient Romans and Japanese in this regard, one must remember that in the case of ancient Rome and Japanese tradition, it was a case of personal honor and not merely gratuitous. In the case of the ancient Romans and the Japanese, these honorable suicides were not sanctioned because the individuals involved were desperately unhappy with their lives.

In the case of the astronauts – leaving aside the civilians, the secondary school teacher and the journalist who have no such honor code – *they do not lose their honor if they refuse to board* *an unsafe vehicle*. However, being young, adventurous and headstrong, they may not be able to understand that the risk that they are taking is an unnecessary one. *It should not be their decision. It should be the decision of the engineers.* The managers, in this situation, have the responsibility to follow the opinions of their experts. That some scholar/commentators have argued some of these experts concurred that it was safe to launch or they questioned the senior engineer's warnings about the safety to launch is based upon uninformed, unsupported or prejudicial evidence. This issue is discussed in detail in the author's *Saving Human Lives: Lessons in Management Ethics*. The prejudicial evidence is examined in Robert Elliott Allinson, *Saving Human Lives: Lessons in Management Ethics*.

Ultimately, the deciders, those who are taking the risk of other people's lives in their own hands, have the responsibility of *not risking other people's lives.* It is as simple as this. *There is no such thing as an ethical choice to risk another person's life, including one's own, when it is not necessary.* If managers do so on account of cost-saving, they are making unethical choices if we define the basic principle of ethics to value human life as the highest priority.

We should not leave the discussion of this one classical example without reiterating that the astronauts and the civilians aboard could have easily been provided with proper space suits, parachutes and ejection chairs. No one died as a result of the explosion. Horrifically, all died because of too fast a collision with the ocean. All were breathing until ocean contact. *It is important to emphasize that the crew and the passengers were alive, were conscious after the spectacular explosion.* As pointed out above, during their three minute descent, some crew members had actually activated and used their emergency air packs.[33] That such a provision of space suits, parachutes and ejection chairs were considered and then rejected by management is another risky decision that resulted not in the safe abortion of the mission, but in the deaths of every person aboard. Earlier spacecraft had been equipped with launch escape systems, thus proving that escape systems were not only possible, but were actual. The decision not to equip the Challenger with an abort system was not a decision based on possibility, but a decision based on policy.[34] This death, or should it be said, manslaughter of the astronauts and passengers, was not the result of high-risk technology, but the result of a cost-benefit analysis that took into account the benefit of profit, not the benefit of life. This horrific outcome was the result of risky and unethical decision making, not the result of improper risk management. This risk did not have to exist in the first place. It was a management decision that decided not to include these safety precautions: *it was risky decision making*, not a lack of risk management. (All later missions were equipped with such life-saving devices).

13. Risking funds

Classic examples of risk management would appear to be the actions of fund managers. Under the arguments advanced above, if one likens one's funds as the means to one's livelihood, one should never take risk with other people's money. While this would appear to be impossible without doing away with capitalism altogether, there is one interesting counter example that should give pause to one's thought. There is one hedge fund company

[33] *Ibid*., p. 188.
[34] *Ibid*.

on Wall Street that guarantees any loss of any client's investment up to four million US dollars.[35] How does this company do this? While the author of this chapter cannot answer this question, it would be assumed that this company is sufficiently profitable such that it can make this guarantee. What is notable is not that it is capable of making this guarantee; what is notable is that it makes the guarantee. By making such a guarantee this company is making the statement that it regards the potential loss of its client's funds as completely unacceptable. This standard of ethics is apparently commercially viable. This is an example, not of practicing risk management in the context in which it is most commonly accepted (that of a fund manager), but rather of practicing not taking risk in the first place (with other people's money).

14. On the hindsight objection and the problems with risk management

Before departing from this topic, we must address an objection commonly brought forth, which in the designation of the author may be labeled, the 'Hindsight Objection'. The following is the 'Hindsight Objection'. There are always warnings of disasters and they are commonly ignored. When no disaster occurs, these warnings are forgotten. If one went back to every successful venture; e.g. a space flight, one would find the ignored warnings. Therefore, if one infers from warnings that an unsuccessful flight occurred, one is inferring from unwarranted premises. Whenever one goes back and traces the causes of disasters to unheeded warnings, one is justifying warnings through hindsight which is always 20-20. Hence, the label, the Hindsight Objection.

There are two major replies to this objection. The first reply is that it is entirely hypothetical. In order to make good on the Hindsight Objection, one must bring forth evidence to support it. In other words, one must take a successful venture; e.g., a space flight, and show first of all that there were red-flagged warnings that the flight should not have taken place. If the 'red-flagged' designator is not in use, the warnings must be shown to be of high-priority status, i.e., it should be demonstrated that such warnings were equivalent in status to the warnings not to launch disastrous flights, such as the dire warnings that were lodged in the attempt to stop the launch of the *Challenger*. Warnings must be vetted in terms of the criteria spelled out above, in terms of source, form and content. What is commonly brought forth as "hindsight evidence" in the case of the Challenger are operational parts which are designated as Criticality 1, of which there were many such on the Challenger which did not fail. A Criticality 1 designation indicates that the failure of such a part would result in the loss of the vehicle and human life. But a Criticality 1 designation is not equivalent to a red-flagged warning that the part designated as Criticality 1 was likely to malfunction. One cannot point to parts with a Criticality 1 designation that did not fail as evidence that there were warnings that failure was imminent. A part with a Criticality 1 designation could, theoretically, be extremely safe. To count as a legitimate warning, there must be a specific, high priority (red-flagged) warning concerning the faulty design or operational capacity under certain weather conditions, etc., of a part which possesses a Criticality 1 designation.

---

[35] *Cf.*, Peter Lattman and Jenny Anderson, 'For 92nd St. Y, a Break from Wall Street Worry,' *NY Times*, November 29. 2011. In this article, this guarantee is offered by the hedge fund manager, John A. Paulson.

This first condition, the prior presence of red-flagged warnings that meet the above criteria in terms of source (knowledgeable expert witness), form (red-flagged), and content (spelling out the specific flaw in the part), to the best of my research and knowledge, have never been met. The Hindsight Objection is always made as a purely hypothetical objection without any evidence for its truth value ever put forth. In the case of space flights (this example is used because it is best known to the author), one must show that there were other launches in which the senior scientist issued repeated red-flagged alerts and all of the engineers and managers voting on the planned launch voted unanimously against the launch. Has anyone ever brought forth any evidence in support of the Hindsight Objection claim?

The second major objection is that even if one could offer an example of a flight which met the equivalent set of warnings that were issued in the case of the *Challenger*, one might miss the point that such warnings are not intended to be construed to be 100% reliable predictions concerning the particular flight in question. Such warnings are not meant to be on the level not of, if this flight will take place, but of when this flight, or a similar flight, takes place, such a likelihood of such a flight ending in disaster is a horrific eventuality. If the warnings are not justified on the basis of one flight, they might well be on the basis of the third or the seventh flight. When one leaves the auto mechanic's shop with faulty brakes, the service manager might warn the customer that the brakes may fail. If the brakes do not fail on the first hill, that does not mean that her or his warning was without value. They may fail on the ninth hill. Thus, even if one could find a case in which dire warnings that fit the above criteria were present and the flight or sail went without incident, that would not prove that all such flights were to be considered safe.

It is informative to reflect on the fact that Boisjoly's warnings cover both the likelihood of the possibility of the *incidence* of the occurrence of the horrific failure of the mission and the *consequence* of the death of the crew and passengers. Thus, both aspects of risk are covered: the possibility of the likelihood of the incidence and the gravity of the consequence. This example of risk taking fits both the criteria of the likelihood of the possibility of the incidence and the enormity of the consequences. Such risk taking is entirely incompetent and unconscionable. *That it need not be restricted in its eventual occurrence in this flight is evidenced by the fact that it was warned against eight years previous.* The resistance to launch on this flight was based on the fact that the weather conditions compounded the already existent risk. *The problem was that the case of the decision to launch the Challenger was an iconic case where it was thought that risk could be managed. In other words, to some extent, the very concept of risk management was at fault.* A better example of a safeguard would have been never to have installed this unsafe part in the first place. That would have been an example not of risk management, but rather of not taking risk in the first place. (After that fatal flight, one of the original designs that was originally rejected was chosen to be used). More obviously, *not launching in adverse weather conditions would be an example of not taking risk. Launching in dangerous weather conditions is an example of attempting to manage risk*. It comes under the thinking of, 'the weather is not good, but we can manage it'. It is not clear what this means. It seems to suggest the belief that 'the weather is not good, but we can chance it.' 'Risk management', when properly analyzed, seems to be equivalent to 'risk chancing' or gambling'. The eminent Nobel laureate physicist, Richard Feynman likened the decision to launch the *Challenger* space shuttle to playing Russian roulette. The proper way of thinking would have been, 'the weather is not good; we are not going to *take* the risk'. *When we frame the decision to be taken in term of risk taking rather than risk managing, it is far more likely that we will act not to take the risk rather than to act to take the risk and then attempt, somehow, to "manage" it.*

It is therefore worthwhile to consider abandoning the concept of risk management and replacing it with the concept of risk taking. When one removes the euphemism of risk management and replaces it with risk taking, it becomes abundantly evident that it is actual human life with which one is taking risk. One is playing G-d with human life. *The very concept of risk management is itself too risky*. The seemingly objective social science language of 'risk management' is in reality a license to treat human life lightly.

15. In conclusion

Whenever, under whatever circumstances, risk management is being considered, one must first consider whether any risk must be in existence in the first place. In a condition of great advantage, such as commercial air travel, the existence of risk is minimized in the first place. After its minimization, it is not a case of risk management that is practiced as much as it is a case of safety protection. It is not a case of a mere shift in wording, because if we employ the language of safety we will be reminded of the ethical priority of the preciousness of human life (together with its corollaries of future generations and the health of the planet). The risk of airplane travel is minimized with proper attention to mechanical safety, pilot training, etc. *The airplane pilot has the right to make the decision not to fly when she or he decides weather conditions are not safe. Such an important decision is placed in the commercial pilot's hands. The judgment as to whether it is safe to fly and therefore to risk lives of the crew and passengers is given to the most experienced and most expert person among the primary actors who will be actually taking the risk.*

To make a brief shift to another example to ensure that the problems with the concept of risk management are not confined to the case of the *Challenger*, many years ago, the author of this chapter personally recommended to cruise ship management, that all commercial cruise ships be outfitted with cockpit voice recorders (black boxes) as are airplanes. Originally, when the author made this recommendation, this author was told it was against strong, naval customs. It is refreshing to learn that strong customs can be changed. This provision of a black box is, of course, strictly speaking, a device that serves as a prevention of future disasters rather than a safeguard for the sail in question.

Whatever we do in life, we cannot close our eyes and pretend that we can remove all dangers from any human endeavor. But a step forward can be made if we consider safety protection as our highest priority. In doing so, we must pay attention to the *taking* of risk in the first place and not be content with *managing* risk that need not exist in the first place. When we do take steps that minimize risk, we should not consider these steps to be "managing" risk, but rather as steps that *reduce the consequences of risk*. While cruise ships are now required to have enough life boats to accommodate 100% of their passengers, the diesel engines with which they are equipped can only power them up to 6 knots and cannot carry passengers to a faraway shore if harsh wind and sea conditions exist.

The idea of "risk management" implies that risk must be present. We must ever guard against complacency. If we keep uppermost in mind that the preciousness of life is our highest priority, we will ensure that the presence of the possible incidence and the consequence of risk is kept at the lowest possible point. *If we change our language from the language of "risk management" which implies that somehow there must always be a risk present, and it is our task to manage it, to risk taking, we will be more alive to the ethical responsibility which is involved in taking risks with our own or other people's lives.* We will be more inclined to work sincerely to minimize the possibility of risk and to reduce the effect of the consequences of risk.

The notion of providing enough safety boats for half of the passengers (the model of the *Titanic*) fits perfectly into the concept of "managing risk". When we do this, we have performed some kind of cost-benefit analysis, or, to speak more strictly, some kind of probability-benefit analysis, and have reached a decision that by providing life boats for only half of the passengers we are fulfilling the responsibility of "managing risk". It is not clear how this decision was reached. Perhaps, it was reached by assuming that only half of the passengers would make it to the life boats, hence, by providing half of the needed boats, we have "managed" the risk.

In conclusion, if we use the language of risk taking, the provision of life boats for half of the passengers on board is still a case of *taking risk with half of our passengers' lives*. On the other hand, if we take the language of risk taking seriously and consider that human life is precious, we would not sail, that is, we would not take risk, unless there were enough life boats provided for 100% of the passengers. (If we rely upon life preservers, for example, those in ocean waters would only live for a few minutes because of the icy temperature of the sea).

Everything in the area of risk management is a matter of ethics. Do we value human life? What do we mean by the phrase 'acceptable risk'? Such a concept can only be tolerated if there are safety provisions that will fully protect and preserve life if the consequences of the risk taking threaten human life. Otherwise, there is no such concept as an 'acceptable risk'. Was the decision to provide life boats for only half of the passengers aboard commercial vessels based on past performance data? In the case of the *Titanic*, lifeboats for only half of the passengers were provided and, as a result, the lives of 1,523 men, women and children were lost.[36] 1,523 men, women and children were killed based on risk management. The *Titanic* disaster was in 1912. Now, 100 years later while the custom is to provide enough lifeboats for all passengers, the lifeboats are prepared with engines that will leave their occupants at the risks of the high seas.

It is hoped that this introduction to the new idea of risk taking as opposed to risk management will create reflection and commitment to a greater ethical sensitivity when we consider how our decision making may affect other people's lives. Do we ever have a right to take decisions that risk other people's lives? If the message of this chapter is heard, then we do not ever have the right to risk other people's lives unless we provide full and adequate protection for the consequences of the risk that we are taking with other people's lives. If we change our language habits from commonly using the phrase 'risk management'

---

[36] *Ibid*., p. 87.

to never using this term, but rather replacing it with the phrase, 'risk taking', we may have a more ethically responsible world.

Author details

Robert Elliott Allinson
Soka University of America, USA

16. References

Allinson, Robert Elliott, Diane Vaughan, The Challenger Launch Decision, Risky Technology, Culture and Deviance at NASA, Chicago and London: The University of Chicago Press, 1996, pp. 575, Business Ethics Quarterly, Vol. 8, No. 4, 1998, pp. 743-756.

Allinson, Robert Elliott, Diane Vaughan, The Challenger Launch Decision, Risky Technology, Culture and Deviance at NASA, Chicago and London: The University of Chicago Press, 1996, pp. 575, Society, Volume 35, Number 1, November-December, 1997, pp. 98-102.

Allinson, Robert Elliott, 'Risk Management: demythologizing its belief foundations,' International Journal of Risk Assessment and Management, Volume 7, No. 3, 2007.

Allinson, Robert, Elliott, Saving Human Lives, Lessons in Management Ethics, Dordrecht: Springer, 2005.

Bell, Trudy E., Esch, Karl, 'The Fatal Flaw in Flight 51-L,' IEEE Spectrum, February, 1987.

Cook, Richard, Challenger Revealed, An Insider's Account of How the Reagan Administration Caused the Greatest Tragedy of the Space Age, New York: Avalon, 2006.

Corrigan, Grace, George, A Journal for Christa, Christa McAuliffe, Teacher in Space, Lincoln and London: University of Nebraska Press, 1993.

Harris, Charles E., Pritchard, Michael S., Rabins, Michael J., Engineering Ethics, Concepts and Cases, Belmont: Wadsworth Publishing Company, 1996.

Jensen, Claus, No Downlink, A Dramatic Narrative about the Challenger Accident and Our Time, Barbara Haveland (trans.), New York: Farrar, Straus and Giroux, 1996.

Lattman, Peter and Anderson, Jenny'For 92nd St. Y, a Break from Wall Street Worry,' NY Times, November 29, 2011.

Lewis, Richard S., Challenger, The Final Voyage, New York: Columbia University Press, 1988.

Slay, Alton, Post-Challenger Evaluation of Space Shuttle Risk Assessment and Management, Washington, D.C.: National Academy Press, 1988.

Schinzinger, Roland and Martin, Mike W., Introduction to Engineering Ethics, Boston, New York, Toronto: McGraw Hill, 2001.

Vaughan, Diane, The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA, Chicago: University of Chicago Press, 1996.

Section 2

Risk and Supply Chain Management

Chapter 8

Supply Chain Security – Threats and Solutions

Daniel Ekwall

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/48365

1. Introduction

In recent years, the cargo transport process has improved mainly in the areas of logistics efficiency and documentation handling. The World Trade Centre terror attack in 2001 changed the world and with it the conditions for logistics world-wide. The logistics consequences were according to[1]: *It is instructive to note that these disruptions were not caused by the attack itself, but rather by the government's response to the attack: closing borders, shutting down air traffic and evacuating buildings throughout the country*. The aftermath to the attack brought needed attention to the vulnerability of modern supply chains. Supply chain vulnerability reflects sensitivity of the supply chain to disruption [2]. This vulnerability can in many cases be described as "unwanted effects" in the supply chain caused either by internal or external forces that create disturbances larger than the supply chain is designed to handle. The objective of Supply chain security is to prevent antagonistic threats from affecting the supply chain performance. Antagonistic threats and other risks and uncertainties are demarcated by three key words: deliberate (caused), illegal (defined by law), and hostile (negative impact for transport network activities) [3].

This chapter presents first the major antagonistic threats to the supply chain and secondly how these threats should be prevented. This leads to the current development of different supply chain security programs.

2. Supply chain and the transport network

[4] defines the supply chain as, "*The network of organisations that are involved through upstream and downstream relationships in the different processes and activities that produce value in the form of products and services in the hands of the ultimate customer*". These processes can be in different companies or in the same company. The different building blocks in a supply chain can, literally, be located throughout the world and connected through the use of a transport network. The transport network is designed to use economy of scale when moving products

from consignor to consignee in a supply chain, through nodes and links. This means the transport network only physically integrates the supply chain with the fulfilment of its transport demands [5]. Therefore, several different supply chains can be present at the same time and the same place in the transport network. This indicates that the relationship between supply chains and transport activities would be better described with complexity theory, especially if the interactions between components are the object of the research [3].

Looking at transport from a system perspective, we find that logistics is made up of different levels, infrastructure, resources and material known as the three levels of logistics [6]. A logistics system consists of links and nodes, where nodes are geographically fixed points such as factories and terminals, while the links are the elements connecting the nodes, i.e., the modes of conveyance. The flow of materials is the first level of the system, because it is the reason for the system's existence. Moving material from one place to another requires a flow of movable resources such as Lorries, trains, airplanes, and ships. These movable resources need infrastructure like roads, harbours, airports, and terminals [7, 8].

The complexity in logistics can be explained by displaying the four flows always involved in logistics activities. The flows of material and resources are mentioned already. These two flows represent the "physical" part of logistics, but the other two flows, monetary stream and flow of information, are just as necessary to make the system work [6]. The four flows of logistics need geographical fixed constructions and infrastructure to fulfil the scope of logistics. Some of the infrastructure is owned and used exclusively by one company while some is co-owned or owned by governments. The four flows of logistics and the necessary infrastructure are the five needs for logistics fulfilment [9].

The cargo thief aims to remove goods from the goods flow by attacking the movement of resources and/or the infrastructure it uses. A potential perpetrator can also utilize the information flow in order to better plan the theft of goods or commit a fraud which targets the flow of capital. This paper uses primary the three elements, flow of goods, movement of resources and infrastructure of the five needs for logistics fulfilments. The frame of reference uses the routine activity theory from criminology to explain the interaction between the transport network and potential perpetrators, where the theft opportunity is determent by each unique configuration of the five needs for logistics fulfilments and then exploited by a potential perpetrator.

3. Threats to supply chains

3.1. Usage of official macro statistics

The usage of general statistics (mainly different types of criminal statistics) can provide a hint about the general criminal threat in a country or local area [9]. This fact is well known. Sometime is the relationship between reported crimes and public fear of crime direct whiles other times more indirect. Thus, it is possible that inverts that relationship and use macro level statistics in order to hint criminal hotspots in general. Important to remember is that this only provides a general hint about criminal threats and this clearly limited the possibility to draw far-fetched conclusions unless more detailed data is added. A good indicator for criminal threats and common distrust in the society is level of corruption in a country [10]. The higher level of corruption, the higher is the distrust in the society and this may lead to a lower will to report crimes to the police. This is only a weak indicator.

A good indicator of the violence present in a society is the number of homicides per 100 000 citizens (lower number means safer). The average homicides rate 2007 in EU which is 1.4 in general but increases to 1.9 in capital cities [11]. This indicates that the country sides are safer than bigger cities, and the long term trend is a lower average homicides rate, which means a safer society (relatively). Other good indicators of criminal threats are the reported total number of crimes as well as the reported number of thefts and robberies. Risk rating (both relative and absolute) are good indicators on more transport related crimes. It is important to remember that each country may (are likely to) have different definition of crimes and guidelines for data collection [9].

The general criminal trend in EU is that crime is declining from the pike around 1995. According to [11] have their being an increase in reported crimes types violent crime (up 3%), drug trafficking and robbery (both up 1%) in the period 1998-2007. During the same period has seen a decrease in motor vehicle crimes (down 7%) and domestic burglary (down 3%). There seems to be a time difference between countries in the trend but nevertheless the declining in reported crimes and even criminal patterns are surprisingly similar between member states [12]. The reason for this decline is according [12] the "*changing demographics, among other factors, have played a causal role in the decreases in crime across the Western world. Since the bulk of common crimes are committed by young males*". There is also a suggestion that the better policing and/or more severe sentencing contributes to the declining criminal trend [12]. The different official statistics supports these conclusions.

3.2. Shrinkage and theft

The term inventory shrinkage is the loss of products between the point of manufacturing or purchase from supplier and the point of sale. According to [13], the average shrinkage rate is 1,8 percent of total annual sales. This means a total loss of $33,21 billion annually in the U.S. The report points out four major sources of shrinkage: employee theft, shoplifting, administrative error, and vendor fraud. Therefore, three of four sources for shrinkage are criminal actions. The losses in the European fast moving consumer goods are for 26 percent in manufacturing, 8 percent in distribution, and 66 percent in retail [14]. Shrinkage during distribution/transport is approximately 0,14 percent of annual sales for all types of products. According [15] is the worldwide loss ratio as 0,025 percent of the total revenue ($307 billion revenue and $77 million in losses). Benchmark participant loss rates varied from 0,0038 percent to 0,25 percent of total revenue. The three different reports [13, 14, 15] indicate that the annual shrinkage during distribution/transport would be 0,025 percent to 0,14 percent of annual sales. This loss ratio is compared with the loss ratio for retailers (1,75 percent) and manufacturers (0,56 percent) [14].

There is a significant problem with the theft of cargo worldwide. It is estimated that theft represents a loss of at least US$10 billion per year in the United States and US$30 billion worldwide [16]. The value of cargo theft for the European Union is estimated to be €8.2 billion annually, an average value of € 6.72 per trip [17]. Gathering accurate numbers for cargo theft losses is difficult or impossible in many cases, due to limited reporting by the transport industry and the lack of a national law enforcement system requiring reporting and tracking uniformity [18]. Despite these figures, cargo theft generally has a low priority status in most countries and is often perceived largely as the cost of doing business.

A problem related to cargo theft is the theft of vehicles and the lorry-driver's private property. There are many reasons behind a truck theft, but they can be described basically with three main characteristics - value, cargo carrying ability, and valuable documents [3]. *The first characteristic,* value, represents the truck's value as all objects and can be sold and exchanged for money. *The second characteristic, carrying ability,* refers to the general propose of a truck. The vehicle and its load were targeted in 63 percent of the attacks [19]. The truck can be stolen with the current load where the goods are the desirable object and the truck is only the simplest method to move the goods to a warehouse or to another truck for further movement. A truck also can be stolen for other criminal activities.

The third characteristic of theft problems toward freight is the attack for the lorry-driver's private property or other types of valuable documents such as credit cards, mobile phones, and digital cameras stored in the truck during transport. According [19], 17 percent of all drivers suffered an attack during the past five years, 30 percent were attacked more than once. Of all drivers attacked, 21 percent reported they were physically assaulted during the attack [19]. This type of attack represents a considerable amount of the total, but nothing was stolen in 38 percent of attacks against trucks [20]. However, even if nothing was stolen it was still a crime against a part of the transport network and therefore shall be seen as an antagonistic threat. 70 percent of attacks against road transports occur between 22:00 in the evening and 06:00 in the morning [19]. Therefore, it is possible to state that time of day plays an important role in antagonistic threats. According [19], the direct cost for an attack is approximately €25000 per attack, including theft of vehicles, load and the driver's personal belongings.

Regardless of which of the three characteristics of theft problems the motivated perpetrator uses, there is a number of commonly defined modus operandi or methods to attack trucks. These different modus operandi are used differently depending on where the attack is executed. The different locations are described in terms of different steps in a road transport from consignor to consignee, which starts with loading the goods and ends when unloading them. Eurowatch has developed a threat/risk matrix based on the data on cargo theft in road transports over a seven-year period [23]. The matrix presented in Table 1 maps modus operandi and location of attacks against each other.

A quick analysis of the matrix points out some obvious relationships. The method *fake accident* is best suited to deceive a truck driver to stop during driving and then conversion to a hijack. The same course of events can be created with the use of *fake police* tactics. The threat/risk matrix points out the most dangerous location to be *near end location* or at *insecure parking* depending on which modus operandi is considered most threatening.

	Hijack	Robbery	Theft from vehicle	Theft of vehicle	Fake police	Fake accident	Deception
Load point	2	3	2	3	1	1	4
Driving	4	1	1	1	4	4	2
Insecure parking	2	4	4	4	3	1	2
Secure parking	2	2	3	3	1	1	2
Near end Location	4	3	3	4	3	1	3
Unload point	2	3	2	3	1	1	4

Table 1. Threat/risk matrix, road transport using Eurowatch data 2002-2009, 4 represents the highest risk [23]

The greatest source of risk for businesses is trusted insiders [21]. Some authors consider insiders to be involved in approximately 60 percent of all losses [22]. According to [23] is 65 percent of all "whole load losses" related to the use of inside information. Others claim there are no reliable figures [24]. This is interesting when considering the fact that most countermeasures are implemented to reduce external theft [25]. An internal perpetrator acts not randomly or in an unstructured way, but more as a response to social and environmental factors present in the work environment [22]. According [26] is the complexity around insiders and drivers expressed: *"Some estimates indicate a high level of driver involvement, but drivers are possibly the weakest link in the security of the supply chain. They are also the first line of defence and there is a need to train and educate them on cargo crime and personal safety issues whilst on the road."* This leads to that the potential perpetrator both can be external to and internally involved in the supply and/or the transport chain.

3.3. Terrorism

The word "terror" is a Latin word meaning "to frighten." Consequently, a terrorist is a person that intends to frighten others through fear. The term terrorist/terrorism is itself controversial because its key signature is political and it has been used by states to illegitimatize political opponents. This leads to a vindication of the state's own use of terror against its opponents [27]. The lack of a universal definition of consequence of this is best explained with the cynical comment "*that one state's terrorist is another state's freedom fighter*". Regardless, terrorism is definition by [28] as, "*Terrorism is not an ideology or movement, but a tactic or a method for attaining political goals.* Terrorism is one of the major obstacles for meaningful international countermeasures.

The World Trade Centre terror attack in 2001 changed the world and the conditions for logistics worldwide. The aftermath of the terrorist attacks clearly indicated that logistics operations will suffer consequences of an attack. [29] state that, *"over the longer term, there is a question of whether the attacks can have a negative impact on productivity by raising the costs of transactions through increased security measures, higher insurance premiums, and the increased costs of financial and other counterterrorism regulations"*. The motivation for the majority of terrorist attacks is because the perpetrator intends to influence and alter the current balance of power in a certain direction [30]. Both the current balance of power and the potential effects on it may only be understood within the perpetrator's own mind. Therefore, the non-economically driven antagonistic threat is more nuanced, uncertain, and harder to predict that other types of antagonistic threats. According to [31] the fear for terrorist attacks is an extreme form of perceived risk. The definition of terrorism influences the difficulty to present valid statistics for the category, because one source may classify an incident as a terrorist act while another considers it to be a "regular" crime. According [28] is the official reported number of terrorism attack in EU declining (581 attacks in 2007 and 294 attacks in 2009) and the major threat (in numbers) comes for separatist movements (Basque and Corsican) whiles Islamist terrorism is still perceived as the biggest threat.

The modern or new thing with terrorism is not the use of violence to influence and alter the current balance of power in a certain direction. According to [32] is: *"Terrorism in all its forms, by its very nature, an asymmetrical response to superior force, and terrorist have always used their capabilities as force multipliers – usually through the exploitation of terror. The generation of fear, in effect the use of purposeful violence as a form of psychological warfare can now be carried much further, enhanced by the modern media and the proliferation of mass media as much as by the proliferation of weapons"*. The new thing with terrorism is therefore more related to development in media technologies than to vulnerabilities in supply chains. Nevertheless, terrorism is a special form of antagonistic threat that needs to be managed, in one way or another.

Reviewing official terrorist statistics from one global source (MIPT Terrorism Knowledge Base) leads to the following transport related conclusions:

- Transport activities represent 4% of the targets in 2006 and 5% in 2007.
- The main modus operandi for attacks is armed attack (38% 2006 and 31% 2007) and bombing (51% 2006 and 54% 2007).
- The main area for terrorist attacks are Middle East/Persian Gulf (68% 2006), and Asia (24% 2006).

The terrorists prefer to use bombs and armed attacks because 80-90 percent of all attacks used these tactics. The targets are rarely transports but more in form of police and other governmental or religious institutions (~53% of the attack 2006). The interesting feature is related to the geographic side of terrorism, because the terrorist threat is mainly linked to local/country/regional contexts.

The official statistics for terrorist attacks indicates that, in order to understand terrorists, it is better to focus on possible attacks instead of probability for attacks. In accordance with this, there is no objective [33] way to determine who is a terrorist and who is not. It has all to do with the context to the terrorist threat. This follows the same logic when terrorism is presented as black swan problem [34]. Nevertheless direct and indirect effects from a terrorist attack or threat will affect the global flow of goods and thereby, to different extent, the global economy.

3.4. Smuggling of goods

The primary target of illegal goods is the black market. The black market consists of places and situations where products with doubtful or no legality are traded for money. This market is subject to the same forces of supply and demand as legal ones. Buyers of these illegal products are everywhere. Statistical reports show that counterfeited and pirated items amounted to $176 billion in Europe in 2007 [35]. According to [36], in 2006, nearly 3 million pharmaceutical products were found to be counterfeit. Product smuggling does not necessarily mean that the product is illegal everywhere. What is legal in one country can be illegal in another, which creates the possibility that the actors in smuggling can be legal companies that are trying to access a market that is prohibited for them. An example of this is Western companies that smuggled products into former communist countries during the Cold War era.

The supply of a typical black market (both authorities and customers knows that the product is illegal) can be illustrated with the illegal smuggling of cocaine to USA. The illegal drug supply chains come mainly from South America. This depends on that the raw material, coca leaf, is grown there. The smugglers use land, sea, and air routes to get past US authorities. The whole distribution of cocaine is controlled by Colombian-based organized crime, but in recent years it has started to cooperate with Mexican criminals to streamline the logistics and share the risks. The Colombians have organized their operations in a business-like manner, creating cells for special purposes like warehousing or transport [37]. According [38], the illegal drug markets are best understood as having high adaptation and great resilience to always supply their products to the end user. This resilience and adaptive ability is clearly found in the logistics system setup and can be understood and explained with the concept of risk for detection presented in this paper.

The supply of a typical gray market (only authorities know that the product is illegal) can be illustrated with the illegal smuggling of counterfeited products. The gray market involves the diversion of goods from legitimate supply chains [39]. The only distinction is the risk for discovery from the authorities or the company whose products are copied. This diversity leads to a different design of the supply chain. The location of the production facilities is subject to the risk of discovery. Normally, counterfeited production units are placed where the risk for detection is low combined with the normal business problem as different types of costs and quality aspects. A counterfeited supply chain uses the freight routes and port activities in the same way as legal supply chain does. Among the receiver countries, Europe and the US are favorites, just as Africa is the favorite for transit activities. Confiscated products that have not been produced in Africa, like jewelry and CDs, show this, because the African market does not have the ability, in general, to buy that type of product. Countries in Central and South America act like magnets for counterfeited products. Purchases of counterfeited goods to launder money occur in larger numbers there than anywhere else in the world [40]. Large stocks of illegitimate products are easily shipped from parts of South America to Central America, where they are big consumers of that type of product. Organized crime also uses Central America as the base for shipments of illegitimate goods to North America. The situation in Europe makes it the most lucrative market for counterfeited products. The types of confiscated goods at the external borders of the EU are different from other places in the world. This indicates that the dealers of counterfeited products adjust products to each market's special condition. They look at the fashion, culture, and buying habits of individual countries [40].

Both types of Illegal supply chains use the international flow of containers to transport their products all over the world, regardless if the product is counterfeited or an illegal drug. Criminals try to delude customs' watchfulness by "breaking" their way through from the area of production to the area of supply, and avoiding direct paths that are well known to the authorities.

A problem linked to smuggling is the manufacturing of products without intellectual rights, or the production of counterfeited goods. Everything that has been produced can be reproduced by someone else. In order to bring counterfeited products from the production site to the end user, they may have to cross several national boundaries as well as intellectual property legislation. The counterfeiting business evolves constantly within current trends and technologies. The production and distribution of illegal products is performed under the risk for detection and this diversity leads to a different design of the supply chain [41]. The location of the production facilities is subject to the risk of discovery. Normally, illegal production units are placed where the risk for detection is low, in line with the normal legal business problem of where to produce according to different types of costs and quality aspects. Then, the illegal products are distributed by trade routes and port activities in the same way as legal products [42]. The pollution of illegal products in the legal transport network is a serious problem. The most common countermeasure against smuggling is the inspection of cargo carriers when they cross a national border. The mere existence of these inspections creates disturbances in the transport network, even if no illegal product is discovered.

The counterfeiting business evolves constantly within current trends and technologies. The illegal products are then distributed by trade routes and port activities in the same way as legal products [42]. According [43], the discourse on the gray market is filled with the idea of a criminal underworld in order to separate it from the legal/normal upper-world. In reality is it very difficult to establish the underworld/upper-world image. The legal companies naturally are not pleased with the competition from black market actors. They may not be pleased with legal competition either, but that is another question. Striving for better business deals and the globalization trend that started centuries ago led to an embedment of illegal actions within legal markets [42]. This implies that the old black markets have been integrated with legal transactions, and today's markets contain every shade of gray when referring to the legality of the markets as a whole [42]. Legal businesses are concerned especially with the problems of counterfeit branded products and the theft of their own products. To increase the efficiency of detecting counterfeited products, legal businesses use cutting-edge technology and security actions [97].

3.5. Piracy

The modern types of pirates do not act officially of any specific courtiers' order but research has indicated relationship between piracy and weakness of central governances [44]. In essence is piracy an international crime against all states and the perpetrators can be brought to justice in all everywhere [45]. In recent years has the threat from piracy against sea shipment (direct threat) and also against the different supply chains utilizing sea shipment (indirect threat) received increasing attention [44, 46, 47, 48]. The real increasing threat from pirates, primary at the horn of Africa [45, 47], has resulted into both a changes in shipping routes (Sullivan, 2010) and also a naval response from several countries [46] that are depended on a smooth passage of cargo carrying ships. According [49], was there 489 attacks last year (2010) and over a twelve year period was there about 347 attacks/year globally [43]. Piracy is an increasing problem, especially near Somalia [48]. The pirates are changing their tactics and targets to use more sophisticated weaponry and apply more advanced techniques, all in order to improve their own success ratio. Table 2 presents the current trends and patterns in piracy.

Year	**Number of acts**	**Lives lost**	**Wounded crew**	**Missing crew**	**Crew hostage/ Kidnapped**	**Crew assaulted**	**Ships hijacked**	**Ships missing**
2006	254	17	23	0	224	225	10	0
2007	310	22	75	57	223	39	18	0
2008	330	6	22	38	773	21	47	1
2009	406	8	57	9	746	2	56	2
2010	489	1	27	0	1027	30	57	12

Table 2. Trends and numbers in piracy 2006-2010 [48]

As stated in table 2, piracy is an increasing threat and especially alarming is the increasing use of violent and kidnapping of crew members. According to [48], piracy cost the maritime industry between 7 - 12 billion dollars a year. In addition to this shall also the cost for rerouting to avoid pirate infested waters, ransom payments and support from various organizations be added.

4. Preventing the threats - solutions

4.1. Interdisciplinary research mixing criminology into logistics

According to several authors [50, 51, 52], criminology is interdisciplinary research of the history and future of crimes, and this paper follows this tradition by using theories from criminology to strengthen the field of logistics. This mix of theories also challenges the predominant research approach in logistics related to tangible artefacts [53], and human intervention or influence to a smaller extent [54]. The reason for this approach is that violation of law is considered a human attribute. Criminology distinguishes three elements of a crime that are present in all sorts of crime ranging from occasional violence to advance and complex economic crimes [55, 56, 57]. The elements are:

1. Motivated perpetrator
2. Suitable object
3. Lack of capable guardian

These three different elements can be described as:

Motivated perpetrator: The perpetrator is an individual that, based on the outcome of the own decision process, commits a certain action or prepares for a certain action that is prohibited by locality or country of international law. The perpetrator can be modelled with two different categories depending on how decisions are made by each individual, namely rational choice theory (also known as the economical man theory) or determinism [54, 58]. It is commonly agreed that different crimes demand various mixtures of rational choice and determinism from the perpetrator's side, where crimes of passion (sexual crime, etc.) are considered more deterministic than property crimes (economic crime, etc.), which are more rational [59]. Thus, therefore can the general description of human behaviour be described as acting rational on the margin or limited (by circumstance, choice or mixture of both) rational choice [60, 61].

Object: The desirable outcomes or objects for the motivated perpetrator differ greatly depending on the motivated perpetrator's decision process. Normally is it suitable to describe the object as the primary or direct reason for the action, but also as secondary or indirect reasons [56]. The primary objects can be shipped products, resources used, infrastructure, or even the media attention an attack will receive (terrorist attacks, action junkies etc.). It is in the relationship between object and motivated perpetrator that the categorisation of the antagonistic threat is found [3].

Lack of capable guardian: The preventive measures that can be induced to alter the motivated perpetrator's decision process are called security [62]. If the security measures are considered insufficient by the motivated perpetrator, then there is nothing to prevent the crime [56].

Most important to remember about the elements of crime is that it is first when all three elements comes together at the same time that a crime is possible. This means that if one of the three elements is missing than is crime impossible. Any combination of lack of security and target are normally referred to as a crime opportunity.

4.2. Opportunity to crime

Crime opportunity is a cornerstone of criminal behaviour. There are ten crime opportunity principles as follows [63]:

1. *Opportunity plays a role in causing all crimes,*
2. *Crime opportunities are highly specific,*
3. *Crime opportunities are concentrated in time and place,*
4. *Crime opportunities depend on everyday movements,*
5. *One crime produces opportunities for another,*
6. *Some products offer more tempting crime opportunities,*
7. *Social and technological changes produce new crime opportunities,*
8. *Opportunities for crime can be reduced,*
9. *Reducing opportunities does not usually displace crime,*
10. *Focused opportunity reduction can produce wider declines in crime.*

Some of these principles are self-explanatory and easy to understand. All of them are valid for every type of crime and therefore they are also valid for crimes committed against the transport network. The more interesting examples of these opportunities will be explained and described later in this thesis. The most important thing to remember about crime opportunity is that an opportunity alone does not explain why a crime occurs because a crime needs a motivated perpetrator and opportunity to occur [64]. The theory of crime opportunity also refers to the fourth principle of microeconomics [57] - *people respond to incentives* - and there the degree of necessary opportunity or incentive depends on the individual. The incentives could range from vindication to morality, ethics, altruism, or determinism [3, 64]. Altogether, this leads to that the relationship between threats (motivated perpetrator) and countermeasures (security) linked around a desirable outcome or object, are complex and contextual depended.

4.3. The two different outlooks on mankind in criminology

It is possible to separate mankind into two different categories, depending on how decisions are made. This separation is a theoretical construction and its validity varies for every person in every situation.

Modern criminology uses rational choice theory as the basis for research. Rational choice theory, also known as rational action theory, is a framework for describing and modeling social and economic behavior. This theory originates in the idea of the economic man in economic research, primarily microeconomics. This theoretical model is also central in modern political science and scientific fields such as sociology and philosophy. Rational choice theory assumes that individuals choose the best action according to the constraints, opportunities, functions, and abilities they face. Today, rational choice theory in microeconomics is defined best with the first four principles of microeconomics [58]. In short, the theory states that every presumed criminal is should be considered a rational person who makes decisions about potential crime from relationships between the benefits of the crime and the troubles and risk it brings. By increasing the perceived trouble and/or risk, it is possible to reduce criminal activity with this perspective [55].

The opposite of rational choice is determinism. The idea is that the course of events depends completely on existing conditions. This approach refuses the idea of free will - everything is predetermined. In reality, every individual is a mixture of their ability to be influenced and the lawful. Among social scientists today, the idea of restricted free will is a common and useful insight. The cause of these restrictions can be found in the individual biological or psychological vulnerability, way of life, upbringing, social group, ethnic background, or society in which the individual lives, and how this affects his or her life with regard to their ethnic background, gender, and social position [55].

These opposing perspectives of the human being as either the master of his own life or as a victim of circumstance can be found in every aspect of criminology. This contrast affects not only how we see the causes of criminal behaviour, but is also important with respect to the social response to criminal behaviour [55]. According to [65], the current approach toward crime, punishment, and pardon is a good way to understand the surrounding society. It is commonly agreed that different crimes demand various mixtures of rational choice and determinism from the perpetrator’s side, where crimes of passion (sexual crime, etc.) are considered more deterministic than property crimes (economic crime, etc.), which are more rational. That being said, the big difference appears when discussing the possible punishment for a certain crime. If an individual is responsible for his or her actions (rational choice), then the possible punishment will deter the crime; however, if the individual is a victim of circumstance (determinism), then it is useless to punish the individual. Therefore all crime prevention methods assume that an individual is responsible for his own actions and that he can perceive the consequences of those actions. The big question for the rational choice perspective is how each individual estimates the risk in a rational way. An individual that has received a formal or perceived punishment for an action previously should be less likely to commit that action again. The outcome of a formal punishment on the perception of risk is mixed [66] and each individual should be considered not rational, but to act rationally on the margin [3].

4.4. Criminal prosecution and punishment

This discussion of the two outlooks on mankind in criminology clearly demonstrates the need to add theories from criminology into logistics. Several logistics authors claim that the weak prosecution of criminals is one reason behind the increasing need for security [16, 67, 68]. In the context of the two outlooks on mankind in criminology, [16, 67, 68], state the threats against logistics activities is simply based on the rational choice theory, while authors in criminology [55, 63, 66, 72, 74] refer to an individual acting rationally on the margin which eliminates the deterring effect of potential punishment that [16, 67, 68] suggest. The deterrent effect that a punishment can have on a perpetrator not to commit a certain crime is very low, due to that the perpetrators do not plan to get captured. Therefore has the risk for detection a bigger deterrent effect that the potential punishment [41]. The authors' [16, 67, 68] request for stronger prosecution of criminals can be seen as a way to understand the surrounding society in general [65]. Compare that to [69] who state that risk management and TQM systems are normally linked to a punishment and reward system for the users. The difference between legal prosecution achieved by a state and a punishment and reward system controlled by a company or similar organization is that legal prosecution involves a weaker relationship between action and consequence for the perpetrator.

4.5. Perpetrator, opportunity and security

The relationship between security and opportunity is the predominant understanding of security in different contexts. This depends on the premise that security only can deter or repel a motivated perpetrator from committing a crime by limiting the opportunities for a certain crime. The most important thing to remember with crime opportunities is that an opportunity alone does not explain why a crime occurs because a crime needs a motivated perpetrator and opportunity to occur [62]. Opportunity plays a role in causing crime, and these opportunities are highly specific, concentrated in time and place, and depend on everyday movements. These opportunities can be reduced and focused opportunity reduction can produce wider declines in crime. It is the theory behind security. The real problem occurs when an organisation's security capability is lower than the capability of the potential perpetrator. The driving force within each potential perpetrator can be vindication from morality, ethics, altruism, or determinism. Therefore, the relationship between security and opportunity must be understood from the viewpoint that the parties involved (stakeholders, actors, and humans) have different individual incentives to exploit opportunity that security needs to address. Consequently, security cannot be seen as only opportunity limiting but also as incentive limiting, making security a preventive factor on both sides of the opportunity (pre- and post-event).

Crime opportunities depend on routines or predictability within certain boundaries. This is the routine activity perspective in criminology and it argues that normal movement and other routine activities play a significant role in potential crime. The routine activity theory states that potential perpetrators may seek locations where their victims or targets are numerous, available, convenient, and/or vulnerable. [70] uses the illustration of "*how lions look for deer near their watering hole*" to explain the practical relevance of the routine activity perspective. Social disorganization in combination with the routine activity theory can provide a wider and better explanation of property crime.

4.6. Incident preventing

The security of freight transport was long under-developed, but since terminal security has improved, theft incidents have increased in the links between terminals [71]. This development is also valid from a supply chain perspective; while security in manufacturing facilities normally is focused and well-managed, the rest of the chain is without security [72]. Security during transport is necessary to prevent unwanted negative disruption in the flow of goods. Transport security means the interaction between physical obstructing artifacts (locks, fences, Closed Circuit Television (CCTV) etc.) and the intervention of humans, with the aim of reducing theft, sabotage and other types of illegal activity. The technological development of the range and sophistication of anti-theft devices and after-theft systems is increasing rapidly [73]. There are different preventive methods that can be used to reduce the risk of a cargo theft incident, but the primary method is to use physical security countermeasures correctly (fences, locks, seals, guards etc). The objectives for these types of countermeasures are to make the theft both harder and riskier to commit. The next important countermeasure is the control and trust of employees in the company. This method targets the internal theft problem and can be subdivided in two parts: present employees' supervision and new employee reference checks [74].

The configuration of transport networks leads to the need for security measures in different forms, depending on the exact function and appearance of each nod and link. All different theft preventive methods used can be explained with the basic theory of situational crime prevention. The aim is to reduce factors specific to different types of crimes, locations and situations. The key issue in situational crime prevention is the recognition that a crime often reflects the risk, effort and the payoff as assessed by the perpetrator [75]. The theory does not state that a perpetrator will commit a crime every time an opportunity occurs. Rather, the potential perpetrator makes a calculated decision about the opportunity to commit a crime [76]. In short, a perpetrator acts according to the rational choice theory, seeking to maximize its utility with regards to a particular time and available resources. Since cargo theft is a property crime, situational crime prevention is a useful method to address this problem. Basically this is achieved by applying the following three prevention principles [77]:

Increased perceived effort [78] – Motivation to commit a theft is reduced if the perpetrator believes the crime is too hard to commit. Preventive actions based on this idea can be categorized as physical separation of the potential perpetrator and the object of the theft. This can be accomplished through the use of access control and physical barriers (fences, locks, etc.).

Increased perceived risks [70] – If perpetrators think they will get away with a theft, it is more likely they will commit it. By increasing the risk for perpetrators they are less likely to commit a theft. This can be accomplished with surveillance systems, security personnel, and by increasing employee's security awareness.

Reducing anticipated rewards [70] – People are more likely to commit a theft if they can benefit from it. By making the target for the theft worthless or reducing its resale value it becomes less attractive for potential perpetrators. This can be accomplished by marking the goods with unique numbers or a product destruction device. Good examples of this principle are the safety cases used in transports of valuables and money and the ink tags used in fashion stores.

[71] added a fourth preventive principle based on rational choice theory:

Inducing guilt or shame – A theft is more likely to occur if it can be excused by appeal to reasons such as "the company can afford it" or "I've worked hard for the company but they have not thanked me for it" [79]. This is a form of ethic relativity. Companies and organizations can affect this ethic relativity by using company rules, signs, and regulations that demonstrate the right moral values. When theft is seen as an additional wage benefit for employees, this preventive action has failed [24]. By appealing to people's morals and making it easy for them to do the right thing, it will be more difficult to make excuses [71].

Some of the criticism against situational crime prevention states that this method leads to property crime receiving more attention than is appropriate. Furthermore, situational crime prevention addresses the symptom and not the cause of the crime. This can lead to an excessive trust in technology [80]. Both of these criticisms are valid for the usage of situational crime prevention to hinder cargo theft. This may explain the over confidence in technology in order to prevent terrorism due to that terrorism is not a property crime and therefore has the use of technology an even lower impact on reducing the threat than for a cargo theft incident. The prevention of incidents are traditionally closely linked to the definition of security as a *show of force* which leads to that security work becomes symbolised with uniformed guards and normal police duties [81]. In order to include impact reduction in the term security is the second definition, *freedom from danger*, better to use. This definition embraces all things that allow the organization to act and carryout the business "free from danger". This is from this view point the supply chain security shall be understood.

4.7. Crime displacement

The theory of crime displacement says that crime prevention in one area may have unintentional consequences in other areas or situations. Therefore, crime prevention may not lead to an absolute reduction in crime. The theory of crime displacement is based on rational choice theory, with the following three assumptions about the potential perpetrator and target [69]:

Crime displacement assumes that crime is inelastic [82] - This assumption indicates that the demand for crime is unaffected by preventive efforts. This is not true because all crimes are more or less elastic [83]. Professional criminals are more inelastic while opportunistic criminals are more elastic [75].

The perpetrator has mobility [75] - Perpetrators have flexibility relative to time, place, method, and type of crime. In reality, perpetrators are limited in their mobility, adaptability, and flexibility relative to a particular crime, place, time, and method [76].

There exist unlimited numbers of alternative targets [84] - The perpetrators have unlimited numbers and types of potential targets to choose from. In reality, the number of targets is limited in one way or another [76].

The theory of crime displacement states that rational thinking perpetrators with crime mobility will alter their behaviour in response to crime prevention efforts [69]. Crime displacement will only occur when the alternative crime has a similar cost-benefit structure rationalised within the perpetrator's decision process [85]. Based on the ten principles of theft opportunities presented earlier in this thesis and the configuration of the transport network, it is obvious that all opportunities cannot be eliminated, but they can be reduced by applying substantive preventive countermeasures. The object is to reduce crime opportunities which will lead to a change in potential theft situations. Therefore crime displacement is a valid theory [75].

Crime displacement can occur in several ways. [75] uses five types of displacement - crime, target, method, place, and time. [86] add another type of displacement, the perpetrator. The six types of displacement are explained below [65]:

Crime:	Transfer to other types of crime *Ex: offenders stop doing robberies and instead commit burglaries.*
Target:	Transfer to other types of goods *Ex: offenders stop taking goods and instead target money transports.*
Method:	Better locking devices force the offender to be more innovative *Ex: better doors force the offender to break-in through the windows.*
Place:	Transfer to a less protected target in the same or other areas *Ex: if one area improves security then the offender attacks another area.*
Time:	Transfer to different times of the day *Ex: better night security forces the offenders to strike during daytime.*
Perpetrator:	Transfer to another perpetrator *Ex: Preventing one offender can create an opportunity for another offender.*

The theory of crime displacement does not explain why perpetrators commit a certain crime or why some crimes are more attractive than others. Furthermore, it does not explain the perpetrator's perceptions and reactions to changes in opportunities [76]. Crime displacement is one probable explanation of why criminal patterns change in a certain system. A practical statement about crime displacement is that *if perpetrators have the ability, mobility, and flexibility to exploit the weakest link in the chain, they will do so*. It is the perpetrator's ability to organize a successful crime and their relationship relative to the actors within the transport network that are the fundamental variables for categorizing perpetrators.

4.8. Logistics, risk management and criminology

The three terms security, risk management, and crime prevention often are considered similar and always work together [61, 74]. This idea suggests that security and risk management are good from an ethics point of view because they reduce crime; therefore, more or better security or risk management will reduce crime. The problem here is that crime is defined by a law according to the principle "no crime without a law" [54], while security or risk management has no philosophical attachment to law. Therefore, people on both sides of the law can have better security or risk management and that security and risk management are not necessary against crime.

Security for an individual or a group of individuals can, if unrestricted, jeopardise the security of others by threatening them or transferring threats to them. This type of discussion can be found with philosophers such as Hobbes and Mills. Unbounded or unrestricted individual security could threaten the authority of a state. This problem is demonstrated in the current debate about individual and private secure communication encryption, which some states want to make illegal unless they can break them. Taking this into account, a security problem may or may not be a legal problem. As a concept, security involves a protector or guardian and a threat against an asset or object [74]. This threat can be from either side of the law. To obtain the right security, it is vital to answer who is protecting what, from whom, in which situation, to what extent, and to what consequence [61]. Security can be seen as contextual risk management [3].

Contextual and statistical risk management approaches as crime prevention methods work in different ways and address different types of potential perpetrators. This demonstrates the difference in the philosophical views of contextual and statistical risk management. Statistical risk management needs a fairly predictable world, or at least a larger amount of trustworthy statistics. Since previous events or incidents are the basis of statistical risk management, it cannot deal effectively with a self-inflicted alteration of the threat pattern. Therefore, statistical risk management is effective in crime prevention if the potential perpetrators are limited to unsophisticated and indifferent methods based on opportunistic behaviour [87]. However, as potential perpetrators become more sophisticated and gain more capability, the accuracy of statistical prediction will reduce dramatically. Antagonistic perpetrators study the victim to discover routines and regularities and improve their skills with this knowledge (planning, technologies, and tactics) to maximize their likelihood for success [88]. The prevention of antagonistic threats by following current business trends makes the system even more predictable. Military special forces and similar organizations have proven this time after time [75].

4.9. Transfer of the effect from antagonistic threats

The mitigation of antagonistic threats refers to strategies to transfer the economic impact to another organisation/company. The basic idea is that by transferring the risk to someone else through different types of contracts, it reduces personal risk substantially [89, 90]. The contractual agreements are divided into two different categories - insurance policies and non-insurance contractual agreements between two organisations. Good risk transfer strategy is composed normally of both types of risk transfer.

Transfer of risk by usage of insurance: The insurance principle prescribes that the insurance company takes over the economic impact if something happens that is covered by the insurance contract. Therefore, the risk of this event needs to be rather easy to identify, classify, and determine for the insurance company to estimate the cost related to the risk and determine the insurance premium. This premium is also accompanied by an insurance excess, giving the potential insurance buyer three components to consider: the terms, premium, and excess of specific insurance. The incentive for each individual insurance buyer is the central issue, and in extension, also the potential reduction of a potential loss [3, 91].

Today special insurance policies exist for everything from cargo damage and machinery breakdown to terrorism, war, and general consequence liability. Typically, all insist on special events to be considered valid for a certain incident. Sometimes the terms are so specific that different types of insurance are valid depending on the cause of the incident. One of the best examples is the often-conflicting terms in terrorism contra war insurance. If there is a recognized government behind the incident, then war insurance is valid, but if the incident is caused by someone else, it is the terrorism insurance policy that covers the economic impact. This also depends on the different terms that each insurance company has in their product [3].

Transfer of risk by usage of business power: This risk transfer strategy is commonly used between contractual partners, but is rarely mentioned because it opposes the belief that everybody wants to collaborate fairly and for the greater good [92, 93, 94, 95]. The general idea of this risk transfer strategy is to use the business power of size, information advantage, or control over a critical asset [96, 97] so the business partner can obtain a share of the business risk as a part of the contractual agreement.

4.10. Supply chain security

The terrorist attacks at the World Trade Centre and Pentagon on September 11, 2001, brought attention to security in trade, for more reasons than just the attacks. Three factors can be outlined: first, the globalization of world trade depends on and is generated by the free flow of people, goods, and information; second, the increasing demands from businesses for efficient supply chain operations; and third, the increasing threats of terrorist attacks [98]. The last reason can be seen as an increase in perceived risk for terrorist attacks. This factor can define illegal and antagonistic threats, of which terrorists are one type. Therefore, supply chain security management can be defined as, "*the application of policies, procedures, and technology to protect supply chain assets from theft, damage, or terrorism, and to prevent the introduction of unauthorized contraband, people, or weapons of mass destruction into the supply chain*" [83]. The only problem with this definition is that it does not address the origin of the threat or risk. The five sources of supply chain risks provide that. Supply chain security needs to adjust its policies, procedures, and technology to protect the supply chain from all five risk sources. The flip side of supply chain security is supply chain resilience, or a supply chain's ability to withstand and recover from an incident [83]. Supply chain risk should incorporate security and resilience, where resilience also must handle a near miss incident that affects the performance of the supply chain and from which it needs to recover.

Present supply chain security research outlines several changes for how security in a supply chain should be approached. First, supply chain security should incorporate not only theft prevention but also anti-terrorism measures. Second, the focus is now on global issues and not just local or national issues [99]. Third, when conducting contingency planning, the concept of crisis management is to be included to obtain better resilience. Last, security is no longer an internal corporate question but rather an issue for all actors within the supply chain [83].

[100] suggest that methods and ideas from total quality management can be used successfully to increase supply chain security. The main idea is the lesson from quality management that sample inspection is expensive and useless at the end of the production line. Just like in quality management, supply chain security becomes more effective and less expensive by implementing the right management approach, technology, and re-engineering operational processes. [85] state that security should be integrated throughout the entire supply chain to be successful at a reasonable price. This opinion is supported by several other authors [74, 83, 101, 102, 103].

4.11. Supply chain security programs

Several new security programmes were launched in the aftermath of the World Trade Centre terrorist attack to protect international cargo flow from being abused for criminal (primarily terrorist) intentions without compromising supply chain efficiency. The U.S. Customs Office launched several programmes such as the Customs-Trade Partnership Against Terrorism (C-TPAT), Container Security Initiative (CSI), the 24-hour rule, etc. These security programmes address different aspects of supply chain security and target different parts of a transport chain. The link between these security programmes is that they involve all parties or stakeholders in supply chain security [104]. The effects from these programs both in order to handle security threats and their impact on different logistics processes have been addressed in a few papers [105, 106, 107, 108, 109]. [110] states that the C-TPAT certification will probably have a negative impact, mostly on small enterprises whiles large firms instead may have the possibility to trade-off the security costs with benefits related to supply chain transparency. [111] demonstrates the economical and competitive advantages for large and small shippers becoming FAST-approved (Free And Secure Trade). The acquisition of the FAST status may provide shippers with faster trans-border operations and consequently a substantial advantage on the export market [96]. [112] emphasize that efficiency and security in supply chains are closely related to each others, since higher security may reduce Customs delays. The relationship between security, efficiency and custom activities is clearly found in the AEO-program.

Other types of security programmes existed before the attacks on the World Trade Centre. These programmes were designed primarily to address theft problems within the transport business (TAPA FSR and TSR, etc.) [113]. The big difference between security programmes before and after the terrorist attacks is that afterwards, authorities (mainly U.S.) took the lead in developing and implementing these programmes. Before September 11, 2001, security was something the business itself handled. The implementation of these programmes has so far mainly occurred in the old western countries in North America and in Europe.

Other types of security programmes existed before the attacks on the World Trade Centre. These programmes were designed primarily to address theft problems within the transport business. The big difference between security programmes before and after the terrorist attacks is that afterwards, authorities (mainly U.S.) took the lead in developing and implementing these programmes. Before September 11, 2001, security was something the business itself handled. The implementation of these programmes has so far mainly occurred in the old western countries in North America and in Europe.

The cost for implementing these programs alters the current collaboration models with regards to risk and profit sharing within the supply chain. The risks for antagonist threats are depended on the local environment. This shall be compared with that the security programs advocates one to three different security levels which shall solve the problem. This leads to that the security level is adjusted towards the security programs instead of the local threat. This may lead to that the security cost is higher than needed but a standardisation within the supply/transport chain in security may result in better collaboration. The political reasons behind the different supply chain security programs is most likely the fear for terrorist attacks which according to [32] depends more on the *proliferation of mass media* than development of new weapons. This is the context to understand the content and effects from fear of terrorist against international trade. Nevertheless the supply chain security programs provides collateral benefits like better product control, lesser shrinkage and better incident prevention by reducing threat opportunities.

The compliance with these different logistics security programs are based on different reasons. The compliance with business logistics security programs (like TAPA:s) are based on customer requirements. This depends on the simple fact that these programs are focusing on theft prevention. The governmental logistics security programs like AEO, C-TPAT, CSI etc. are on other hand more focusing on preventing terrorist activities. These programs normally also contains some kind of disadvantage for the own organisation if not compliance. Therefore are the governmental logistics security programs entail with a higher likelihood of compliance due to that need for compliance are not based on a risk assessment about the potential causes and impacts for antagonistic threats but on a general business assessment. For legal businesses, the AEO, C-TPAT, etc., are both a global supply chain headache and a business opportunity, depending on the risk for theft and counterfeiting for that company [114]. Irrespective of the difference in compliance reasons, the governmental logistics security programs may result in collateral benefits like lower cost for theft and better working conditions for the employees but there are still to be demonstrated that the terrorism preventing logistics security programs actually reduces the risk for non-economical-driven antagonistic threats because it's their primer reason for existing.

5. Conclusions

First, [1] clearly points out the effects of the WTC terrorist attacks on the global flow of goods. The effect may be indirect, but it was devastating. This event along with non-antagonistic events such as Hurricane Katrina and other natural disasters demonstrated their power to disrupt or cause uncertainty in supply chains. Second, terrorism conducts fund raising through criminal activities, which leads to that terrorism represents all antagonistic threats. Third, the tools and strategies for handling antagonistic threats are partly governmental (police and justice system) and partly consequence handling (insurance businesses, conventions and business contingency).

For terrorist threats, it is useful to refer to transferred and perceived threatening pictures [31]. The perceived or actual marketing strategies which link a company or organisation with certain values/states/advocates will affect the risks for terrorist threats to the organisation/business/market position. The key issue in this conclusion is that the perceived relationship between a company/organisation and values/states/advocates is made by the potential terrorist. Therefore is it important for a holistic threats assessment to include and consider these different relationships between an organisation and different values/states/advocates with regard to direct and more important indirect connections.

Historically, terrorist groups and organisations have been closely related to a certain geographical area and executed in that area. This is a valid statement for terrorist organisations such as IRA and the Basque separatist movement, even if both have conducted attacks outside their fighting area but within their targets (e.g., in UK but outside Northern Ireland, and in Spain but outside the Basque province). Those types of organisations only constitute a threat if the operations are carried out in that area and if the logistics operation possesses any real or symbolic value for the terrorist organisations. The strong relationship between different terrorist groups and geographical area may be reduced when referring to international terrorist organisations such as Al-Qaeda, but the targets should still possess a symbolic value for the terrorist organisations, even if the value is public fear.

The vulnerability of the supply chain is transmitted to the transport network. This depends on the simple fact that transports and freight activities physically bring the facilities of a supply chain together. Therefore, risks, uncertainties, and vulnerabilities in the supply chain and the transport network affect, contribute, and neutralize each other. Supply chain security is indented to safeguard the supply chain (in this meaning the transport and freight activities) from different antagonistic threats and thereby reduce the vulnerability of modern global trade.

Author details

Daniel Ekwall
School of Engineering, University of Borås, Borås, Sweden
Supply Chain Management and Corporate Geography, Hanken School of Economics, Helsinki, Finland

6. References

[1] Sheffi, Y. (2001), "Supply chain management under the threat of international terrorism". *International journal of logistics management*, Vol. 12, No. 2, pp. 1-11.

[2] Waters, D. (2007),"Supply chain risk management: vulnerability and resilience in logistics". Kogan-page, London

[3] Ekwall, D. (2009) *Managing the Risk for Antagonistic Threats against the Transport network*, Division of Logistics and Transportation, Chalmers University of Technology: Göteborg.

[4] Christopher, M (2005), *Logistics and Supply Chain Management – Creating Value-adding Networks*, Prentice Hall, London

[5] Bowersox, D.J. and Closs, D. J. and Cooper, M. B. (2002), *Supply Chain Logistics Management*, McGraw Hill/Irwin series, Boston (2nd Edition)

[6] Lumsden, K. (2006), *Logistikens grunder*. Studentlitteratur, Lund (*in Swedish*)

[7] Ruijgrok, C. and Wandel, S., and Nemoto, T. (1991). "Advanced Logistics and Road Freight Transport", *OECD Road Transport Research*, Chapter IV

[8] Wandel, S. and Ruijgrok, C. (1995), "Information technologies for the development of transport and logistics". *ATAS Bulletin information for development*, United Nations, New York.

[9] Ekwall, D. (2010), "On analyzing the official statistics for antagonistic threats against transports in EU: a supply chain risk perspective". *Journal of Transportation Security*, Vol. 3, No. 4, pp 213-230

[10] Francis, T. L. (1986), "A dynamic model of corruption deterrence. *Journal of Public Economics*, Vol. 31, pp 215-236

[11] Eurostat, 2009, *Criminal and criminal justice*. Statistics in focus 36/2009

[12] EUICS (2005), *The Burden of Crime in the EU, A Comparative Analysis of the European Survey of Crime and Safety*. (EU ICS).

[13] NRSS (2002), *National Retail Security Survey*. University of Florida

[14] ECR (2003), *Shrinkage: A Collaborative Approach to Reducing Stock Loss in the Supply Chain*. ECR Europe, Brussels, Belgium

[15] TAPA (2006), TAPA Loss Data Benchmark Survey 2006.

[16] Anderson, B. (2007), "Securing the Supply Chain – Prevent Cargo Theft". *Security*, No 5, Vol. 44, pp. 56-58

[17] EP - European Parliament's Committee on Transport and Tourism, (2007), *Organised theft of commercial vehicles and their loads in the European union*. European Parliament, Brussels

[18] ECMT (2001), *Theft of goods and goods vehicles*. CEMT/CM (2001)19, Lissabon.

[19] IRU (2008), Attacks on drivers of international heavy goods vehicles. INTERNATIONAL ROAD TRANSPORT UNION, GENEVA

[20] Dillén, J. and Ekwall, D. (2006), "Brott mot yrkestrafiken". Transek 2006:10 (in Swedish).

[21] Barth, S. and White, M. D. (1998), "Hazardous cargo". World Trade, November 1998, pp. 29,

[22] Tryon, G. and Kleiner, B. (1997), "How to investigate alleged employee theft properly". *Managerial auditing journal*, Vol. 12, MCB University Press.

[23] Robinson P. V. (2009), "Freight crime in Europe: what happens next?". *A presentation at ESCB 09*, Prague

[24] Muir, J. (1996), "Theft at work". *Work Study*; Vol. 45.

[25] Beck, A. (2002), "Automatic product identification & shrinkage: Scoping the potential". *ECR Europe*, Brussels, Belgium

[26] Europol, (2009), *Cargo theft report: Applying the Brakes to Road Cargo Crime in Europe*. Europol, The Hague

[27] Napoleoni, L (2004), *Terror inc*. Penguin books ltd.

[28] Europol (2010), "EU Terrorism Situation and Trend Report 2010". The Hague, Netherlands

[29] Johnston, R. B. and Nedelescu, O. M. (2006), "The impact of terrorism on financial markets". *Journal of Financial Crime*, Vol. 13, No. 1, pp. 7-25

[30] Rystad, G. (2006), *Politiska mord – det yttersta argumentet*. Historiska media, Lund (in Swedish)

[31] Sjöberg, L. (2008), "Antagonism, trust and perceived risk". *Risk Management*, 10, 32 – 55. doi: 10.1057/palgrave.rm.8250039

[32] Gearson, J. (2002), "The nature of modern terrorism". *The political quarterly publishing*, pp. 7-24

[33] Burke, R. J. (2005), "International terrorism and threats to security: Implications for organizations and management". *Disaster Prevention and Management*, Vol. 14, No. 5, pp. 639-643

[34] Aggarwal, R. and Bohinc, J. (2011), "Black swans and supply chain strategic necessity". *Journal of Transportation Security*, DOI 10.1007/s12198-011-0080-5

[35] Rodwell, S. and Van Eeckhout, P. and Reid, A. and Walendowski, J. (2007), *Study: Effects of counterfeiting on EU SMEs and a review of various public and private IPR enforcement initiatives and resources*. Framework contract B3/ENTR/04/093-FC-Lot 6 Specific agreement n°SI2.448309

[36] EU Commission (2008), *Public Consultation in preparation of a Legal Proposal to combat counterfeit medicines for Human Use - Key Ideas for better Protection of Patients against the risk of Counterfeit Medicines*, Brussels

[37] DEA, U.S. Drug Enforcement Administration, (www.dea.gov)

[38] Bouchard, M. (2007), "On the resilience of illegal drug markets". *Global crime*, Vol. 8, No. 4, pp. 325- 344

[39] Huang, J-H. and Lee, B. and Ho, S. (2003), "Consumer attitude toward gray market goods". *International Marketing Review*, Vol. 21 No. 6, pp. 598-614.

[40] EC, European Commission: Combating counterfeit & piracy, The economic consequences, (ec.europa.eu)

[41] Ekwall, D. (2009), “The risk for detection affects the logistics system setup for cargo smugglers”. *In proceedings of Nofoma 2009*, Jönköping.

[42] Naylor, R (2004), *Wages of crime.* Cornell University Press. Ithaca.

[43] Antonopoulos, G. A. (2007), “Cigarette smugglers: A note on four “unusual suspects””. *Global crime*, Vol. 8, No. 4, pp. 393-398

[44] Jannati, F. and Salimi, M. (2011) *Modern Sea piracy - Modus operandi and economical and development state backgrounds*. School of Engineering, University of Borås, Sweden

[45] Abeyratne R (2010) Managing the twenty-first century piracy threat. In: Thomas A (ed) Supply chain security, Vol. 1. ABC-CLIO, Santa Barbara

[46] Sullivan, A. (2010), “Piracy in the Horn of Africa and its effects on the global supply chain”. *Journal of Transportation Security*, Vol. 3, No. 4, pp. 231-243

[47] Onuoha. F, C, (2009), “Sea piracy and maritime security in the Horn of Africa: The Somali coast and Gulf of Aden in perspective”. African Security Studies, Vol. 18, No. 3, pp. 31-44

[48] Madden, M. (2009), "Trading the Shield of Sovereignty for the Scales of Justice: A Proposal for Reform of International Sea Piracy Laws". *University of San Francisco Maritime Law Journal*, Vol. 21, No. 2, pp. 139-166.

[49] IMO (2011) *Ship and Shipping Facts & Figures – Information Resources on Trade, Safety, Security, and Environment*. Maritime Knowledge Centre, IMO

[50] Stock, J. R. (1997), “Applying theories from other disciplines to logistics”. *International journal of physical distribution & logistics management*, Vol. 27, No. 9, pp. 515-539.

[51] Solem, O. (2003), ”Epistemology and logistics: A critical overview”. *Systemic practice and action research*, vol 16, no. 6, pp. 437-454

[52] Arlbjörn, J.S. and Halldorsson, A. (2002), ”Logistics knowledge creation: reflections on content, context and processes”. *International journal of physical distribution & logistics management*, Vol. 32, No. 1, pp. 22-40.

[53] Gubi, E., and Arlbjørn, J.S., and Johansen, J. (2003), “Doctoral dissertations in logistics and supply chain management. A review of Scandinavian contributions from 1990 to 2001”. *International Journal of Physical Distribution & Logistics Management*, Vol. 33, No. 10, pp. 854-885.

[54] Aastrup, J. and Halldórsson, A (2008), ”Epistelmological role of case studies in logistics: a critical realist perspective”. *International journal of physical distribution & logistics management*, Vol. 38, No. 10, pp. 746-763.

[55] Sarnecki, J. (2003), *Introduktion till kriminologi*. Studentlitteratur, Lund (*in Swedish*)

[56] Sherman, L.W. and Gartin, P. R. and Buerger, M. E. (1989), “Hot spots of predatory crime: routine activities and the criminology of place. *Criminology*, Vol. 27, No. 1, pp. 27-55

[57] Sampson, R. and Eck, J.E. and Dunham, J. (2010), “Super controllers and crime prevention: A routine activity explanation of crime prevention success and failure”. *Security Journal,* Vol. 23, No 1, pp. 37–51.

[58] Mankiw, N.G. (1997), *Principles of microeconomics*. The Dryden Press, Fort Worth Texas.

[59] Bodman, P. and Maultby, C. (1997), "Crime, punishment and deterrence in Australia". *International journal of social economics*, Vol 24. pp 884-901.

[60] Gigerenzer, G. and Reinhard, S. (2002), *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

[61] Simon, Herbert A. 1982. Models of Bounded Rationality. Cambridge, MA: MIT Press.

[62] Manunta, G. (1999), "What is security?". *Security journal*, Vol. 12, No. 3, pp 57-66, Perpetuity Press.

[63] Felson, M. and Clarke, R.V. (1998), "Opportunity makes the thief: Practical Theory for crime prevention". Home office police and reducing crime unit: London

[64] Kroneberg, C. and Heintze, I. and Mehlkop, G. (2010), "The interplay of moral norms and instrumental incentives in crime causation". *Criminology*, Vol. 48, No. 1, pp 259-294

[65] Bergman, M. (2007), Straff och nåd speglar samhället. *Svenska Dagbladet*, november 2 (*in Swedish*)

[66] Horney, J. and Marshall, I. H. (1992), "Risk perceptions among serious offenders: the role crime and punishment". *Criminology*, Vol. 30, No. 4, pp. 575-594.

[67] Badolato, E.V. (2000), "Smart moves against cargo theft". *Security Management*, No 7, Vol. 44, pp. 110-115

[68] Tarnef, B. (2006), "Combating Cargo Theft". *American Agent & Broker*, Vol. 78, No. 10.

[69] Williams, R. and Bertsch, B. and Dale, B. and van der Wiele, T and van Iwaarden, J and Smith, M. and Visser, R. (2006), "Quality and risk management: what are the key issues?". *The TQM Magazine*, Vol. 18, No. 1, pp. 67-86

[70] Felson, M. (1987), "Routine activities and crime prevention in developing metropolis". *Criminology*, Vol. 25, No. 4, pp. 911-932

[71] Ekwall, D. (2009), "The Displacement effect in cargo theft". *International Journal of Physical Distribution and Logistics Management*, Vol. 39, No. 1, pp. 47-62

[72] Purtell, D. and Rice, J. B. (2006), "Assessing cargo supply risk". *Security management*, November, pp.78-87, ASIS international.

[73] Urciuoli, L. (2008), *Security in Physical Distributions*. Department of Industrial Management and Logistics, Lund University: Lund

[74] Speed, M. (2003), "Reducing employee dishonesty: In search of the right strategy". *Security journal*, Vol. 16, No. 2, pp 31-48, Perpetuity Press.

[75] Clarke, R. V. (1995), "Situational crime prevention". In Tonry, M. and Farrington, D.P. (eds), *Building a safer society: strategic approaches to crime prevention*. Chicago: University of Chicago press.

[76] Lab S. P. (2000), *Crime prevention: Approaches, practices and evaluations*. Fourth edition. Andersson Publishing Co, Cincinnati US.

[77] Clarke, R. V. (1992), *Situational crime prevention: Successful case studies*. Harrow and Heston, New York.

[78] Clarke, R. V. and Homel, R. (1997), "A revised classification of situational crime prevention techniques". In, Lab, S. P. (ed.) *Crime prevention at a crossroads*. Andersson Publishing Co Cincinnati US.

[79] Tyska, L.A. and Fennelly, L.J. (1983), *Controlling cargo theft*. Woburn: Butterworth Publishers, Boston.

[80] Crawford, A. (1998), *Crime prevention and community safety: Politics, policies and practice*. Longman London.

[81] Borodzicz, E. P. (2005), *Risk, crisis & security management*. Wiley, Chichester

[82] Reppetto. T. A. (1976), "Crime prevention and the displacement phenomenon". *Crime and Delinquency*, Vol. 22, pp. 166-177.

[83] Hesseling, R. (1994), "Displacement: A review of the empirical literature". In, Clarke, R. (ed.) *Crime prevention studies,* Vol 3, pp. 197-230. Monsey, NY: Criminal Justice Press.

[84] Clarke, R. V. and Cornish, D. (1985), "Modelling offenders decisions: A framework for policy and research". In Tonry, M and Morris, N (ed.), *Crime and justice*, Vol. 4, Chicago: University of Chicago press.

[85] Clarke, R. V. and Cornish, D. (1990), "Crime specialization. Crime displacement and rational choice". In, Wegener, H. (eds), *Criminal behaviour and justice system*. New York: Springer-Verlag.

[86] Barr, R. and Pease, K. (1990), "Crime placement, displacement and deflection". In Tonry, M and Morris, N (ed.), *Crime and justice*, Vol. 12, Chicago: University of Chicago press.

[87] Manunta, G. (2002), "Risk and security: Are they compatible concepts?". *Security journal*, Vol. 15, No. 2, pp 43-55, Perpetuity Press.

[88] Clutterbuck, R. (1987), *Kidnap, hijack and extortion*. Basingstoke Macmillan, London

[89] Ekwall, D. and Nilsson, F. (2008), "Using business complexity to handle supply chain risk: Dealing with borders of cargo liability". *In proceedings of Nofoma 2008*, Helsinki

[90] Ekwall, D. and Nilsson, F. (2011), "Reallocation of risks within supply chains: The practice of enhanced liability clauses". *In proceedings of Nofoma 2011*, Harstadt

[91] Kelly, M. and Kleffner, A. E. (2003), "Optimal Loss Mitigation and Contract Design". *The Journal of Risk and Insurance*, Vol. 70, No. 1, 53-72

[92] Mears-Young, B. and Jackson, M. (1997), "Integrated logistics - call in the revolutionaries". *Omega,* Vol. 25, No. 6, pp. 605-618.

[93] Gentry, J. (1996), "Carrier involvement in buyer-supplier strategic partnerships". *International Journal of Physical Distribution & Logistics Management*, Vol. 26, No 3, pp.14-25.

[94] Lambert D. and Cooper, M. and Pagh, J. (1998), "Supply Chain Management: Implementation Issues and Research Opportunities". *The International Journal of Logistics Management*, Vol 9, No 2, pp.1-20

[95] Yu, Z. and Yan, H. and Cheng, E. (2001), "Benefits of information sharing with supply chain partnerships". *Industrial Management & Data Systems*, Vol. 101, No. 3, pp. 114-121.

[96] Cox, A. and Ireland, P. and Lonsdale, C. and Sanderson, J. and Watson, G. (1992), *Supply chains, markets and power: mapping buyer and supplier power regimes*. Routledge, London

[97] Cox, A. (2001): The Power Perspective in Procurement and Supply Management. *Journal of Supply Chain Management*, Vol. 37, nr. 2, ss. 4-7

[98] Closs, D. and McGarrell, E. (2004), "Enhancing Security Throughout the Supply Chain". IBM Centre for the business of government.

[99] Sweet, K. (2006), *Transportation and Cargo security*. Pearson Prentice Hall, New Jersey.

[100] Lee, H.L. and Whang, S. (2003), "Higher supply chain security with lower cost: Lessons from total quality management". *International journal of production economics*, 96, pp. 289–300.

[101] Manuj, I. and Mentzer, J.T. (2008), "Global supply chain risk management strategies". *International Journal of Physical Distribution & Logistics Management*, Vol. 38, No. 3, pp. 192-223

[102] Ritter, L. and Barrett, J. M. and Wilson, R. (2007), *Securing global transportation networks*. McGraw Hill, NY

[103] Sarathy, R. (2006), "Security and the global supply chain". *Transportion journal*, fall, pp.29-51

[104] Sheu, C. and Lee, L. and Niehoff, B. (2006), "A voluntary logistics security program and international supply chain partnership". *Supply Chain Management: An International Journal*, Vol. 11, No 4, pp.363–374.

[105] Fletcher, T. (2007), "Authorised economic operator (AEO) programs: IBM's perspective". *World Customs Journal*, Vol. 1, No. 2, pp 61-66

[106] Gutiérrez, X. and Hintsa, J. and Wieser, P. and Hameri, A-P. (2007), "Voluntary supply chain security program impact: an empirical study with basc member companies". *World Customs Journal*, Vol. 1, No. 2, pp 31-48

[107] Grainger, A. (2007), "Supply chain security: adding to a complex operational and institutional environment". *World Customs Journal*, Vol. 1, No. 2, pp 17-30

[108] Tweddle, D. (2007), "Logistics, security and compliance: The part to be played by authorized economic operators (AEOs) and data management". *World Customs Journal*, Vol. 1, No. 2, pp 101-105

[109] Mazeradi, A. and Ekwall, D. (2009)," Impacts of the ISPS code on port activities – A case study on Swedish ports". *World Review of Intermodal Transportation Research, Vol. 2, No. 4,* pp. 326-342

[110] Powanga, L. (2006), "A business perspective of US International Seaborne Security Measures: Impact on Importers". *Journal of Global Business*, Vol. 33, pp.63-75

[111] Haughton, M.A. (2007), "Examining the business case for shipper participation in Canada USA trade security programmes". *International Journal of Logistics Research and Applications*, Vol. 10, No. 4, pp.315-331

[112] Willys, H.H. and Ortiz, D.S. (2004). "Evaluating the Security of the Global Containerized Supply Chain," RAND Corporation, Santa Monica, CA.

[113] www.tapaemea.com

[114] Holmes, J. L. (2004), " The container security initiative: is container security a global challenge or a supply chain headache". *Fleet Equipment*, August 1

Chapter 9

Dynamic Risk Management Strategies with Communicating Objects in the Supply Chain of Chemical Substances Within the European Union

Omar Gaci, Hervé Mathieu, Jean-Pierre Deutsch and Laurent Gomez

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50138

1. Introduction

Supply chain is a set of activities involving a group of commercial actors to create a product or a service to satisfy a customer demand. The actors are the ones who form the supply chain, they are suppliers, transporters, manufacturers, distributors, retailers and customers.

In this chapter, our objective is to describe dangers for people, environment and goods from dangerous goods exploitation. We want to show how dangerous goods can be source of accidents during storage activities. To understand the potential of chaining event we elaborate accident scenario to describe the consequences of such accidents.

This risk study in warehouse provides insights to elaborate a risks management strategy relying on communicating objects such as RFID tags and wireless sensors. Then, by exploiting the technical features that provides these smart items, it becomes possible to detect in real time any accident risks and to react in consequence.

This chapter is organized as follows. Part 2 introduces briefly what supply chain management is and gives some definitions. Part 3 describes dangerous goods in logistics and notably the risks inherent to their storages. Current regulations such as CLP is presented, it provides a standard in the European Union to classify and identify dangerous goods. Risks relative to dangerous goods manipulation are presented, segregation strategies and storage constraints must be respected to maintain the security within a warehouse. In part 4, a risk study is lead to extract three scenarios that describe the domino effect consecutive to accidents. These scenario are treated further in part 5 when communicating objects are integrated at pallet level. Then, these smart items allow elaborating real-time risk assessment that contribute to detect early accident risks and to deploy emergency procedures to mitigate that risks.

This work is funded by German BMBF and French ANR as part of ReSCUeIT project.

2. Supply chain management in logistics

Depending on actors involved in Supply Chain (SC) (suppliers, transporters, customers), SC definitions can differ regarding the described interactions and dependences between actors. A definition is given by [1]: *"A network of organizations that are involved, through upstream to downstream links, in the different processes and activities that produce value in the form of products and services in the hands of the ultimate customer".*

With this definition, it appears that each company owns an internal SC and participates to another SC. A SC is a succession of activities related to a specific function performed by one or several actors in order to satisfy customer demands. Then, each actor belongs to a global flow, he evolves according to three interactions: his internal tasks, his provider status and his customer status.

2.1. Supply chain management

A SC exists if at least two companies work together to the production of a particular product or service. If this association is explicitly guided to improve the performance then we describe this as supply chain management, also called SCM.

From a general point of view, the SCM can be defined as the coordination between companies internal and external activities. The goal of this management is to improve the SC performance on long term basis so that each actor of the SC can take benefit of this global management.

The SCM consists of managing the whole organizations involved in the delivery of a final product or service. Its aim is, on one hand, to produce products relying on information received from the customer needs and, on the other hand, to minimize the different supply, production, warehousing and delivery costs. The SCM gathers two parts: the integration of the company along the SC and the coordination between the physical, information and financial flows.

The main objective of the SCM is to improve the competitiveness of companies by minimizing the costs while the quality of service required by the customer is guaranteed.

3. Dangerous goods management in logistics

A good is considered as dangerous when it may present a danger on population, environment or on infrastructure according to its physicochemical properties or because of the reactions it can imply.

Activities involving dangerous goods concern all parts of the world. A global regulation is needed and should be coordinated with local authorities to make laws more reliable and respected. Since the logistics tend to be global, the control should be also global for a better efficiency.

Laws related to the use, the loading, the unloading, the storage, the transportation and the handling of dangerous material differ according to activities, status or modality of transports. Countries rely on international recommendations proposed by the Organization of United Nations to regulate the operation on dangerous goods.

The UN proposes recommendations for the dangerous goods since 1957 and updates periodically its texts. It is a reference which provides the main recommendations related to the dangerous goods notably about the different methods of transport: air, road, railway, canals and sea. A specialized authority is dedicated for each means of transportation, they are the followings:

- International Air Transport Association, IATA, is for transport by airplane;
- European Agreement concerning the international carriage of Dangerous goods by Road, ADR, manages the European transport of dangerous goods by road;
- Regulation concerning the International transport by railway of Dangerous goods, RID, is for the international transport of dangerous goods by railway;
- European Agreement concerning the International Carriage of Dangerous Goods by Inland Waterways, ADN, manages the international transport on internal canals;
- International Maritime Dangerous Goods Code, IMDG Code, is for international transportation by water.

Due to the existence of international regulations, countries or groups of countries (the European Union for example) have adapted their laws to harmonize them with the UN model. Thus, country legislations tend to follow the international regulations to make the management of dangerous goods more standard and more visible by a third.

3.1. Dangerous goods identification

Considering the important number of substances, there is a clear need for dangerous goods classification. Existing classifications of dangerous goods are based on chemical families (acid, alcohol, amide, etc.), chemical reactions (oxidation, reduction, combustion) or also on chemical compositions.

The CLP (Classification, Labelling, Packaging) regulation is relative to the chemical substances imported or commercialized in the European Union. This regulation entered into force in January 2009 and will be totally applied in 2015.

3.1.1. Obligations under CLP

CLP provides a global obligation for all suppliers in the supply chain to cooperate. This cooperation is necessary to make the different suppliers meet the requirements for classification, labelling and packaging.

3.1.2. Terminology

A new terminology is used, terms of existing regulation are kept whereas news are adopted. The term substance is used to designed hazardous material and the transformation of these substances into a new one is called mixture.

As well, the properties of substances are described according to three properties: physicocochemical, toxicological and ecotoxicological. According to these three criterion, the definition of hazard classes helps to classify a substance. Then, a hazard class defines the nature of a hazard, it can be physical, on health or on the environment.

3.1.3. Classification of substances

CLP possesses specific criteria of classifications that are rules that allow associating a substance to a class of hazard or a category in this class. In particular, the classification process is based on the substance concentrations to establish the effects of those substances on the health and the environment.

CLP defines three hazard classes and 28 categories, such as:

- 16 categories for physical hazards;
- 10 categories for health hazards;
- 2 categories for environmental hazards.

For example, the physical hazards regroup explosives, flammable gases, solids, aerosols, liquids. The health hazards are relative to acute toxicity, skin corrosion, irritation and sensitization. The environmental hazards address hazardous to the aquatic environment and hazardous to the ozone layer.

3.1.4. Labelling

A substance contained in packaging should be labelled according to the CLP rules with the following information (called labelling elements):

- the name, address and telephone number of the supplier of the substance;
- the quantity of the substance in the packages;
- hazard pictograms;
- signal word;
- hazard statements;
- appropriate precautionary statements;
- supplemental information.

A substance contained in packaging is labelled according to the CLP rules and contains a set of information such as name of the supplier of the substance, quantity of the substance in the packages or hazard pictograms, see Fig. 1.

The CLP regulation helps then the identification of chemical substances through the supply chain since it provides a standard framework for the classification, the labelling and the packaging of substances.

3.2. Risks in dangerous goods storage

In the context of logistics the main considered risks are relative to the physical or information flows disruption. The case of a supply chain whose physical flow manages dangerous goods, the risk becomes different and takes another dimension.

Figure 1. Pictograms used in CLP regulation.

3.2.1. Risk definition

The Process Safety Management (PSM) defines risk as follows:

"Risk is defined as a measure of frequency and severity of harm due to a hazard. (...) In the context of public safety, risk is commonly characterized by fatalities (and injury) to members of the public".

According to this definition, several factors make difficult to assess the risk for dangerous goods manipulation. Among these factors, there are the followings:

- Because of the diversity of hazards, the chemicals physicochemical properties are different and consequently risks evolve also;
- The localization of potential accidents stays uncertain, it is not possible to determine where an accident can occur (warehouse, highways, county roads, local roads, etc.);
- The large diversity of causes implies that it is impossible to enumerate all cases and to treat these risks.

3.2.2. Risks in warehousing

Among dangerous goods, products can react when they are in contact. For these reasons, they must be stored in separate places. Strategy of storage consists in avoiding that incompatible products are neighbors. To this end, a first step is needed to identify substances as a function of their potential chemical reactions.

In order to prevent any storage of incompatible chemicals and risk of chemical reaction in case of wrong handling, seggregation strategies are used. As shown in Fig. 2, it exists

Danger Code	F	F+	T	Xi	O	Xn	N	C
F	+	+	-	+	-	+	-	-
F+	+	+	-	+	-	+	-	-
T	-	-	+	+	-	+	-	-
Xi	+	+	+	+	-	+	+	-
O	-	-	-	-	+	-	-	-
Xn	+	+	+	+	-	+	-	-
N	-	-	-	+	-	-	+	-
C	-	-	-	-	-	-	-	+

Figure 2. Identification of storage compatibilities between dangerous goods. The letter F means inflammable, F+ means very inflammable, T means toxic, Xi means very irritant, means O oxidizing, Xn means noxious, N means polluting and C means corrosive.

incompatibilities for storage that is why the maintain of segregation between products constitute a risk assessment strategy to mitigate the risk of chemical reactions.

Once dangerous goods are stored in separate places of a warehouse, other sources of risks remain:

- Container falls;
- Container damages;
- Storage conditions (humidity, heat, cold, light);

4. Risk study in warehouse by scenarios

Within the framework of scenarios modeling, three different scenarios are built, including causes and effects of accident risks that define a basis for further risk assessment by communicating objects.

The methodology used is the one usually practised in audit or studies of risks and dangers and also in the impact procedures.

4.1. Methodology

The methodology used for scenarios definition aims at characterizing, analyzing, and assessing the risks in warehouse. Those risks can have two different causes: due to logistic exploitation (e.g., shipping, handling, storage), or due to external causes such as flooding, lightning.

This first analysis, find all the regulatory constraints aims to identifying and at recording, through various tools, the potential hazards that may occur in the retailer warehouse and its operating system. The natural and human environment hazards are also analyzed. These steps are intended to identify external attackers at the retailer warehouse.

4.1.1. Risk evaluation

An analysis of the warehouse neighborhood is conducted in order to identify targets and their vulnerabilities and to characterize the severity of the dangerous phenomena. Equipments that may be impacted (eg. racks, clarks, trucks) due to internal or external logistics processes are also identified.

4.1.2. Preliminary risk analysis

A rating of hazardous phenomena identified enable to identify major accidents and through a combination of different criteria:

- Probability of occurrence of a dangerous phenomenon;
- Intensity of these effects;
- Vulnerability of the target impacted by these effects.

Following this preliminary risk evaluation, we can discard non significant risk, which are evaluated as low gravity, and probability. Therefore, we perform a detail analysis over the major identified risks.

This rating corresponds to a hierarchy through a matrix (gravity/probability) allow identifying two categories of risk:

- risk level is considered generally sufficient;
- risk has to be subject to a closer examination.

4.2. Scenario 1: Aerosol explosion

The warehouse is likely to receive products classified under the heading 1412 UN code, bottles generating aerosols. The propellant gases contained in these bottles is most of the time of butane or propane under pressure.

In this scenario, we describe the explosion of aerosol (air freshener). It is considered as an inflammable.

4.2.1. Risk evaluation

The principle of a generating bottle of aerosols is to allow the propulsion of the product (lacquer, deodorant, adhesive, maintenance product, ...) out of the bottle thanks to a gas under pressure contained in the bottle. Common bottles have a varying volume from 50 to 500 ml and contain between 30 and 150 G of product plus propellant gas.

The behavior in fire of aerosols generators of depends on one hand nature on propellant gas and on the other hand on the nature of the conditioned liquid. Therefore, if the propellant gas is a flammable gas standard butane or propane, there is a risk of explosion. The bottles are then dispersed by missile effect and can be in their turn a propagator of fire.

In the same way, if the liquids contained are flammable liquids, they will support the fire. The generators of aerosols are thus the subject to specific storage, depending on the nature of the containing products (gas and/or liquid).

4.2.2. Preliminary risk analysis

Based on the calculating probability scale and the calculation of gravity scale, this scenario occurred and may occur throughout the life of the installations and no person will be impacted.

The method used to assess the effects of an explosion of aerosols in a warehouse is multi-energy model. Explosion cloud formed by the gas contained in a bottle would not reach the thresholds of overpressure. Regardless of the distance to the cloud, a target suffers less than 20 mbar pressure which causes no human and material damage. We consider the explosion of the cloud formed by the simultaneous outbreak of 100 bottles of aerosols, which is extremely unlikely even when taking into account that the bottles are for most made on aluminium.

4.3. Scenario 2: Fire truck incident

This scenario develops the consequences of the fire on a truck loaded with pallets, we assume that the truck is loaded exclusively with generic aerosol or fertilizer.

Nitrate-based fertilizers (that do not contain high nitrate content) are not capable of detonation. The ammo-nitrates high or medium dosage, French standards or European are difficult to detonate in the absence of contamination. Furthermore, dust fertilizer containing ammonium nitrate in the air are not combustible and do not present an explosion risk, unlike the dust of grain or organic fuels.

4.3.1. Risk evaluation

The detonation is possible in some cases:

- Contamination of the nitrate high or medium dosage of substances fuel (fuel, oil, plant protection products) or incompatible;
- heating of the nitrate high or medium dosage especially if it is contaminated, and maintenance of a containment of the gases emitted by the combustion;
- Severe impact of projectile on the product or shock very violent in contact the nitrate.

4.3.2. Preliminary risk analysis

Based on the calculation probability scale and the calculation of gravity scale, this scenario event thought: occurred and may occur throughout the life of the installations or may unlikely event: something similar already encountered in the sector of activity or this type of organization at the global level, without any possible corrections made since bring a significant reduction in its probability is guarantee and no person will be impacted.

The following scenario develops the consequences of a fire truck loaded with pallets at a unloading dock site. The fire or explosion of aerosols and fertilizer is due to inflammation of a palette (mechanical heating, inflammation, engine, non-compliance with the prohibition of smoking). We left the assumption that a cigarette would be has the origin of the starting point of fire. We leave the principles which another vehicle charged with manure would be been parked on the quay dimensioned. The starting point of fire and the propagation with the truck containing of manure cause effects missile, explosions, and dangerous heat fluxes for the man as well as the toxic dispersion of smoke, of projection has missile effect as well as discharge of water and product polluting for environment.

4.4. Scenario 3: Hazardous liquid spill

Accidental discharge is related to the presence of liquid products on site. The spill may trigger the leakage of an important quantity of hazardous liquid. A leakage of hazardous substances can occur due to the fall of one or more pallets of a rack but also due to a weakness in on (or a series) of container (bottle, can). A spill of hazardous substances can have an impact on the environment and on people.

4.4.1. Risk evaluation

Because of the loss of integrity of a container, a liquid will flow and form a pool of dangerous liquid which mainly function of the topography, the viscosity of the product and quantity.

The multitude of references and stored products do not enable to study exhaustively all dangerous chemical reactions. However, all possible precautions are taken to avoid storing incompatible products in a single cell.

4.4.2. Preliminary risk analysis

This scenario occurred and may occur throughout the life of the installations and no person will be impacted. We will find a fall of pallet that will provoke liquid spill which is the starting point of fire, projection missiles effect, various explosions and propagation of smoked and gas dangerous for the man and the environment.

According to the nature of fire, it must there have a respect of the means of extinction, in this case of figure it will be impossible to respect this standard, and the means of extinction set up, will cause a pollution of the natural environment.

4.5. Scenarios modelling

In this section, we model the scenarios presented previously by UML diagrams to describe the sequential chain of events following an accident. In particular, we are interested in identifying the main damages and consequences of accidents involving dangerous goods.

As we said, the accidents are evaluated by several measures, by the place and by the number of persons within this place. The type of accidents we consider and which can occur during the manipulation of dangerous goods is correlated to the chemical features of those goods. Further, the main effects we consider in our scenarios are the followings:

- Release of substances which are toxic to health and to the environment;

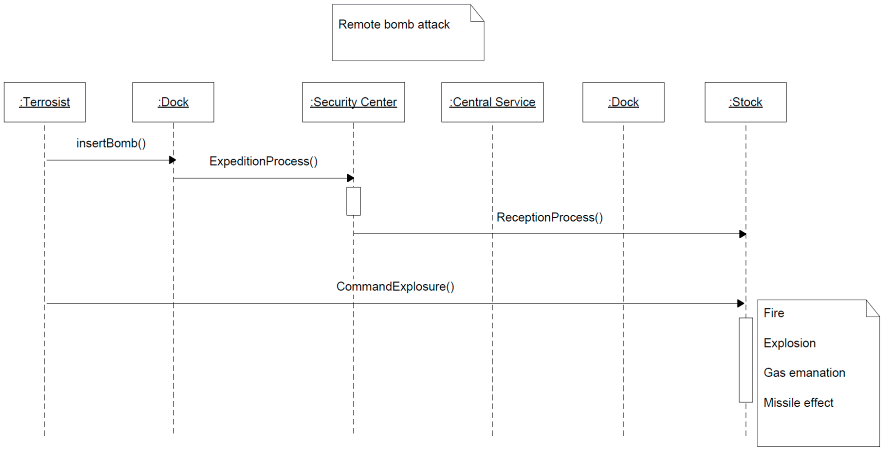


Figure 3. Scenario 1: a bomb attack triggered from a stored pallet.

- Release of thermal energy;
- Release of pressure.

The consequences that derive from an accident depend obviously on the type of dangerous goods, the area population, the accident time, the accident place, etc. Here, we are interested in describing the impacts of accidents by scenarios. In particular, we study the consequences of energy release according to the two phenomenons of Unconfined Vapor Cloud Explosion (UVCE) with the first scenario and the Boiling Liquid Expanding Vapour Explosion (BLEVE) with the second scenario.

4.5.1. Scenario 1: Aerosol explosion

Based on the scenario introduced as Scenario 1, we describe how a bomb attack can occur in a warehouse. We propose a first sequence diagram; see Fig. 3, to model this scenario.

We remark that the bomb is transported from a first place to the final storage place where the bomb attack is triggered. Nevertheless, by considering that terrorism attacks have a probability to miss their targets, we can assume that a remote bomb attack can potentially become a road accident when the bomb explodes at an untimely moment.

We describe by an activity diagram, see Fig. 4, the events following a bomb explosion once it has been stored. The first events are relative to the way that the bomb is susceptible to explode when the trigger is activated; it can burn first before exploding. Once the explosion happens, the shock wave provokes damages on near products and a domino effect starts from the pallet to the rack beams. With time the whole pallets are susceptible to burn and then to explode. The fire and explosions provoke pressure effects and thermal effects which have consequences on persons present in the stock. The damages caused by this bomb attack are measurable by three dimensions: the human, the structure and the merchandises damages. Here the humans damages are relative to the persons working in the warehouse and who

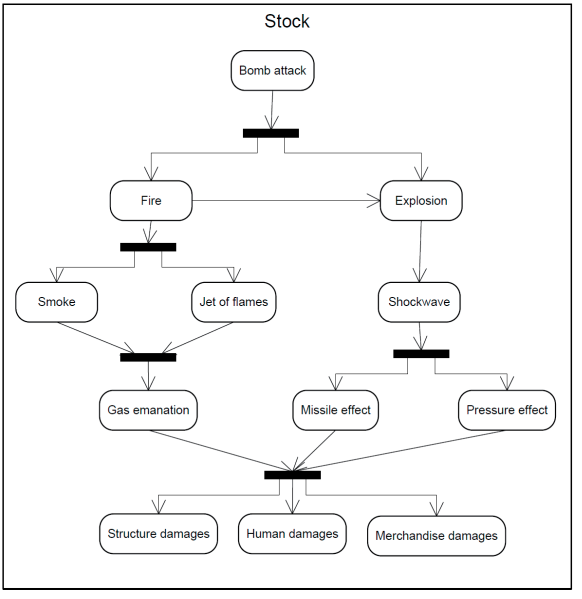


Figure 4. Scenario 1: activity diagram for a bomb attack.

have: inhaled gases, been victims of missile effects or building collapses, been burnt. The structure damages regroup the building destructions notably the racks, the traverses and the building structure. The merchandise destructions are easily evaluable and correspond to the pallets implied during the fires and the explosions.

This first scenario describes actually the consequences of a Boiling Liquid Expanding Vapor Explosion, BLEVE. This scenario describes the chain reaction explosion generated from the expansion of inflammable vapors produced by gas substances (previously kept under pressure in a liquid state). The effects that derive from this type of explosion are effects of excess pressure and fire balls projections and provoke damages on people, structures and goods. This type of event entails three main dangers: the shockwave from the explosion, the thermal flow from the fire and the projectiles from the damaged goods.

4.5.2. Scenario 2: Fire truck incident

This scenario involve a set of actors during the loading or the unloading of a truck. Then, the transporter is present and follows the operations while the driver handles pallets between the dock and the stock. This scenario is modeled by a sequence diagram in Fig. 5. The accident occurs in the dock after the driver knocks its cargo against another one on the dock.

We describe by an activity diagram, see Fig. 6, the chain of events from this accident. As we see, the collision is the trigger of this scenario and occurs on the dock. The presence of a cigarette increases strongly the accident consequences. The shock provokes the release of gas which concentrates to form a cloud around the dock and the truck. The cigarette provides

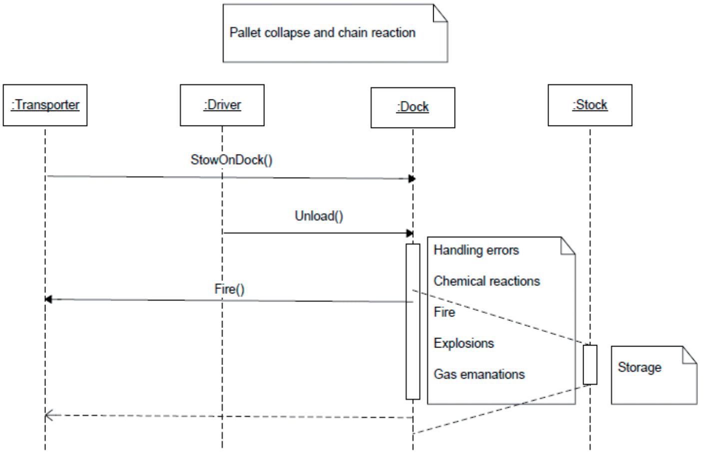


Figure 5. Scenario 2: a pallet collapse and a chain reaction on dock.

the spark so that uncontrollable events start. In presence of the cigarette, the released gas provoke thermal and pressure effects. The damages caused by this accident spread on four dimensions: the human, the structure, the merchandises and the environment damages. The human damages are relative to the person evolving around the accident place. Among them, we can cite the transporter, the driver, the person who works near the dock and near the stock. The structure damages implies the dock itself which can be deformed by the explosion, the truck which can burn and also the warehouse building which can suffer from the explosion. The merchandises destroyed are the ones present on the dock or the ones which burnt during the fire. This type of accident provokes the formation of toxic gas clouds which are polluting the environment. Before and after the cloud explodes the mixed gases stay in suspension and moves with the ambient air. It constitutes then environment pollution.

4.5.3. Scenario 3: Hazardous liquid spilt

The third scenario describes how a rack failure is susceptible to involve a collapse of pallets and then provoke an accident. The pallets fall over each others contributing to gas emanations and liquid spreads, see Fig. 7.

We propose an activity diagram to describe the actions, see Fig. 8, leading to an accident. Here, we don't use any aggravating facts and we consider that the gas emanation and liquid spreads stay stable and don't provoke yet any explosions. Nevertheless, this unstable situation represents a danger for different reasons. Thus, the gas emanations are toxic for the persons working in the warehouse and the different pallet collapses and rack damages constitute structure and goods damages.

5. Integration of communicating objects for a dynamic risk management

Over the past few years, communicating objects have become a new emergent solution to secure the international trade by providing tracking tools [2–4] and environmental monitoring

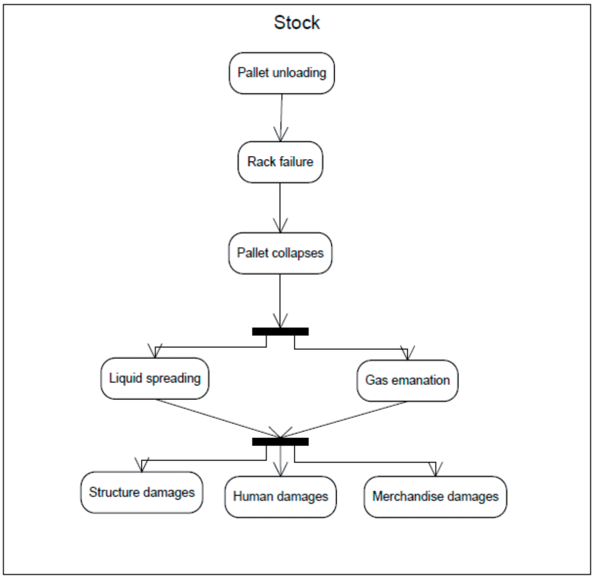


Figure 6. Scenario 2: activity diagram for a pallet collision on dock.

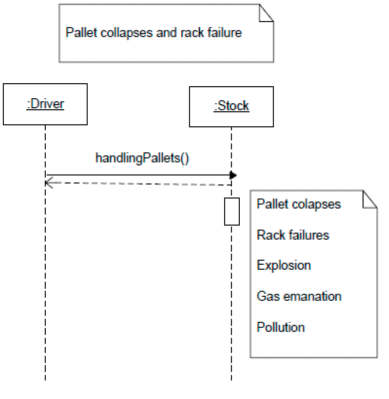


Figure 7. Scenario 3: rack failure and chain reaction.

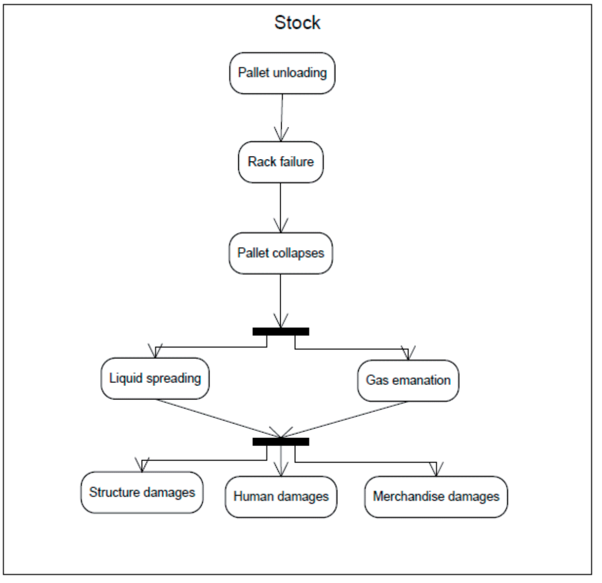


Figure 8. Scenario 3: activity diagram for a rack failure.

[5, 6]. Then, tracking tools provide geolocation solutions that can be used to detect any risks inherent to goods location. environmental geolocation gives the possibility of detecting any "unusual" or "dangerous" environmental conditions such as high temperature or incorrect constraints.

In this chapter, a SCM approach that exploits RFID tags and wireless sensors is presented. The studied supply chain manipulates chemical substances that represent potentially hazard for persons and environment. The developed approach integrates constraints from existing regulations and complies with them to finally propose a dynamic risk assessment.

5.1. Supply chain actors and activities

The studied supply chain involve different actors that manipulate chemical substances:

- Chemicals providers are located in Asia. They prepare pallets by loading chemicals on communicating pallets and organize the container loadings in the Shanghai harbour;
- Transporter by road or by ship, transport containers and deliver them in Le Havre port or to the retailer warehouse;
- Retailer in France for big-box stores. This actor organizes the pallet storages. Operators unload containers to store pallets in racks, load pallets in containers and finally send them to the downstream supply chain actor.
- Customers, they are big-box stores around Paris. They receive containers, unload them and transfer chemicals from communicating pallets into common ones and send back them to the pallet-provider.

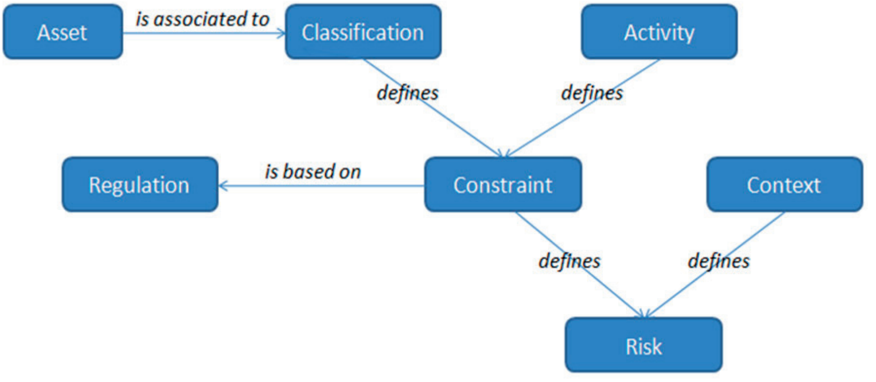


Figure 9. Risk assessment description. The risk assessment is based on classification and activities that define constraints to be respected along the supply chain.

The communicating objects are integrated at the pallet level. The cost is shared by the different actors and the programming of communicating objects is realized by an upstream actor: the pallet-provider. This last is responsible of writing information in RFID tags and initializing wireless sensors.

5.2. RFID technologies

A basic RFID system contains three elements; a tag, a reader and a middleware. A tag is mainly formed by a microchip attached to an antenna. An RFID tag is read when it receives radio signals from the reader and sends data back to the reader. The reading and writing process between the tag reader and the RFID tag can be realized from centimeters up to meters depending on the system characteristics.

5.3. WSNs: Wireless Sensor Networks

Wireless Sensor Networks (WSANs) refer to a group of sensors linked by a wireless medium to carry out phenomenon sensing and acting in consequence. The devices deployed in the environment are considered as nodes. Nodes are sensors whose main features are: low-space, low-cost and low-power, they are able to collect environmental data.

5.4. Risk assessment with communicating objects

The proposed risk assessment is based on the chemicals physicochemical properties. Depending on these properties, the substances are classified to a specific class of hazard according to the CLP regulation. From this classification, constraints are deduced and must be respected all along the supply chain. Then, chemicals evolve in an environment that defines their "context". The combination of the identified constraints and the environment of chemicals are the source of risks, see Fig. 9. The deployment of communicating objects for a dynamic risk assessment allows monitoring in real-time pallet environments.

Different sensors are associated to form a cluster of sensors able to monitor a set of constraints. Sensors send data to a centralized software that translates them into accident risks in case of unusual values. By this way, the software is able to send alert messages to supply actors that are currently responsible of goods so that they can intervene on goods.

The accident risks treated by the software are associated to sensors, they are the following:

- Overheating (temperature sensor);
- Pallet squashing (pressure sensor);
- Pallet overturn (stability sensor);
- Pallet lost (gps sensor);
- Gas release (gas sensor);
- Fluid leak (fluid sensor);
- Solid product dumping;
- Humidity constraint (humidity sensor);
- Incompatible product (RFID tag).

5.5. Scenarios: Explosion and chemical reactions on stow

In the first scenario, the events occur at dock where the truck is stowed and a forklift driver is unloading pallet and put them on the dock before storing them further on rack. The dangerous goods considered are aerosols and are commercialized as air fresheners.

Sensors present on pallets send periodically their data to the software. A terrorist triggers a bomb explosion (remote activation by mobile phone for example) from a pallet that is still on the truck cargo. Different events occur from this explosion: thermal and smoke emanation, release of toxic gases and missile effects. Pallets in the container and on the dock will be damaged by the blast wave, the main consequence on goods is the loss of aerosols containment. Consequences are expected on environment, on warehouse workers and on warehouse building structure.

As soon as the explosion occurred, sensors present on pallets that suffer from the bomb consequences will immediately be damaged and will not emit anymore whereas sensors present on dock will emit only few seconds before being destroyed by fire. The software that collects data will receive unusual values such as increasing heat and gas pressure and in consequence will sends an alert message to the warehouse manager. This last is responsible of the security maintain and will understand that an accident occurred on a dock. He will then deploy emergency procedures and will be helped by the software that provide strategic information with the alert messages. The software will send information about the pallet whose sensors are broken, the CLP designation of goods present on these pallets will be sent. By this way, the retailer manager will know exactly what the involved goods are, he will adapt emergency procedures as a function of physicochemical properties of goods.

In the second scenario, the context and the assumptions are the same than the first one: the action takes place on the stow where the truck is being unloaded by an operator. Goods considered are housekeeping products, they are flammable liquids. An handling error such as a pallet is poorly loaded and taken by the forklift driver with a slight offset provokes a collision between the handled pallet and others already unloaded on the dock. This collision provokes a liquid leakage on the dock and a trigger event like a spark or an electrical arc created by a cell phone provokes a fire.

This accident is treated by the same way than the scenario 1. The pallet collision and collapse on dock correspond to unusual stability values. The software sends an alert message to the warehouse manager who will be able to locate the place where the accident occurred and will coordinate the emergency services relying on qualitative and quantitative information about chemicals involved in the accident.

5.6. Scenarios: Pallet collapse and domino effect in stock

The scenario 3 takes place in the warehouse storage area and involves operator, pallets and racks. During operator movements, we assume that the rack structure can be damaged by forklift collisions and the accumulations of such shocks can potentially provoke rack damage that can lead to the rack destruction due to the supported pallet weights.

The detection of a collision between a forklift and a rack can be detected relying on the stability sensor values transmitted by sensors. If more than four pallet have their stability values modified between 20% and 40%, the software translates this situation into "rack stability" alert and sends it to the retailer manager. From this alert, the retailer manager will trigger emergency procedures to check the involved rack. Then, the software provides different information relative to pallets whose stabilities have evolved, these information are exploited by the retailer manager to locate them using his internal WMS (Warehouse Management Software). Then, if the pallets whose stabilities have evolved are neighbours and located near the rack structure, an internal alert of "rack collision" is emitted and the forklift drivers that operate in the cell are convoked for a further debriefing.

In the scenario 4, we assume that forklift drivers move in the stock area when a forklift loses its cargo that fails. Stability values transmitted by sensors of this pallet will trigger a "pallet fail" by the software. The software provide also gps location of the involved pallet but this information cannot be exploited by the warehouse manager because none mapping exists between the geolocation and the location in the storage area. For that reason, an internal alarm is triggered in the corresponding cell to inform forklift drivers about a pallet fallen but none other accurate information are available.

6. Conclusion

In the global context of supply chain management, the three common flows of goods, information and physical are subject to modelling regarding optimization concerns. Risk studies allow identifying logistics missions that may represent danger on persons, environment and goods. Scenarios are then extracted to describe risk management strategies and then improving emergency procedures.

In this chapter, three scenarios are developed to answer specific events, located in quite precise places and representing real risks. The chosen scenarios presented three types of accidents:

- the creation of an explosive area resulting from a leak of aerosol;

- a departure of fire from a truck stowed;
- a breakdown of racks causing a toxic mix dangerous goods;

In the presented scenarios, only the product concept and application of procedures and risk prevention differ. It is the analysis of risks, its causes and its impacts which will enable to set up one of the mechanisms of detection and prevention.

The integration of communicating objects such as RFID tags and wireless sensors offer new features for a dynamic risk assessment. Sensors are then used for accident risk detections, collected data are transmitted to a centralized software that is able to compare them with confidence intervals and in case of mismatches alerts are sent. Then, communicating objects are used to risk detection and also to risk mitigation. Risk mitigation consists in providing strategic information (involved goods, their quality and quantities, accident place, etc.) to emergency services that can then plan and adapt their intervention.

Author details

Omar Gaci and Hervé Mathieu
ISEL, France

Jean-Pierre Deutsch
LogPro Conseil, France

Laurent Gomez
SAP, France

7. References

[1] Christopher, M. (1998). *Logistics and Supply Chain Management: Strategies for Reducing Costs and Improving Services*, Editions Pitman Publishing, ISBN 978-0273630494, London, UK.

[2] Huanjia, Y.; Yang, L.; Yang, S-H. (2011). Hybrid Zigbee RFID sensor network for humanitarian logistics centre management. *Journal of Network and Computer Applications*, Vol. 34, No. 3, 938–948.

[3] Lin, L.C. (2009). An integrated framework for the development of radio frequency identification technology in the logistics and supply chain management. *Computers & Industrial Engineering*, Vol. 57, No. 3, 832–842.

[4] Hong, I-H.; Dang, J-F.; Tsai, Y-H.; Liu, C-S.; Lee, W-T.; Wang, M-L.; Chen, P-C. (2011). An RFID application in the food supply chain: A case study of convenience stores in Taiwan. *Journal of Food Engineering*, Vol. 106, No. 2, 119–126.

[5] Shin, T-H.; Chin, S.; Yoon, S-W.; Kwon, S-W. (2011). A service-oriented integrated information framework for RFID/WSN-based intelligent construction supply chain management. *Automation in Construction*, Vol. 20, No. 6, 706–715.

[6] Marucheck, A; Greis, N.; Mena, C.; Cai, L. (2011). Product safety and security in the global supply chain: Issues, challenges and research opportunities. *Journal of Operations Management*, Vol. 29, No. 7, 707–720.

Section 3

Enterprise Risk Management

Chapter 10

IA OM® as an Enterprise Risk Management Metric

David R. Comings and Wendy W. Ting

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50880

1. Introduction

Ting and Comings [1] described how to use the Information Assurance (IA) Object Measurement (OM®) metric as a tool to measure the monitoring step (Step 6) described in the United States (U.S.) National Institute of Standards and Technology's (NIST) Risk Management Framework (RMF)[1] [2]. This chapter expands the applicability of the IA OM® metric and shows how it may be used as an enterprise-wide information security risk management metric.

Risk management is concerned with the identification of risks, the avoidance, mitigation, transference, or sharing of unacceptable risks, and the acceptance of risks that are within an organization's risk tolerance. However, just as with information system controls within NIST's RMF, it is necessary to monitor the risk posture of systems, maintaining an ongoing assessment of the level of risk they represent within and to an organization. This risk posture changes with changes to the hardware and software employed by the organization, as well as when patches and updates are released that are intended to be applied to deployed software. Changes can also occur from vulnerabilities identified with no patch available, or when new types of information are allowed on a previously authorized or accredited information system. Different types of information are of varying interest to an organization or adversary. More valuable information generally has a higher impact on the organization when it is compromised[2], and can increase the threat level of an information system. From an information system perspective, many of the monitoring activities, conducted to ensure the systems remain operational and maintain an acceptable security posture, are also activities involved in the management of information system risks.

[1] The RMF is described in detail in NIST Special Publication (SP) 800-37, available from: http://csrc.nist.gov/publications/nistpubs/800-37-rev1/sp800-37-rev1-final.pdf

[2] Compromise is used in this chapter to indicate a loss of confidentiality, integrity, or availability of the information.

The IA OM® metric is a good choice for use as an enterprise risk management metric, as described in this chapter, due to its versatility as a management metric. This metric:

- Measures information security risk management activities within an organization;
- Shows organizational senior management where their organization currently stands with respect to its risk management strategy, and its monitoring plan; and
- Demonstrates to senior management how such metrics can be used over time to track and improve their organization's ability to meet its overall risk management strategy and risk monitoring plan.

2. Risk management

Risk management focuses on understanding and managing risks to an organization. This chapter focuses on information security (also referred to as information assurance (IA)) risks. Wheeler [3] in his book on risk management stated that "The goal of risk management is to maximize the output of the organization (in terms of services, products, revenue, [mission accomplishment], and so on), while minimizing the chance for unexpected outcomes". This goal is best accomplished through the use of an established, proven framework for managing information security risks to organizations. This chapter proposes an approach based on the structure provided by the NIST.

The approach described by the NIST is used in this chapter due to several factors. First, the NIST approach has been developed to be consistent with and harmonize with international standards to the extent appropriate. These international standards include those of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) [3]. This approach is also being adopted and used by the U.S. government for virtually all government organizations, as well as other private organizations regulated by the U.S. government.

The approach described by the NIST is based on a 4-step process, used within a 3-tiered structure. The 3-tiered structure is used to depict the principal functional areas within an organization as they relate to risk management decision-making – the organization, mission/business process, and information systems Tiers – described in Section 2.2. The risk management process and the 4-steps in the process – frame, assess, respond, and monitor – are described in Section 2.3.

2.1. Risk management overview

Information security practitioners are transitioning from a compliance-based, checklist type of approach to a more risk managed approach to security [3]. This is being done largely due to practicality and resource constraints [3]. It is not possible to eliminate risk in an

[3] The NIST risk management structure is aligned with ISO/IEC 3100, *Risk Management – Principles and Guidelines*; 31010, *Risk Management – Risk Assessment Techniques*; 27001, *Information technology – Security techniques – Information Security Management Systems – Requirements*; and 27005, *Information Technology – Security Techniques – Information Security Risk Management Systems*.

information system, and is resource intensive to try [3, 4]. In order to effectively manage resources and maintain usability of the system, it is necessary to implement a risk managed approach to securing Information Technology (IT) systems.

To understand this change, it is important to start with a good definition of risk. A good definition in this case is one that can be understood operationally and can be easily used to clarify the process. Not surprisingly, there are many different definitions of risk. Wheeler [3] defines risk as: "the probable frequency and probable magnitude of future loss of confidentiality, integrity, availability or accountability". Accountability is not commonly accepted as being a part of the definition of risk, is not included in the definition of risk used by the NIST[4], and thus will not be included in the definition of risk used in this chapter. The definition of risk used for this chapter is therefore:

Risk: The probable frequency and probable magnitude of future loss of an organization's operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, and the Nation resulting from a loss of confidentiality, integrity, or availability.

There are two fundamentally different approaches to Information security risk management activities used by organizations:

1. Compliance-based; and
2. Risk management-based.

Each of these approaches, while working toward the overall objective of ensuring the confidentiality, integrity, and availability of the organization's information and information systems, attempts to meet that objective in a fundamentally different way. Compliance-based approaches default to including controls from the best practice or other framework they are implementing – not including a prescribed control is the exception [5]. On the other hand, risk management-based approaches default to not including a control unless its need and utility can be justified by a risk analysis [5].

2.1.1. Compliance and best practice frameworks

Failure to comply with the legal and regulatory structures confronting an organization can result in penalties, loss of contracts, loss of confidence, loss of business, and stock price declines [6]. Some of the requirements that organizations may need to maintain compliance with include:

- U.S. Legal requirements (e.g., the Sarbanes-Oxley Act, the Health Insurance Portability and Accountability Act (HIPAA), Gramm-Leach-Bliley Act (GLBA);

[4] The definition of risk provided by the NIST is: "A measure of the extent to which an entity is threatened by a potential circumstance or event, and typically a function of: (i) the adverse impacts that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence. [Note: Information system-related security risks are those risks that arise from the loss of confidentiality, integrity, or availability of information or information systems and reflect the potential adverse impacts to organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, and the Nation]" [17].

- Non-U.S. Legal requirements for organizations operating outside the United States;
- International frameworks (e.g., Basel, Basel II); and
- Industry standards (e.g., the Payment Card Industry Data Security Standard (PCI or PCI-DSS)).

Compliance with these requirements is often a part of an organization's due diligence [7]. The focus on best practice frameworks frequently focuses on satisfying the auditor/examiner, helping to meet compliance requirements rather than the organizations genuine information security needs [8, 9]. Unfortunately, these checklist or compliance-based approaches to information security risk management provide static, "one-size-fits-all" information security "solutions" [3]. One result of this approach is the common perception among information security practitioners that "if you are secure you *may* be compliant, but if you are only compliant you are certainly not secure" [6].

There are a number of best practice frameworks for information security, including the Information Technology Infrastructure Library (ITIL) [10], COBIT [11], and the ISO/IEC 27001 [12], 27002 [13], and 27005 [14]. Unfortunately, these approaches are dated almost as soon as they are published due to the speed of change on the Internet and within the IT security arena [15]. Attackers are very adaptable, and change their tactics quickly. In addition, a checklist or compliance-based approach assumes that every system requires the same protection as every other system, without regard for cost, information sensitivity, and mission or business impact [3]. However, there are often considerable differences in the types of systems deployed in an enterprise, and the information they contain. Since the information contained in the IT system is normally the critical asset requiring protection, the protective mechanisms that should be implemented will depend largely on the sensitivity of the information processed, stored, or transmitted by the system [8].

However, best practice frameworks do provide a useful method for ensuring that all aspects of an IT information security program are considered when using a risk managed approach to IT information security requirements development. Due to the need of organizations to remain compliant with respect to specific legal requirements and industry frameworks some form of compliance-based approach is likely to remain necessary.

2.1.2. Risk managed approaches

Risk-based approaches to securing information systems allow organizations to customize their information security protections, based in the needs of their organization. Using a risk-based approach requires consideration of the information processed, stored, or transmitted by the information system, as well as consideration of the IT system's environment, connectivity, and threat environment [7, 16]. Other considerations include the cost of implementing security controls weighted against the impact on the mission and business operations of the organization should a loss of confidentiality, integrity, or availability of the information occur [7].

A risk-based approach to information security is what the U.S. government is transitioning to with the release of the NIST Special Publication (SP) 800-37 [2] and SP 800-39 [3, 17]. The NIST has developed a series of SPs to focus on information security risk management, starting with an enterprise view in SP 800-39 [17], an information system view in SP 800-37 [2], and by providing an approach for performing risk assessments in SP 800-30 [18]. NIST SP 800-53 [19] provides a catalog of security controls, and recommends "baselines" of security controls based on the sensitivity of the information on the system. These baselines are intended to be customized, or "tailored" to meet the needs of the information and the information system when consideration of the system's environment, connectivity, and threats are considered [19].

2.2. 3-Tiered risk management structure

When addressing information security risk management activities, the NIST and other authors divide organizations into three levels [17, 20]. The NIST [17] identifies these tiers as the:

1. Organization;
2. Mission/Business Process; and
3. Information Systems.

Each tier has different organizational risk management responsibilities. However, despite their different perspectives and roles, they all use the same 4-step risk management process described in Section 2.3 for risk management actives within their Tier. This 3-tiered structure is depicted in Figure 1.

Figure 1. Risk Management Tiers[5]

[5] This figure is adapted from Figure 2 on page 9 of NIST SP 800-39 [17].

2.2.1. Organization tier

At the organization tier, risk to the entire organization is considered and managed. Part of the responsibility for managing risk throughout the organization is the process of "risk framing", establishing the context within which all organizational risk management activities will be conducted [17]. Risk framing establishes the governance framework from which are derived the risk management activities and the risk tolerance of the organization. Other activities occurring at Tier 1 include:

- Establishment and prioritization of activities and programs at the Mission/Business Tier;
- Determination of organization-wide common controls[6]; and
- The Risk Executive (Function)[7] recommended by the NIST is established in and is located within the organization at this level [17].

The Risk Executive (Function) (REF) is an individual or group within the organization that serves in an advisory role to organizational decision makers at all 3-tiers. The REF does not make decisions for the organization; rather it informs decision makers about the risks to the system, network, and the organization that may result from a particular risk decision. The REF considers risk from a holistic perspective, considering mission and business risks, in addition to security risks. This risk consideration is done with an organization-wide perspective, allowing the REF's recommendations to evaluate the potential impact of accepting risks in one area or system on other systems or the organization.

2.2.2. Mission/business process tier

Tier 2 is where the mission/business processes necessary to implement the strategic goals and objectives established at Tier 1 are defined and prioritized [17]. This is also where the types of information required, the information sensitivity, and information flows necessary to support the Tier 1 goals and objectives is determined [17]. The IT enterprise architecture is defined and established at this tier, to include the implementation of the common controls identified in Tier 1. The decisions and activities at this tier have a direct effect on the activities undertaken at Tier 3.

2.2.3. Information systems tier

Tier 3 is where the information systems that support the organization reside. The decisions and prioritizations established in Tiers 1 and 2 are implemented in information systems at this tier. The activities required for each of the steps in the Risk Management Framework

---

[6] Common controls are security controls implemented in a way that they are available to information systems across the organization. Common controls can be "inherited" by a system, meaning that the system itself does not need to implement the control the system can leverage/use the control as implemented by the organization.

[7] The Risk Executive (Function) is an individual or office responsible for considering risk across the organization, to include mission, business, and security risks, balancing them appropriately for the organization, and making recommendations to decision-makers within the organization based on their risk determinations.

(RMF)[8] and the system development life cycle [9] are performed here, to ensure each information system meets its technical, mission, and security requirements. In this tier, "information system owners, common control providers, system and security engineers, and information system security officers make risk-based decisions regarding the implementation, operation, and monitoring of organizational information systems" [17].

2.3. Risk management process

The risk management process is comprised of a number of discrete steps. These steps take place at different times within the process, and possibly at multiple times in the process due the iterative nature of risk management activities. It is important that all of the steps are completed for a risk management program to be fully effective. These steps apply to risk management activities taking place at each tier within the organization, so this process is equally applicable to risk management activities taking place at Tier 1 as it is at Tier 2 or 3.

NIST describes four distinct steps in the risk management process. These steps are:

1. Risk Framing
2. Risk Assessment
3. Risk Response
4. Risk Monitoring

2.3.1. Risk framing

Risk framing is a governance activity that is performed at Tier 1. Its principal output is a risk management strategy "that addresses how organizations intend to assess risk, respond to risk, and monitor risk" [17]. The risk management strategy is created as a joint effort between an organization's senior management and/or executives in conjunction with the risk executive (function) [17]. The risk management strategy explicitly states the assumptions, constraints, risk tolerances, and priorities or trade-offs used in making investment and operational decisions for the organization [17]. It also details what types of risk responses are supported, how risk is assessed, and how risk is monitored for the organization [17].

2.3.2. Risk assessment

Risk assessment is the process of:

1. Identifying risks (threats and associated vulnerabilities) to an organizational asset, activity, or operation;
2. Estimating the potential impact and likelihood of the risk materializing, and

---

[8] The RMF is described in detail in NIST Special Publication (SP) 800-37, available from: http://csrc.nist.gov/publications/nistpubs/800-37-rev1/sp800-37-rev1-final.pdf

[9] The system development life cycle established by the NIST is described in NIST SP 800-64, *Security Considerations in the System Development Life Cycle*, Oct. 2008 [23].

3. Prioritizing the identified risks according to their severity to the organization [3, 17].

Risk assessments can and should be conducted at every Tier of the organization. However, the objectives of risk assessments conducted at different Tiers will reflect the differences in responsibility and objectives for the Tier being assessed [17]. For example, a risk assessment at Tier 3, the Information Systems Tier, will go into considerable technical detail on a specific information system and the risks involved in its operation. Whereas a risk assessment at Tier 1, the Organization Tier, may address information systems from the perspective of the organization's enterprise architecture or common control framework, but will not go into significant technical detail, nor will it address a specific information system. However, a Tier 1 risk assessment will address business risks, and the risks involved in investing in particular missions or business areas.

2.3.3. Risk response

Responding to risk involves deciding on and implementing a course of action to address the risk within the organization. The options available to the organization for risk response are defined in the organization's risk management strategy. The courses of action available to address risk as identified in NIST SP 800-39 [17] are:

1. Accept – take no action when the risk is within the organization's risk tolerance;
2. Avoid – eliminate the activities or technologies resulting in risk that exceeds the organization's risk tolerance, or reposition the activities or technologies into areas or positions where the risk is avoided;
3. Mitigate – apply security controls, safeguards, or process re-engineering to reduce the risk to a level acceptable to the organization;
4. Transfer or Share – shifting all or part of the liability for risk, respectively, to another organization.

2.3.4. Risk monitoring

Risk monitoring is conducted on an ongoing basis to ensure that the organization's risk posture remains within the organization's risk tolerance. The risk monitoring process allows an organization to:

- Ensure compliance with national laws, regulations, organizational policies, and mission/business functions;
- Determine how effective its risk response measures are;
- Identify changes to organizational assets or their operating environments that result in changes their risk postures [17].

Risk monitoring activities are normally conducted on a periodic basis, with the period determined in accordance with the organization's risk management strategy and the sensitivity of the information or business process being protected [16, 17]. Monitoring risk and changes to operating environments includes identifying changes to the threat environment, and determining whether a change in the threat environment requires a reassessment of the risk posture of an organizational asset, activity, or operation earlier than normally scheduled [3].

These risk monitoring activities can be implemented at any tier in the organization [17]. The objectives and process for each tier will differ according to their respective needs, and, particularly at Tiers 1 and 2, are likely to involve cross-tier monitoring – as the upper tiers are directly affected by operational or process-level changes to the organization's risk posture at the lower levels [17]. An example of cross-tier monitoring would involve the monitoring of the risk posture of an information system authorized to operate within an organization. The organizational official responsible for the authorization decision will want to monitor the risk posture of the information system to ensure it continues to operate within an acceptable level of risk, and that any changes to the risk posture of the system or its environment are properly evaluated and addressed.

3. IA OM®

The metrics or results derived from the IA OM® methodology are meaningful to the organization because the measurements themselves are "tied directly to questions that are important to the organization"[21] . The results are also useful to organizational management since they indicate the degree to which specific information security risk management goals are being met as action is taken to improve an organization's overall information security posture in terms of its information security objectives [1]. In this instance IA OM® can be conceptually expressed as providing a measure of the degree to which the organization's information security risk management objectives are being met [1].

The creators of the OM® methodology, Donaldson and Siegel [21, 22], have 20+ years of professional software engineering experience at Science Applications International Corporation (SAIC) and in the Department of Defense (DoD). The OM® methodology enables one "to measure software products and software systems development processes in everyday terms familiar and – therefore meaningful – to your organization" [21]. The OM® methodology measures software products and software systems development processes as part of a continual process improvement exercise. The OM® framework derives its effectiveness as a "management process" tool in that the OM® Index, akin to the Consumer Price Index, folds in a number of individual measurements into a single overall value [1, 22]. The OM® Index can also be deconstructed to gain insight into the elements comprising the index value. By looking at trends in the index values, it is possible to determine the effect or outcome of changes within the organization.

The OM® quantifies software *product* "goodness" and software *process* "goodness", where an object (i.e., an attribute, component, or activity) is measured through its characteristics. "For products, these characteristics are called attributes; for processes, these characteristics are called components and activities" [22]. Software product "goodness" is the degree to which the product satisfies the customer and meets the customer's requirements. Software process "goodness" is a measure of the product creation process' ability to consistently and reliably create good quality products within budget [1].

The IA OM® metric can be used with existing organizational objectives or industry best practices. However, industry best practices are not appropriate for many organizations, and as described in Section 2.1.1, are often not sufficient in and of themselves to meet the an organization's requirements. Each organization must determine what is most appropriate for its needs. This is best accomplished when requirements are evaluated in the context of their budget and a thorough risk analysis [7]. Evaluating an organization's information security risk management posture, based on how well its systems comply with the organization's information security risk management objectives, provides a metric with greater versatility and applicability across a wider range of organizations [1].

Implementing IA OM® at an organizational level provides Senior Management with a high-level or strategic-level view of where its organization's risk management program stands and how well it is meeting its stated information security risk management objectives. IA OM® is the OM®[10] metric created by Donaldson and Siegel [21] adapted to information security and information security risk management activities as IA OM®[11]. The IA OM® metric can be used to:

- Quantitatively determine the degree of risk identified within an organization's information security risk management program;
- Characterize the organization's risk management policy elements as they are applied to its information security risk management program for IA OM® evaluation;
- Determine the weighting factors, based on a determination of the relative importance of each component, for the identified characteristics of the organization's risk management strategy and policy;
- Periodically present the results of the evaluation, the current risk management posture of the enterprise, in a balanced scorecard-like format that is familiar and easy to understand and interpret by Senior Management;
- Analyze the metrics data, identify the strengths and weaknesses of proposed metrics approaches; and
- Suggest areas for future assessment and evaluation.

IA OM® will help answer Senior Management's questions regarding the state of their organization's information security risk management program enabling them to determine their organization's current risk posture, identify areas needing improvement, and prioritize the allocation of organizational resources in addressing identified risks.

Any effective risk management program requires periodic monitoring and re-assessment, including risk monitoring at the organization, mission/business process, and, information systems Tiers. The components of risk management at these multiple-tier levels are specifically identified[12] and encompass the following:

---

[10] OM® was developed by Donaldson and Siegel, SAIC, [21] to evaluate system development life cycle (SDLC) processes.

[11] See Ting and Comings [1] for a complete description of IA OM®.

[12] NIST SP 800-39 [17], *Managing Information Security Risk*, describes the fundamentals of risk management in Chapter 2 and the process for framing, assessing, responding and monitoring risk in Chapter 3.

- Tier 1 – Addresses risk from an organizational perspective by establishing and implementing governance structures including: 1) Establishment and implementation of a risk executive (function); 2) Establishment of a risk management strategy including a determination of organizational risk tolerance; and 3) Development of organization-wide investment strategies for information resources and information security.
- Tier 2 – Addresses risk from mission/business process perspective by designing, developing, and implementing the processes supporting the mission/business functions defined at Tier 1 including: 1) Risk aware processes designed to manage risk according to the risk management strategy defined at Tier 1 and explicitly accounting for risk in evaluating the mission/business activities and decisions at Tier 2; 2) Implementing an enterprise architecture, and, 3) Establishing an information security architecture as an integral part of the organization's enterprise architecture.
- Tier 3 – Addresses risk from an information system perspective, guided by risk context, risk decisions and risk activities at Tiers 1 and 2, risk management activities (i.e. activities at each step of the Risk Management Framework (NIST SP 800-37 [2]) and in the systems development life cycle (NIST SP 800-64 [23])

These three Tiers, properly integrated, provide the capability to establish a strong, risk-based security infrastructure for an organization.

Risk management monitoring must be conducted at all three Tiers, making sure that all key activities are performed properly within each Tier and across Tiers [17]. Monitoring at the information systems level must take into account those controls that are "common" to an organization or enterprise [2]. Common controls are those controls that are established by an organization itself, or an element within an organization, and are made available for use by other elements and information systems within the organization. It is often not possible for an individual system owner to monitor common controls – such monitoring must be provided by the Common Control Provider.

With monitoring taking place at many levels within the organization, the need for an enterprise-wide risk management solution is even greater. Without a big-picture view of the information security risk posture of the systems within an organization, there may be systems operating that are creating significant risks to the organization without anyone in a position to address the problems realizing a problem exists.

At the Tier 1 and 2 levels, the concept of common controls is not the same as at Tier 3, however, there is still the need to ensure that organizational guidance and organizational functions remain aligned. When strategies, policies, and other organization-level guidance changes, the changes need to ripple through the organization – updating policies, programs, investments, etc. to ensure they remain in alignment with the top-level guidance.

IA OM® addresses these needs by aggregating the results of risk monitoring programs occurring throughout the organization, rolling them up, and presenting them as a set of summary statistics indicating where an organization stands with respect to remaining within its:

- Overall risk tolerance
- Risk tolerance within organizational elements
- Risk tolerance for individual systems

Applied this way, IA OM® allows Senior Management to readily assess their current information security risk management posture and determine whether it fits within their risk tolerance. It also identifies those areas, programs, and systems of greatest risk to the organization – allowing Senior Management to quickly and easily prioritize their remediation efforts.

4. Using IA OM® as an enterprise risk management metric

The process for using IA OM® as an enterprise risk management metric involves a number of steps. The overall steps in the process, as adapted from the OM® process, are:

1. Decide what questions you need or want answers to regarding your organization's risk management program;
2. Identify the organizational assets, processes, or operations that need to be measured to answer the questions from step 1;
3. Identify any characteristics and sub-activities of the organizational assets, processes, or operations to be measured from step 2;
4. Define an activity value scale for each activity or sub-activity in terms that make sense within the organization;
5. Determine the current value (or location along the value scale) for each activity or sub-activity being measured;
6. Calculate the value for each asset, process, or operation identified in Step 2 using the formulas provided in Section 4.4 and the activity (or sub-activity) values from step 5. Weighting factors are selected based on the organization's determination of the relative importance of the activities;
7. Combine the values for each activity (asset, process, or operation) into an overall IA OM® index value to be reported and analyzed.

These steps and activities are applied to organizational risk management using IA OM®, resulting in a metric that provides:

- The ability to evaluate the risk posture of a specific organizational asset, process, or operation;
- A set of organizational assets;
- All activities within a risk management tier; or
- Risk management activities across tiers.

This process is shown in Figure 2 [24] using an activity from an organization's personnel security process as an example. The individual process steps are examined in more detail in Sections 4.1 – 4.7.

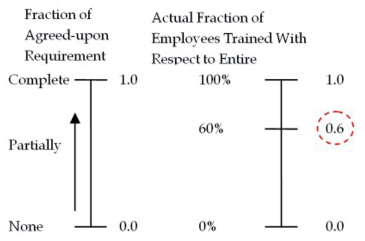
IA OM® Process[13]	
Steps	**Examples**
1. Decide what questions you want answered	Am I complying with my organization's personnel security policy?
2. Identify the organizational assets, processes, or operations that need to be measured to answer the questions from step 1	1) Employee job description 2) Employee information security training
3. Identify any characteristics or sub-activities of the organizational assets, processes, or operations to be measured from step 2	1) Extent to which security in job definition and resourcing are met 2) Actual fraction of employees trained with respect to entire organization
4. Define an activity value scale for each activity or sub-activity 5. Determine the current value (or location along the value scale) for each activity or sub-activity	
Associate observable events with value scale numbers	1) Number of requirements fulfilled in employees' job description 2) Date that employee trained records in training log that employee received
Measure each characteristic	1) Filled positions addressed 7 of the 10 employee's security requirements called out (Value = 0.7) 2) The employees' training log showed that 1 employee did not receive training
6. Calculate an index by substituting measured values into the appropriate IA OM® equation	$PersonnelSecurityComplianceIndex = \frac{\sqrt{0.7^2 + 0.6^2}}{\sqrt{2}} = 0.65$
7. Combine the values for each activity (asset, process, or operation) into an overall IAIndex value[14] to be reported and analyzed	$IARMIndex = \frac{\sqrt{\sum_{i=1}^{n} w_i^2 da_i^2}}{\sqrt{\sum_{i=1}^{n} w_i^2 (\max[da_i])^2}}$

Figure 2. IA OM Process with Personnel Security Process Example

[13] Adapted from interview with Stanley G. Siegel on Jan. 7, 2004 [24].

[14] This equation and its components are described in Section 4.5.

4.1. Step 1 – Ask questions

IA OM® is like other investigative ventures – the first step in the process is determining what you want or need to know. This chapter focuses on evaluating and understanding the ongoing risk management activities within an organization. As a result, the questions framed for use with IA OM® in this chapter focus on what the organization's executives, program managers, or system-level managers want or need to know about the risk posture of the portion of the organization they are responsible for. As such, the questions to be addressed by IA OM® need not to be restricted to Tier 1 – they can, and ultimately should, be spread across all three tiers so that the managers at each tier have the answers they need to be successful in the organization's risk management program.

Examples of questions that might be asked at each Tier are presented in Table 1:

Tier 1	Tier 2	Tier 3
Is our risk management strategy aligned with our organizational goals and objectives?	How well do our mission/business processes align with our organization's risk management strategy?	Are our information systems properly operating within the risk tolerance of our organization's mission/business processes?
How well aligned is our risk management strategy with our mission and business programs?	How well does my mission program align with our organization's risk management strategy?	Is the risk posture of this system within established boundaries to support this mission?
Is our common control strategy effective?	Which common controls are cost effective?	How fully am I using the common controls available to me?

Table 1. Example questions by Tier

4.2. Steps 2 and 3 – Identify assets, processes, or operations to measure

Now that the questions to be answered have been identified, the next step is to identify the assets, processes, or operations that can be measured to obtain answers to those questions. For example, to determine the effectiveness of an organization's common control strategy, you could examine:

- The common controls called for in the common controls strategy/policy;
- Their alignment with current organizational goals and objectives; and
- Their utilization.

An example analysis of Tier 1 risk monitoring activities, their components, and their subcomponents is provided in Figure 3. The common controls assets are presented as Characteristic_n in the right hand branch of the figure. These characteristics and subcharacteristics will also be referred to as diagnostic areas, where diagnostic area 1 (da_1) is equivalent to characteristic_1; da_2 is equivalent to characteristic_2; and so forth. The subcharacteristics for each area follow a similar numbering scheme where, for example, $subcharacteristic_{11}$ is equivalent to da_{11} and so forth. This is done to simplify the abbreviations used in the equations in Steps 6 and 7.

It is important to remember that the identification of organizational assets, processes, and operations are organization-specific. This is one of the strengths of the IA OM® process, since it enables the abstraction of these key activities – specifically identified by the organization as being of interest – into a metric that shows Senior Management where their organization stands with respect to its risk management strategy and policies. If low level technical metrics exist, they can be combined and abstracted into the IA OM® process. Metrics produced through the use of the IA OM® process can be deconstructed into their component areas, allowing Senior Management to identify the areas needing attention. If further improvement in their risk management program is required, the IA OM® and its component measures can be used to track and improve the organization's risk posture over time.

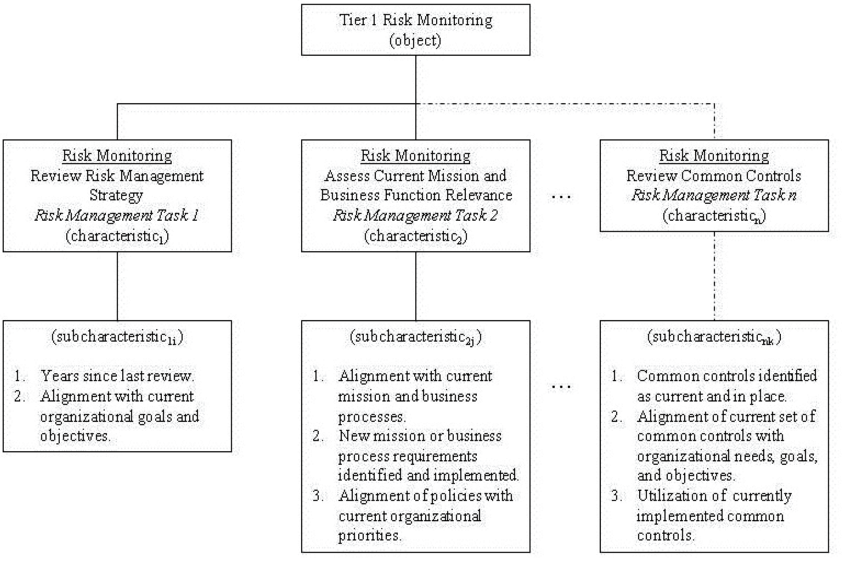


Figure 3. Example analysis of risk monitoring activities and their subcomponents

4.3. Steps 4 and 5 – Define activity value scales and determine activity values

Activity value scales used for IA OM® activities are normally scoped to range between 0 and 1. This makes comparison easy and allows them to be aggregated and rolled-up into measures that are easy to use and understand. Also, these values can be readily seen as percentages to further enhance their understanding. However it is perfectly acceptable to have more important characteristics have values greater than 1 if desired to indicate their relative importance to the organization. Alternatively, the relative importance of different characteristics/subcharacteristics can be accounted for using the weighting factors discussed in Steps 6 and 7.

For activities that can only assume a specific set of values (e.g., Yes/No, or high, moderate, and low), the value scales can be adapted to accommodate them. For example, with a binary set, like Yes and No, it is common to use Yes = 1, and No = 0. Sets like high, moderate, and low could be represented with high = 1, moderate = 0.5, and low = 0. It would also be possible to decide that high = 0.9 (since even high is not definite like "yes"), moderate = 0.5, and low = 0.1 (since low is also not definite like "no"). The decision on value scales is made by the organization and is made to maximize the utility and understandability of the measurements in the context of their organization.

Continuing with the example of the effectiveness of an organization's common controls strategy, value scales for each of the three subcomponents can be defined, and values determined for each of the subcharacteristics assigned as shown in Figure 4. Each subcharacteristic/diagnostic area (da_{ij}) will be evaluated separately and then combined in Step 6 to provide an overall value for each characteristic or top-level diagnostic area (da_i). These values are then combined in Step 7 to provide an overall IA OM Index Value.

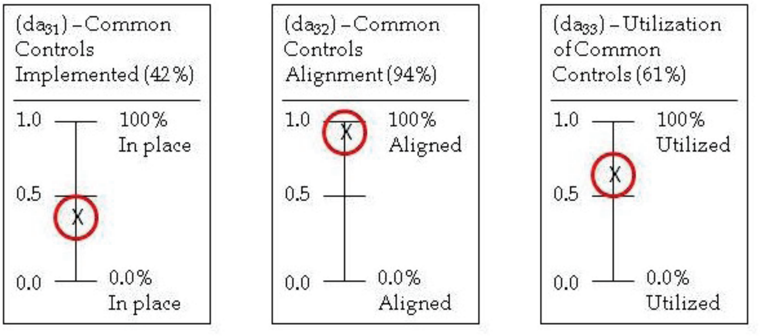


Figure 4. Common Control Value Scales Example

4.4. Step 6 – Calculate an asset, process, or operation index

Steps 3 through 5 have shown how to derive values for each of the assets, processes, and operations identified in Step 2. Each of these values represents a characteristic or subcharacteristic for an asset, process or operation. In this step, any values for subcharacteristics (da_{ij}) will be combined with weighting factors and aggregated to provide values for each characteristic (da_i).

The weighting factor value (w_{ij}) for each subcharacteristic (da_{ij}) is assigned by organizational management to represent the organization's determination of the relative importance of each subcharacteristic to the characteristic being evaluated.

Equation 1 – Calculating diagnostic area (characteristic) values:

$$da_i = \frac{\sqrt{\sum_{j=1}^{ni} w_{ij}^2 da_{ij}^2}}{\sqrt{\sum_{j=1}^{ni} w_{ij}^2 (\max[da_{ij}])^2}}$$

Where:

n_i= number of diagnostic criteria for diagnostic area da_i
w_{ij} = weighting factor for diagnostic da_{ij} of diagnostic area da_i
da_{ij} = the j[th] diagnostic criterion of the of the i[th] diagnostic area da_i
max [da_{ij}] = maximum value of da_{ij}

For the common control example, using organizationally determined weightings and subcharacteristics/diagnostic areas values provided in Table 2:

Weightings	Subcharacteristics/ diagnostic areas
$w_{31} = 1$	$da_{31} = 0.42$
$w_{32} = 1$	$da_{32} = 0.94$
$w_{33} = 2$	$da_{33} = 0.61$

Table 2. Weightings and Subcharacteristics Values

The equation works out to:

$$Characteristic_3 = \frac{\sqrt{(1^2 \times 0.42^2) + (1^2 \times 0.94^2) + (2^2 \times 0.61^2)}}{\sqrt{(1^2 \times 1.0^2) + (1^2 \times 1.0^2) + (2^2 \times 1.0^2)}}$$

Thus:

$$\text{Characteristic}_3 = \frac{\sqrt{2.55}}{\sqrt{6}} = .65$$

Inserting values for the other characteristics, and arranging all of the values into a fishbone diagram for clarity of presentation provides breakdown of the process as shown in Figure 5.

4.5. Step 7 – Calculate the IA OM® index

Calculating an IA OM® index provides a concise, high-level assessment of the enterprise risk management posture of the organization. The IA OM index, in this case referred to as the IA risk management index, or IARMIndex, is defined in terms of:

1. Diagnostic areas
2. Diagnostic criteria for each diagnostic area
3. Diagnostic criterion value scales

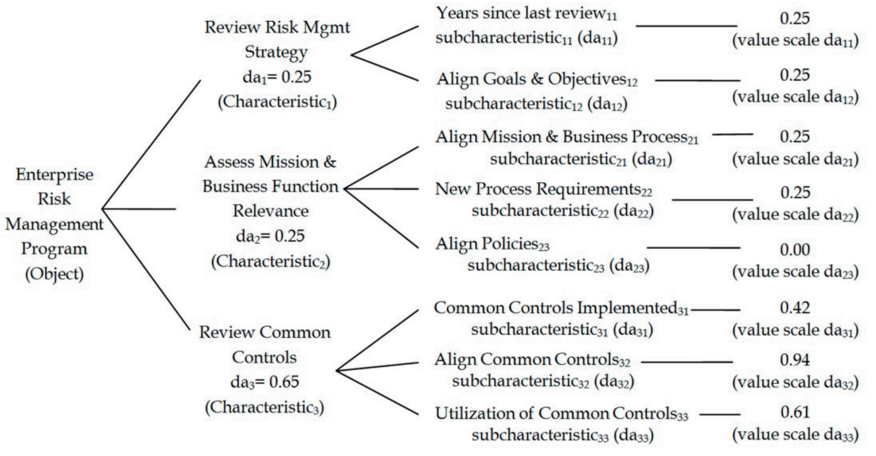


Figure 5. Fishbone Diagram with All Risk Monitoring Characteristics and Subcharacteristics Values

All of these items are provided in the fishbone diagram provided as Figure 5. A weighting factor can be applied to each characteristic/diagnostic area. As noted above, the weighting factor is an organizationally defined value indicating the relative importance to the organization of the particular characteristic. In this example the IARMIndex is normalized to one (i.e., ranges between zero and one). In all cases, value scales are defined in terms familiar to corporate management.

The process for calculating the IARMIndex is shown in Equation 2. For activity element characteristics, the IARMIndex[15] is normalized to one:

Equation 1 – Calculating the IARMIndex value:

$$IARMIndex = \frac{\sqrt{\sum_{i=1}^{n} w_i^2 da_i^2}}{\sqrt{\sum_{i=1}^{n} w_i^2 (\max[da_i])^2}}$$

Where:

da_i = diagnostic area (characteristic)
n = number of attributes
w_i = weighting factor for attribute da_i
max [da_i] = maximum value of da_i

Using the values from above and the organizationally defined weightings as shown in Table 3:

[15] Taken and adapted from Donaldson & Siegel [22] (p. 420).

Weightings	Characteristics/ diagnostic areas
$w_1 = 2$	$da_1 = 0.25$
$w_2 = 1$	$da_2 = 0.25$
$w_3 = 1$	$da_3 = 0.65$

Table 3. Weightings and Characteristics Values

Results in a value for **IARMIndex = 0.35**

4.6. Evaluating the IA OM® index

The following steps describe the analysis process allowing organizational management to quickly uncover areas needing attention and prioritize which area(s) to address first to obtain the greatest benefit:

1. The IARMIndex value is examined to determine whether it is within the range determined by the organization to correspond to its risk tolerance;
2. If the IARMIndex value is within the risk tolerance of the organization, the current values are included in the trending information and no further action is required until the next review cycle is initiated.
3. If the IARMIndex value is not within the risk tolerance of the organization:
a. The individual component index values that were aggregated to derive the IARMIndex value are examined and the value(s) furthest from the normalized value of 1.0 (the values closest to 0) is singled out for further analysis;
b. The selected component index value(s) is then unfolded (if applicable) to find the subcomponent(s) with the lowest value(s);
c. The component(s)/subcomponent(s) with the lowest value(s) is analyzed and a method for improving it is identified and implemented;
d. At management's discretion, reassess and recalculate the characteristic(s) (and any applicable subcharacteristic(s)) value(s) for the component(s) that was addressed and update the IARMIndex value to assess the impact of the changes made.
4. By looking at trends in the IARMIndex values, and the component index values compiled over time, the organization may determine the overall improvement resulting from addressing these components.
5. By looking at the IA OM® mapping and measurement trends in the component values, leadership or management can see which areas have had the greatest improvement, and which areas are most in need of attention.

Continuing with the example from above, the IARMIndex value = 0.35. Assuming the organization has determined that any value under 0.70 is outside of its risk tolerance, the next step is to determine which component(s) to single out for further analysis. Using the values in Table 4, we find that characteristics da_1 and da_2 have the lowest values. However, if we also consider the weightings we see that $w_1 = 1.0$ indicates that it is a higher priority item to the organization than w_2, suggesting that if we cannot address both areas, priority should be given to characteristic1, reviewing and updating the organization's risk management strategy.

Weightings	Characteristics/ diagnostic areas
$w_1 = 2$	$da_1 = 0.25$
$w_2 = 1$	$da_2 = 0.25$
$w_3 = 1$	$da_3 = 0.65$

Table 4. Component Values and Their Weightings

5. Conclusion

The need to initially assess, and then conduct ongoing monitoring of an organization's overall risk posture, the risk posture of its assets, processes, and systems has been clearly established. The more valuable the information, asset, activity, or operation, the greater the need to increase the frequency of monitoring activities to ensure these resources are not compromised, or to identify and respond to any compromises as quickly as possible. This chapter shows how the IA OM® metric herein described provides an enterprise-wide risk management metric that can integrate synergistically with other risk management tools and efforts within an organization to provide monitoring personnel and decision-makers with the timely, accurate, and useful information they need to perform their functions and ensure their organization's mission and business functions are protected. IA OM® not only provides a metric targeted to organizational senior management, but one that can be used by decision-makers at all levels in the organization to ensure the processes and assets they are responsible for are protected in a way that aligns with the risk management strategy of the organization.

Author details

David R. Comings
Tenacity Solutions, Inc., United States of America

Wendy W. Ting
Department of Defense, United States of America

Acknowledgement

David R. Comings, Ph.D., CISSP, CRISC, *Tenacity Solutions, Inc., United States of America*, Dr. Comings specializes in information security, risk management, and strategic planning particularly for customers within the U.S. Government. Dr. Comings earned his doctorate at the University of Fairfax and his Master of Arts at the University of Pittsburgh.

Wendy W. Ting, Ph.D., CISSP, CISM, *Department of Defense, United States of America*, Dr. Ting specializes in performance metrics, cross domain information-sharing solutions, information security and systems security engineering. Dr. Ting earned her doctorate at the University of Fairfax and her Master of Science at the University of Maryland.

6. References

[1] Ting WW, Comings DR. Information Assurance Metric for Assessing NIST's Monitoring Step in the Risk Management Framework. Information Security Journal: A Global Perspective. 2010; 19(5): 253-62.

[2] NIST. SP 800-37, Revision 1: Guide for Applying the Risk Management Framework to Federal Information Systems. In: Commerce Do, ed.: National Institute of Standards and Technology 2010: 93.

[3] Wheeler E. Security Risk Management: Building an Information Security Risk Management Program from the Ground Up. Waltham, MA: Syngress 2011.

[4] Olivia LM, ed. IT Solutions Series: IT Security Advice from Experts. Hershey, PA: CyberTech Publishing 2004.

[5] Straub DW, Goodman S, Baskerville RL, eds. Information Security: Policy, Processes, and Practices. Armonk, NY M E Sharpe Inc. 2008.

[6] DeLuccia IV JJ. IT Compliance and Controls: Best Practices for Implementation. Hoboken, NJ: John Wiley & Sons, Inc. 2008.

[7] LeVeque V. Information Security: A Strategic Approach. Hoboken, NJ: John Wiley & Sons, Inc. 2006.

[8] Pironti JP. Changing the Mind-set - Creating a Risk-conscious and Security-aware Culture. ISACA Journal. 2012 2012; 2: 13-9.

[9] Johnson ME, Goetz E. Embedding Information Security into the Organization. IEEE Security & Privacy. 2007 (May/June): 16-24.

[10] HM Government. ITIL. 2012 [cited 2012 4/17/2012]; Available from: http://www.itil-officialsite.com/

[11] ISACA. COBIT 5: A Business Framework for the Governance and Management of Enterprise IT. 2012 [cited 2012 4/17/2012]; Available from: http://www.isaca.org/COBIT/Pages/ default.aspx

[12] ISO/IEC. ISO/IEC 27001:2005 Information technology -- Security techniques -- Information security management systems -- Requirements. 2012 [cited 2012 4/17/2012]; Available from: http://www.iso.org/iso/catalogue_detail?csnumber=42103

[13] ISO/IEC. ISO/IEC 27002:2005 Information technology -- Security techniques -- Code of practice for information security management. 2012 [cited 2012 4/17/2012]; Available from: http://www.iso.org/iso/catalogue_detail?csnumber=50297

[14] ISO/IEC. ISO/IEC 27005:2011 Information technology -- Security techniques -- Information security risk management. Information Technology 2012 [cited 2012 4/17/2012]; Available from: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56742

[15] Johnson ME, ed. Managing Information Risk and the Economics of Security. New York, NY: Springer Science+Business Media, LLC 2009.
[16] Slay J, Koronios A. Information Technology Security & Risk Management. 3rd. ed. Milton, Qld: John Wiley & Sons Australia, LTD 2006.
[17] NIST. SP 800-39, Managing Information Security Risk: Organization, Mission, and Information System View. In: Commerce Do, ed.: National Institute of Standards and Technology 2011: 88.
[18] NIST. SP 800-30, Risk Management Guide for Information Technology Systems. In: Commerce Do, ed.: National Institute of Standards and Technology 2002.
[19] NIST. SP 800-53, Revision 3: Recommended Security Controls for Federal Information Systems and Organizations Gaithersburg, MD: National Institute of Standards & Technology; 2009.
[20] Jones A, Ashenden D. Risk Management for Computer Security. New York: Butterworth-Heinemann 2005.
[21] Donaldson SE, Siegel SG. Cultivating Successful Software Development. 1st ed. Upper Saddle River, NJ: Prentice Hall 1997.
[22] Donaldson SE, Siegel SG. Successful Software Development. 2nd ed. Upper Saddle River, NJ: Prentice Hall 2001.
[23] NIST. SP 800-64, Revision 2: Security Considerations in the System Development Life Cycle. In: Commerce Do, ed.: National Institute of Standards and Technology 2008.
[24] Ting WW. Interview with Stanley G. Siegel. Arlington, VA 2004.

Chapter 11

Risk Management in Business – The Foundation of Performance in Economic Organizations

Gabriela Dobrotă

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50706

1. Introduction

A world marked by rapid changes of the economic, financial, political and social environment, a world ruled by uncertainty is subject to the emergence of increasingly higher risks, affecting the process of economic development of world economy. The increasingly frequent manifestation of unforeseen events caused high interest for research in the risk identification, quantification and prevention at the microeconomic level. In this context, risk management can be considered the art of taking decisions in an uncertain environment, on the background of the identification, quantification, analysis and management of the risks which affect an organization.

Why is it necessary an active management of risk? The globalization process, the interdependence between economies in a regional and global plan, the problems arising from the need to ensure compatibility between legislative previsions, the effects of free labor movement, the macroeconomic context located in an accentuated dynamic, the fierce competition at the level of participants from the economic circuits, the limited degree of the resources and unlimited of the needs, the need to adapt to technological changes, the challenges generated of climate change, the high degree of complexity of the factors which influence economic and financial results of the business, the diversity of international economic flows are just some aspects which sustain the organized risk management, training the personnel for managing the activity, the identification of the losses caused by the action of the risk and the insurance of resources necessary to cover them, but also in the identification and communication of the risk, fact which requires the existence of a strong organizational culture oriented to this sense.

Defining risk has been done in different ways over several decades, the polemics continuing today. But regardless of the angle of approach, defining the border between risk and uncertainty, the identification of the management methods, a thing is certain: the existence of risk. The importance of the risk management is even more striking in the current economic climate, amid a crisis which seems to be having just begun. In these conditions, the timeliness and necessity of the topic is obvious. The existence of an arid land in Romanian economy seeing application of some specific techniques and strategies to the risk impose the immediate development of risk management, especially at the level of small and medium size entities (at the level of large companies it is noticed a considerable orientation towards the risk management, being created even different departments to exercise an active management at the level of risks).

The analysis of the economic-financial performances requires taking into consideration the risk factor as well, and hence of its possible implications on profitability and financial stability. At the level of any economy it cannot be discussed about the existence of some investments. Estimated return does not coincide with the actual one. Therefore, in addition to an analysis of profitability, the manager should also focus his attention on potential risk factors of the investment. In addition to estimating the expected future return it is imposed also a quantification of the risk associated with its achievement. Profit variability may involve achieving a lower return than expected; the higher the profit variation range is, the more risky the investment. Thus, the profitability - risk trend analysis provides a better classification of the investment opportunities. Rational investors, generally characterized by an aversion to risks, will exclude from the list those investments that offer the same return, but with a higher risk, and the option for one of the remaining opportunities will ultimately depend on their degree of risk aversion. Also, some authors consider that the risk is measured as variability in comparison with the profitability average of the last exercises and in foresight as variability of the profit in relation to the hope of profitability or as variability of the profit in relation to the volume of activity (turnover) of the company [1].

Avoiding risks due to significant changes in the activities concerned can be considered a beneficial strategy for the trader, in the context of the possibility of occurrence of problems with serious consequences. Against the background of strong economic and financial instability or when carrying out complex activities involving the development of partnerships, strategies may be adopted to follow fundamental indicators and to develop action plans for risk occurrence (for example, we can predict the resources used in case of higher costs or stocks of spare parts can be constituted to repair equipment and avoid large interruptions in production, are identified the opportunities for rescheduling of exchanges in case of defects which do not allow obtaining the quantity of finished products necessary to comply with the contractual obligations and so on). A commonly used option is that of transferring potential risks to specialized companies or individuals (insurance companies, experts in the field) or to some businesses which provide service for the technological equipment used, even if this generates additional costs.

The emphasized development of the international economic relations, the existence of a complex ensemble of agents that influence the economic-financial performances of the economic organizations, the economic - financial instability, the competitive environment's volatility, the extension of information and communication technologies are just some aspects that support unconditionally the necessity of risk management.

The research conducted aims to clarify an issue considered previously, but also to identify new issues, allowing further future approaches. Thus, the paper will be structured in several important parts, concerning the:

- Presentation of the concepts and theories regarding risk and highlighting the need to manage it;
- Elements of risk identification, with customization at the level of economic entities;
- Risk probability theory in risk study;
- Opportunities for risk assessment;
- The system of indicators that allow the determination of the economic - financial performance of economic entities, and the foundation of risk management decisions;
- The possibility of using discriminating statistical models in determining an entity's financial and economic status and the realization of forecasts on the outcome of future events, in order to provide answers to the managers' concerns to identify and use the most viable means of action in order to develop business and ensure sustainability of the entities they manage.

So, we consider that the problem that managers are facing is reflected by the need to manage risk at the level of any company, which implies: tracking to identify the factors that impact negatively on the work performed; the estimates quantifying the consequences of the event risk; basis for a complex of measures to prevent event risk; mitigate damage caused when it realized; using the services provided by the specialized units in risk management if it is not possible by the entity. Each manager must determine a minimum and maximum risk on the scale that is willing to accept, as a result of the company's results are dependent on them (assuming a higher risk lead to better results but and corresponding losses and vice versa).

2. The risks identification at the level of economic entities

Any firm bases their activity on two major coordinates: satisfying consumer requests and maximizing the obtained profitability. In order to achieve these goals, it is imposed primarily, a sizing of the capital necessary and identifying the funding opportunities. The object of the financing decision is realized by the selection of the sources of capital at the lowest cost for obtaining them in terms of risk reduction. So, developing a financial policy aimed at determining the financing needs for a period of time, selecting a financing structure, meaning a way of funding through its own resources or loans, and establishing the ratio between short-term use of resources or long term. But, all these issues require the identification of potential risks and establishing a strategy which allows reducing or even avoiding their effects.

Before performing a risk analysis, a differentiation between the concepts of risk and incertitude is necessary [2]. The unanimous opinion is that both regard essential categories that affect the general policy of a firm, but, often it has been made confusion between them, putting the sign of equality. In reality, the two concepts are different and proper understanding. The theoretical and practical approach of the two concepts has suffered many changes due to the more accentuated degree of complexity of the worldwide economy, relations in social area, technological area, economic crisis event, diversification of financial instruments, etc.

The inability of companies to adapt to changes recorded in the external environment with minimal cost can develop into a risk for this. If we refer to this sense, it is clear that any company (even the most profitable ones) is subject to constant risk, being imposed the development of risk management mechanisms to enable a fast referral of the changes recorded but also developing a mechanism of intervention.

A definition of the notion of risk is based on the changes recorded at the profit level compared to the average achieved in the previous years (this can be implemented also to the level of change in future profitability, of revenues to be obtained, of the results recorded). Another manner of approach, illustrates risk as the possibility of producing a fact which has unintended consequences. It is noted that the risk is regarded as the probability of manifestation of an event (possible to predict or not) with negative implications on the economic activity of a company.

Whatever the way of approaching the concept of risk we observe that it transforms the potential losses according to the probabilities of their manifestation, which are known or determined.

Instead, uncertainty implies inability to estimate those probabilities. It can be considered that the uncertainty is similar to a variable that cannot be defined fully as you cannot identify or predict possible events and neither the probability of their occurrence. In other words, uncertainty is present in the fact that it is not known which of the situations will intervene."[3]

In presenting these concepts it can be observed that the risk results from uncertainty. Thus, the inability to estimate an event to occur, of the time of registration and the size of the effects recorded materializes the state of uncertainty. Following, the adoption of decisions today determines the registration of results in the future, these being subject to a state of uncertainty.

At the level of an economic organization, uncertainty is similar with the risks that cannot be identified and estimated. So, any increase of their weight in the total of risks which affect the activity that takes place determine an increase of the uncertainty state and reverse (knowing and anticipating almost fully the risks involved in the activity that takes place is not similar with the elimination of uncertainty). Uncertainty is seen by the manager as the total amount of the consequences of the existence of some potential undesired situations. As a result, the objective or subjective character of uncertainty is reported to the necessity of correlating the opinions and perceptions of the decision maker with logical argument, based on real dates recorded in the previous period.

Obviously, no one can say that he proposed to avoid the uncertainty state, because, as it has been specified, the profound transformations that affect the entire world have such an accentuated dynamic as it would be utopian such an attitude. The bigger and bigger possibilities of choice make future being uncertain also for who realizes the strictest planning [4]. As a result, it is much closer to reality an acceptance of the uncertainty and maintenance of it in reasonable limits. A study made on a sample of 50 countries has highlighted important differences in the attitude regarding the avoidance of uncertainty, explained by differences of economic conjuncture, but also cultural, educational or behavioral [5]. Clearly, scientific investigation pointed the reduction of uncertainty at different components and behaviors of human mind. But, the activity of people has caused the action of some new uncertainties so that economic and financial crisis, poverty or globalization accompanies us permanently on the evolution scale of the whole society.

The more and more frequent occurrence of unpredicted events has caused a high interest for researches in the field of risk identification, quantification and prevention at microeconomic level. Therefore, ever since 1955, the professor Wayne Snider had defined the concept of risk manager, and in 1956 the concept of risk management appeared. On this background, risk management has known a fast development, being defined also the concept of *global cost of risk.* The definitions given to risk management have either empirical character(”the art of making the right choice, an art based rather on anticipating future events than on the reaction to past ones” ,- or “risk management is just common sense”) or pragmatic character (“the management of global cost of insurable or non-insurable risks, in a company” – [6]).

Although the concept of risk itself expresses a state of uncertainty and indeterminacy, however, it must be made a distinction of it, starting from the reasoning that we can act for the purposes of identifying opportunities for expressing some options, not only favorable but also unfavorable, and also of the degree of likelihood. In the market economy frame, the concept of risk has been widely debated, without existing uniform opinions, but rather diverse and even contradictory.

If we consider the definition of the explanatory dictionary of Romanian language, respectively:” The possibility of reaching a danger, of having to face trouble or harm suffered; Danger possibly more or less predictable”, we can depict the side of probability to express a danger which determines the human factors to identify opportunities in order to prevent and mitigate its effects. Into the specialized dictionaries are meet also other definitions of risk such as, “the possibility that a loan or investment to generate a loss ", and “likely future event, whose production may cause some losses”. It can be predictable when factors that cause losses can be predicted in advance and unpredictable when determined by fortuitous circumstances”. In the classical theory of decision, the risk is considered „**an uncertain element but possible that always appears in the social and human activities whose effects are damaging and irreversible”.**

From the definitions presented we can depict the conclusion that the risk is an uncertain phenomenon but with a certain degree of probability of event that can cause both losses and the effects which can be removed with difficulty or even at all. The area of risk event is broad, covering both the human side and also the social, political or other nature side. As a result, conscious or unconscious actions exerted on the components of a system can cause expected direct effects, appropriate to the objectives followed, but also the evidence of unwanted direct or indirect adverse effects that the specialty literature defines as risks.

Among the economic agents, there is a wide range of risks arising both from their actions and from the external environment. The development of the society has created new forms of risk and imposed the identification of new strategies and methods for forecasting and analysis of risk and uncertainty. Thus, to the risks caused by natural disasters, the ones related to economic processes and phenomena, armed conflicts, etc., were added also other categories that are often difficult to define or to estimate quantitatively. Thus, there is growing evidence of risks caused by increasing complexity of the business environment, changes in legal systems (especially for harmonization with other legal systems), conducting business in different countries which experience different rates of economic development, the action of political, demographic, cultural, juridical factors.

The group of risks can be realized from more points of view, in the specialty literature being meet many approaches. Without planning to repeat them, we shall try to report especially at those typologies that put in light categories which execute a big influence on the economic-financial performance of the economic organizations (table 1).

No. Item	Criterion of classification	Type of risk
1.	The source of the risk factor	□Internal □External
2	Predictability character	□Determinable □Uncertain
3	Nature of risks	□Natural □Political □Social □Juridical □Technological □Economic □Financial □Commercial
4	Level of occurrence	□Macroeconomic □Microeconomic
5	Type of transaction performed with financial instruments	□Price risk □Risk of liquidity □Credit risk □Variability risk of the cash flow

Table 1. Classification criteria of risks that affect the activity of economic entities

The risks specified can be deducted at their turn on subcomponents and they allow an analysis of generator causes. No matter the classification criterion and the possibility of dividing, one thing is for sure: the risk in business is a certain state, a feature of market economy. Its inexistence could lead to methods of performing the activity by the economic entities not according to the economy realities (practically, this thing would equalize with the existence of some administrative intervention leverages at the level of state that could eliminate freedom of action of the economic operators).

The conclusion of the aspects presented converges in only one way, more exactly that one where there is no probability of performing an activity in a certain environment, of course, no matter the field, branch or sector where an economic agent places his activity. As a result, exercising an active risk management can only have favorable effects at the level of economic operators. The structure of such a process can be highlighted schematically like this (Figure 1).

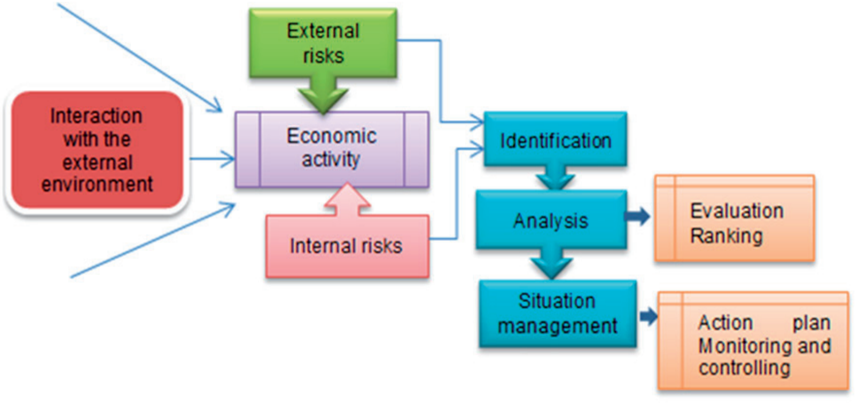


Figure 1. Structure of risk management process

The activity performed by a company is subject to different possibilities of manifestation of the internal risks and of the external ones, generated by the interaction with the environment. As a consequence, the first stage imposes the identification of different risks their notation in a "register of risks". Their analysis implies an evaluation and ranking process according to the effects they can generate. The elaboration of an action plan, with clear action measures, for avoiding risks and reducing the losses caused by their manifestation is the next mandatory step. But, there not enough these actions without a permanent verification of the risk management process and performing a rigorous control over the way the corresponding measures are applies so that they record an improvement of the company's performance.

3. Theory of probabilities in risk review

Can it be estimated and avoided the risk only by using the manager's intuition? The answer is clear, demonstrated by daily realities. An inherent aspect of any activity is the possibility of existence of a not desired event, for which it would be ideal at least determining the probability of manifestation. To the possibility of manifestation of a situation, process, phenomenon, etc. it can be associated an interval or one or more concrete answers. In the probabilities theory, these elements are defined as discrete random variables (the ones for which there can be enumerated the results) and continuous ones (those for which there can be placed the results in certain intervals).In practice, there can be also used mixed variables, which report to the possibility of using a combination of the two types mentioned previously for reflecting the possible results (these are specific to the insurance sector).

These probabilities can be identified, but, the recorded results do not appear at the same time also at the level of all variants. For instance, at a bowling competition, you can strike down between 0 and 6 skittles (there are 7 possible results), but only after the throw you will clearly know their number, without being possible recording in the same time two variants. The question is however to assess the probability of the results to action. In the case of discrete variables, things seem simple; the evaluation is possible by reporting favorable variations at the number of possible cases. For example, consider an economic agent who has an order for making furniture for at least one piece of each, but with a higher production capacity. In this context, he would like to know what is the probability to charge a minimum amount of 3000 euro from the sale of the goods mentioned, knowing that the unit price of sale is 500 euro for product X and 300 euro for product Y, and the quantity that can be offered for sale is of 5 pieces of each product. To identify favorable situations, are followed the dates from table 2.

1X,1Y	1X, 2Y	1X, 3Y	1X, 4Y	1X, 5Y
2X, 1Y	2X, 2Y	2X, 3Y	2X, 4Y	2X, 5Y
3X, 1Y	3X, 2Y	3X, 3Y	3X, 4Y	3X, 5Y
4X, 1Y	4X, 2Y	4X, 3Y	4X, 4Y	4X, 5Y
5X,1Y	5X,2Y	5X,3Y	5X,4Y	5X,5Y

Table 2. The Matrix of possible situations

Given the minimum amount targeted and unit sales prices, the favorable situations are represented by the following cases: (3X, 5Y) (4x, 4y) (4X, 5Y), (5X, 2Y) (5X, 3Y); (5X, 4y), (5X, 5Y). As a result, there are 7 favorable cases from 25 possible. The probability of obtaining the amount targeted is 7/25 = 0.28. Certainly, the example considered is simple, but we took into consideration highlighting a case that allows a quick estimate of the probability of recording a desired situation. And in the case of continuous random variables, the probability theory gives the theoretical foundation for identifying the possibility of manifestation of a situation.

It cannot be questioned the result of applying these mathematical methods for assessing the probability of outcomes to appear. And still, we must consider the following aspect: how many managers know the probability theory, are willing to apply it and are convinced that they can use it in any case, according to the level of complexity of the factors that can influence the work done? If it would be made a survey in Romania, especially at the level of small and medium organizations, the percentage would be extremely low.

Moreover, the risk management issue in our country has acquired a well-defined contour in a recent period, the researches performed being divided and the simple and the legislation is insufficient and permissive. In addition, although the manifestation of a result can be anticipated through these rational methods, even the term used, “probability” "shows that the activity of the economic agents cannot be translated into logical terms to quantify some uncertain aspects. Also, the use of some statistical data for sizing the probability of the appearance of a phenomenon allows obtaining relevant results when reporting to a large number of different cases that are not interrelated.

In this context, our approach supports the need for linking conceptual issues from the economic theory with the creation of a information database as complete as possible (reported not only to statistical data but also to forecasted information), so that economic entities to identify the trend of the evolution of phenomena and processes that influence the market and over the economic – financial situation. Such behavior would allow them to achieve the obtained objectives established with the available resources and opportunities for action. Obviously, the process of obtaining information is costly. But, what is more important - recording some additional costs generated by obtaining essential information and avoiding unintended consequences through the manifestation of some risks that have not been identified and evaluated? The answer can be supported by arguing that the firm size is conditioned by the need and the information possibilities, also by the information costs [7]. We must not believe that the fact of having a large volume of information on possible alternatives, allows adopting decisions in the absence of risk conditions and uncertainty. But, ignoring the information pushed to the extreme by the rationing “anyway it will happen, with or without strategy "can only be counterproductive attitude, with negative repercussions that can result even in bankruptcy.

4. Possibilities for risk assessment

Is risk identification sufficient to determine a course of action so as to avoid or minimize losses? Obviously, the answer is no. An important aspect of risk management is represented by dimensioning its size so that there can be evaluated the effects generated by its manifestation. In general, risk quantification starts from certain assumptions:

- The existence of multiple risks does not determine the reduction or increase of losses specific to a certain category (e.g., the loss caused by the manifestation of interest rate risk does not action over the one arising from environmental risk);
- The manifestation of a risk does not automatically act on the probability of the appearance of another risk (the increase of price of used raw materials does not directly or indirectly influence the risk of the firm's reputation);
- The losses caused by the manifestation of a risk must be under the level of available assets owned by the company.

On any business, the activity is reported on the components, differentiated in relation to the business established at its set up. Considering an economic agent engaged in the activities of production and marketing, we can identify and assess major risks that may occur in the economic, financial, investment and trade area. For this, it is imposed stating the possibilities of capital (Figure 2).

This structure allows capital allocation process targeting specific risk assessment activities of the two components. As a consequence, we shall report to the main risk categories that action at the level of each component.

Running operating cycle determines both revenue and expenditure, activity level all affect profitability. Obviously, these are uncertain and do not remain unchanged from one period to another. The revenue is subject to combined influence of several factors: the volume of production conducted, quantity sold, market demand, sales price. The level of expenses acts on profitability both through the volume and the structure. Thus, a high level of fixed costs will pose a risk for a major operating company. It will be even greater as the sales volume will be lower, the fixed costs will be higher and the output will be made near the critical point (that level, on account of production or profit made are not recorded, no loss). The economic theory offers solutions for the determination of the point where the developed activity does not generate any profit or loss and the economic realities have shown a tendency of bringing forward the growth rate of expenses of the corresponding increase in production due to higher labor productivity, improving the organization of work or assimilation of technical progress. In addition, the increase in the level of production leads to a lower cost per production unit, the allocation base for fixed costs being higher. In this context, companies which registered a high level of fixed costs are much more exposed to risk than those who have a high share of variable expenses. For the operational risk assessment the operators can monitor the level of turnover compared to breakeven. From the surveys conducted it emerges the idea that a difference between them up to 10% reflects an unstable situation, between 10 and 20% - a relatively stable situation and over 20% a comfortable situation. Also, another indicator that allows sizing the operating risk is the operating leverage which reflects the percentage increase of the operating result due to the increase by one percentage of sales.

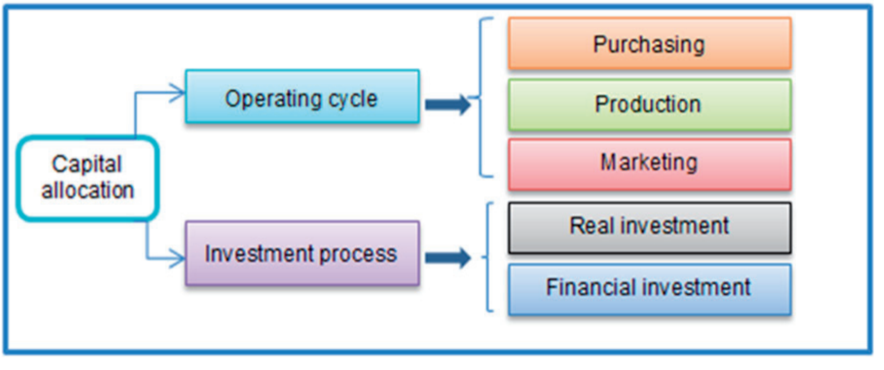


Figure 2. Use of capital at the level of economic agents

The risk manager must analyze all possible risks that can influence the profitability of the operating activity at the level of each phase. In this way, there can be identified and structured a series of risks as presented in table 3.

Phases of the operating cycle	Supply	Production	Sale
Potential risks	▪ Choosing wrong the supplier ▪ Contracting an insufficient quantity of raw materials; ▪ Supplying at intervals that do not allow continuity in ensuring the necessary resources; ▪ Concluding contracts with delay and non-compliance with de delivery clauses; ▪ Recording losses on the time of transport, handling and storage; ▪ Accidents of transport time; ▪ Recording a high level of expenses with supply and storage.	▪ Obtaining some products at a lower quality level; ▪ Insufficient use of production capacities; ▪ Use of nonperforming equipments; ▪ Use of production direct personnel with low qualification level; ▪ Increasing time of cancellation; ▪ Supplying the labor job with insufficient quantities of raw materials, ▪ Reducing labor productivity; ▪ Recording some high specific consumptions;	▪ The low transformation of finished products into monetary facilities; ▪ The reduction of the demand for finished products; ▪ The increase of the period of time between the moment finishing and that of dispatching; ▪ Not complying with the contractual clauses; ▪ Increasing the period of settlement period, ▪ Not cashing the receivables as a consequence of bankruptcy of clients.

Table 3. Potential risks associated to the operating activity

One of the most important decisions at the micro-economical level is the allocation of capital for the course of the investment processes because they reflect the performance of current costs in order to achieve future revenue. So, the quantity of resources allocated for the investment achievement can determine their success and the business development. Real investments, embodied by essential assets for the development of the production and sale processes, are the subject of differentiated risks, the ability of managers in the selection of projects that ensure the maximization of the economic agent value being essential. To substantiate the investment decision, the practice imposed certain elements that allow the sizing of the efficiency of the analyzed project, namely: the amount of invested capital, life of purpose, net cash flows, residual value and the discount rate [8]

The criteria that can be applied to analyze the investment project must relate to an uncertain environment because a certain environment does not exist in reality. Building some models on the assumption that future cash flows are certain, it has an explanatory nature. In a deterministic economic and social environment is assumed the previous and certain knowledge of the interest rate and future monetary flows of the investment projects. Possible variations in interest rates and the risk of an improper activity, that of achieving breakeven (the dead point) and/or to become insolvent and to go bankrupt are not included in the calculation base for the investments efficiency. As a result, the foundation of the investments decision in a certain environment has more a theoretical nature, of understanding the essential tools of the analysis of investment projects. It is about using the actuarial computations in selecting investment projects or project portfolios that maximize the wealth of equity investors (shareholders and creditors) and thus maximize the enterprise value. Concepts of interest rate as rate of productivity without risk (Rf) and net present value (NPV) are in the center of the analysis in the certain environment.

The determination of cash flows generated by the investment project and of the moment of their action, the balance of the capital market and the practice of an unchanged interest rate for a long time that allows the exact identification of the borrowing costs are items that cannot be recorded on different market segments. In this context, the estimation of future positive and negative flows is a difficult operation which can be completed with an unreal perception and therefore a failure in the investment process. Estimation can be achieved by identifying some probability elements in the registration of certain levels of the used indicators even in an uncertain environment. The states of uncertainty can be generated by the investment project or the market. So, the manner in which is done the investment projection, project implementation and its use, and also the capital market fluctuations, the competitors' actions, the manifestation of possible financial and economic crisis can generate risks and differences between predictions and reality. For the embracement of an investment decision in the probabilistic environment are required the following steps: the sizing of the positive and negative flows in different probability assumptions, the measuring of the investment risk, the adoption of investment decision for a specific project [9].

To achieve the first stage it is necessary the sizing of the mathematical hope, respectively the expected annual average value of the cash flows and their comparison. The comparison can be done with the update of the cash flows or without this operation. The mathematical expectance of the cash flows of the investment projects is determined by the weighting of the annual flows (estimated in three ways: one positive or optimistic, one pessimistic and one medium) with their probability of implementation (this is determined either statistically or by consulting specialists with experience in the field or on the investor's intuition and the sum of the probabilities must be equal to the unity). When there are analyzed several projects, it is obvious that it will be chosen the one that generates the greatest mathematic hope of realizing the future returns. Comparing the projects allows better results if it is used the technique of flows update. The drawback of this criterion is that it does not take into account the variation of the results around the average. As a result, in financial practice, it can be used the criterion of dispersion of the total net discounted absolute cash flow. In this sense, it is considered the best project the one which determines a total discounted net cash flow that has the smallest dispersion.

The risk reflects the possibility of undesirable results. In the investment process, the risk is much higher than in other economic activities because the effects are recorded for long periods (resulting the difficulty of accurate predictions of the obtainable effects) and the negative flows have comfortable values that can materialize in losses (in case of failure). For the avoidance of the risk, the investor may choose: the increase in the discount rate with a differential percentage according to the degree of risk associated with the type of investment (for example, the replacement investments are associated with a very little risk, the upgrading of the existing equipment in use involves a low risk, new investments have a higher risk while the strategic investments involve a substantial risk), but manifesting a degree of subjectivity in determining its size; the execution of an investment plan which can be varied according to specific conditions both from the point of view of costs that have to be made and also from the point of view of performing gradually the investments during their execution; projects with a lower lifetime (to avoid negative effects of the obsolescence, and also for projects that allow faster recovery of the invested capital).

At the same time, however, it can be measured the risk by using some mathematical methods. The risk can be measured by using the dispersion or the intermediate square deviation of individual flows to the mathematical expectation. The smaller the dispersion is, the lower the investment risk is. This process involves a comparison between the sizes of the dimension of variation period of the rate of return for different investment projects. Practice also stated another procedure that allows accurate identification of optimal investment solution, referred to as the standard deviation [10], which reflects the small deviation of the distribution of probability of the return. For example, we consider two projects, X and Y, for which there are estimated the following rates of return and probability distributions (Table 4).

Projects	Version	Optimistic	Medium	Pessimistic
X	Productivity rate	25%	15%	5%
	Probability	0,3	0,5	0,2
Y	Productivity rate	30%	20%	-10%
	Probability	0,3	0,5	0,2

Table 4. The rates of return and the probability of occurrence for investment options

For the risk assessment of the two projects we proceed as follows:

- It is determined the weighted average of the return for the two projects:

$$R_{mX} = 0{,}25x0{,}3 + 0{,}15x0{,}5 + 0{,}05x0{,}2 = 0{,}16$$

$$R_{mY} = 0{,}30x0{,}3 + 0{,}20x0{,}5 - 0{,}10x0{,}2 = 0{,}17$$

- It is determined the amplitude of the variation of the rate of return with standard deviation (table 5).

Project	Probability	Productivity rate	Average of productivity rate	Medium quadric deviation	Standard deviation
Project X	0,3	0,25	0,16	0,81	0,243
	0,5	0,15	0,16	0,0001	0,00005
	0,2	0,05	0,16	0,0121	0,00242
					0,24547
					δ =0,495
Project Y	0,3	0,30	0,17	0,0169	0,00507
	0,5	0,20	0,17	0,0009	0,00045
	0,2	0,10	0,17	0,0049	0,00098
					0,0065
					δ =0,0806225

Table 5. Determination of risk based on the spread - type (standard approach)

It is found that the lowest standard deviation is recorded in the project Y, the decision maker choosing for the realization of this project.

In the process of adopting an investment decision, the decision makers can use the breakeven analysis, the sensitivity analysis, the decision trees method, the simulation method.

Getting a higher rate of return is an objective aimed of any economic entity. Achieving this is often conditioned by the level of available resources. The insufficiency of own resources determines the use of borrowed resources, which increase indebtedness. If the economic profitability is higher than the financing costs for borrowed resources, the leverage effect is positive (otherwise the effect is negative). The manager must not omit, however, that a change in the financial structure may cause the increase of financial risk, without being possible an exact identification of the period in which insolvency may occur, especially because obtaining a rate of return greater than the cost of borrowed capital is not dependent only on the work done by the economic agent but also on the joint action of some external factors: competition, the evolution of energy prices, fuel, raw materials, the emergence of substitutes; the manifestation of some crisis with large amplitudes, the demand. Therefore, establishing a break-even point as a range rather than a fixed size seems a much more useful solution in avoiding financial risk [11].

The indebtedness and the insolvency (the insufficiency of available resources for the discharge of outstanding debt) are at the basis of the determination of insolvency and bankruptcy. These should be analyzed not only at the level of a financial exercise, because insolvency occurs when liabilities of previous period, to which we add the present and future ones, are higher than the existing availabilities and those that will be cashed in the future. To predict and avoid the bankruptcy risk it can be used a system of indicators represented by solvency ratios (these allow the sizing of the liquidity of a company), leverage, financial debt repayment capacity and the rate of financial autonomy. But, the expectation of a state of financial difficulty should not be synonymous with the liquidation of the firm. The manager can present to the creditors a plan for the recovery of the activity, and if they accept taking risks by granting new loans or rescheduling debts, the bankruptcy situation can be rectified.

In Romania, the insolvency law was adopted in the year 2006, when bankruptcy law was repealed. According to the legal provisions, an economic agent that is unable to pay his debts at maturity is obliged within 30 days from the date of recording this situation, to address the court showing that he wants to enter in the simplified procedure or reorganization one in accordance with a plan, with the to pay its debts. Not respecting the plan or the record of new losses determines the judicial administrator, the creditors committee or the special administrator to enter into bankruptcy.

From the data analysis presented in table number 6 there is a significant growth in 2006 and 2007, explained by the change of legal provisions in the field. The situation was set in 2008, but as a result of the global economic-financial crisis and fiscal measures taken by the legal authorities, the number of bankrupt companies increased considerably in 2009. For a complete vision of the situation recorded in the business environment from Romania as a result of the measures stated, we must provide other information. Thus, in 2009, 133.362 companies have temporarily suspended their activity, marking an increase of 1009,59% compared with 2008, 18766 have requested dissolution (increase of 398,83% compared with 2008) and 43616 were canceled from the Register of Commerce. Also, over 50% of Romanian companies with a total of more than 50 employees have opted for layoffs of staff, and from all these entities only 4.23% have recorded an economic growth.

Year	Number of bankruptcies	Coefficient with the base in series
2006	10431	45,19
2007	14104	35,21
2008	14483	2,69
2009	18421	27,19
2010	21692	17,76
2011	22650	4,42

Source: www. coface.ro

Table 6. Evolution of bankruptcies recorded in Romania between 2006 – 2010

An analysis of the existing data for highlighting the number of bankruptcies for the first 10 fields of activity reveals the most exposed sectors to this risk (table 7).

Field of activity	2006	2007	2008	2009	2010	2011
Retail	1935	2371	3553	3684	4178	4423
Wholesale and distribution	1909	3431	2932	3501	4262	4358
Buildings	588	1066	1666	2497	3172	3548
Transports	383	723	811	1237	1555	1596
Other main business activities in enterprises	354	625	718	979	1340	1402
Hotels and restaurants	322	520	782	1022	1275	1385
Manufacture of wood and wood products	575	810	793	927	955	866
Agriculture	1382	1093	575	934	668	686
Food and drinks	1133	1064	627	573	641	672
Manufacture of textiles, clothing and footwear	570	731	705	762	793	645

Source: www. coface.ro

Table 7. Distribution of bankruptcies recorded by sectors of activity (first 10 positions)

It is noted that the most exposed sectors are: wholesale, retail and construction, these retaining the "leadership" in almost all years of the period analyzed. In 2011 the shares held by them are of 21%, 22% and 17% (Figure 3).

The main factors that have caused major financial difficulties were: reduced possibilities of financing and aggravation of credit conditions, a significant reduction in consumption, with a direct impact on the demand of goods and services, the establish of new tax liabilities, products prices increases, particularly in the energy, fuel and raw materials level, the excessive bureaucracy, lack of an incentive package for the development of the business area, difficulties in the mechanism of absorption of community funds, the examples could continue.

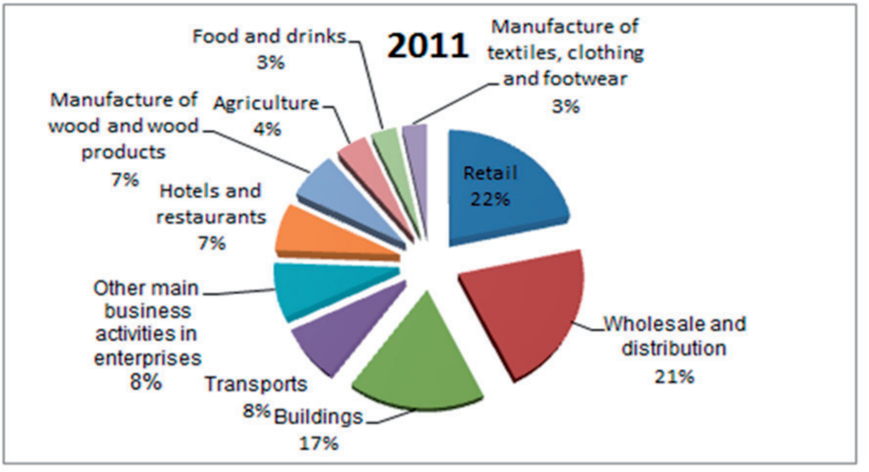


Figure 3. Structure of bankruptcies on sectors of activity in 2011

5. The system of indicators that allow the highlight of the economic-financial performances of the economic entities and the substantiation risk management decisions

Maximizing the efficiency of the activities performed at the firm level requires a constant monitoring of the financial performances. In other words, to optimize economic performances is necessary to implement a financial diagnostic system of the performed activity, resulting in a series of' documents that include a set of indicators requested by the management team. It is obvious the fact that the efficiency of the system is determined by the type of documents requested, by the way in which cooperation takes place between different components of the company but also by the technology used in the development of the reports.

To estimate the financial status of a firm there can be used use a series of indicators that allow highlighting the profitability, solvency, liquidity, turnover of debts and the degree of financial independence. The activity of financial diagnosis of a firm can be made to determine the situations of financial disequilibrium and of the generating causes, so that there can be adopted in time the necessary measures for effective management. In this case, are involved the managers, shareholders and the employees. However, the diagnosis may cover the company's ability to make profit, to record surplus flows in comparison to the negative ones due to the activity performed, to honor payment obligations, cases in which it is performed by persons outside the firm (banks, potential investors, customers, competitors, state). But, even if it concerns an internal or external analysis of the activity undertaken, the objectives followed are the assessment of economic and financial performance and identification of the potential risks that may occur.

But a real analysis involves not only determining the state of the financial balance and liquidity of a firm but it should aim to identify the margins of cash accumulation. To this purpose, there can be used the information provided by the profit and loss account statement which summarizes the outcome of economic and financial flows of entry, the combination of production and output factors over the period considered. Basically, this document creates a group of incomes and expenditures which have determined the overall result for the ended financial year and allows assessment of the overall economic and financial performances achieved by the firm by using some performance indicators. Known as intermediate management account balances these indicators are underlying preparation of financial statements analysis and forecasting, providing managers the possibility to identify deviations from the predictions made by categories of activity. Thus it can follow the degree of turnover achievement and the anticipated gross margin, carrying out the operating expenses in limits set, the impact on the indebtedness degree on current outcome, obtaining the estimated profit.

Sizing the economic and financial performances is based on the diagnostic analysis, which aims to identity responses to a series of questions related to the results obtained, how they were obtained, the levels recorded compared with those projected, of the measures that can be adopted to achieve the objectives, both short and medium term and long term. The diagnosis-analysis is based on synthesis accounting documents (the situation of assets, liabilities and equity - balance sheet, financial results - income statement and attachments) and other information on the evolution of prices and exchange rates [13].

The sizing activity of the financial and economic performances should be carried out continuously and not only in difficult situations, as Jean- Pierre Thibaut stated “Performing some analysis is motivated not only whether the company has difficulties, but also when it wants to improve the results obtained”. Practically, realizing continuously some analysis of the financial statements concerns the achievement of information necessary for taking decisions by managers in a rational manner.

The analysis made must be supplemented by a series of liquidity indicators such as the overall liquidity, immediate liquidity, ability to pay rates, coverage degree of daily expenditure, the rotation period of assets, debt indicators, namely: indebtedness, the degree interest coverage, total debt ratio, leverage, general solvency ratio, management indicators (duration of stock rotation, customer flow, credit suppliers, non fixed, fixed and total assets) and other indicators of profitability.

Whatever the type of indicators used, the decision maker must consider a number of issues related to the need to supplement the fixed analysis with dynamic analysis (permanently there occur changes which can affect the results recorded at a moment in time) but also the use of a complex system of indicators (making analysis based on a single indicator may lead to misinterpretation) which are able to show clearly and accurately the financial and economic performance level of the company.

The activity done by any economic entity can be structured at the level of three fundamental components: operating, financial and extraordinary (Figure 4). The manager must identify the potential risks for each component so that it can base an effective strategy.

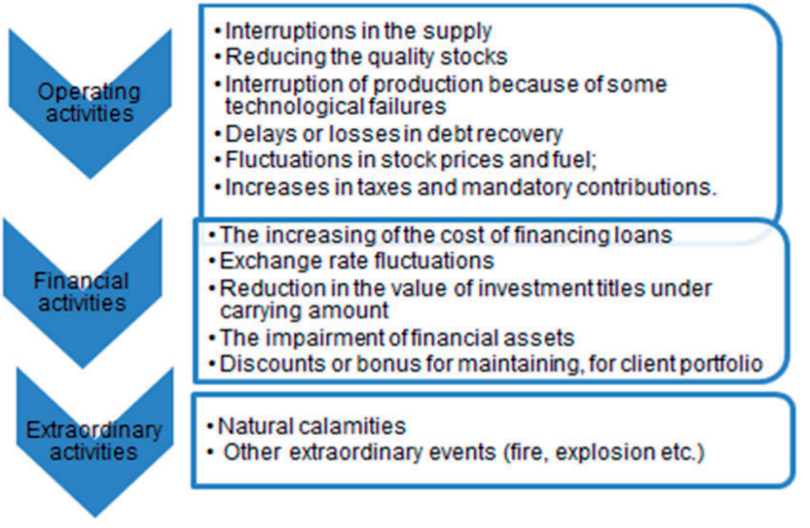


Figure 4. Structure of developed activities and risks afferent

Obviously, is not enough an overview of these risks, without a detailed analysis of the causes. For example, the risk, "cutoffs in the supply "can occur from multiple causes. Thus, the supplier fails to comply with the contract clauses regarding the delivery terms or quantities established , may occur accidents during transport, the quality of goods supplied is not according to the requirements established, which leads to their refusal by the buyer, there take place extraordinary events (fires, floods, earthquakes), etc. Also, these cutoffs may have a short, high or very high length, with direct implications on the production process, or, as appropriate, in marketing (minor, significant or major). The analysis of these issues should be completed with the assessment of the losses caused by the manifestation of the mentioned risk. Another example can be the one of the risk of increasing taxes and fees. The causes in this case are external, but applying the principles of fiscal management can only have a favorable impact on the financial performances of the company.

Risk management should be based on an analysis of indicators highlighting economic financial performance at the level of the activities shown in figure number 5. But it requires a detailed breakdown of the indicators presented in order to identify the contribution of each component in recording a result or another, being not enough to follow-up the synthetic level. In addition, if not takes into account also their dynamics and not is not corroborated with factors of influence and circumstantial situations, it is not possible to determine the causes that generate negative effects and no anticipation of any difficult times.

The main components that are at the base of the formation of the gross result, are presented in Figure 5.

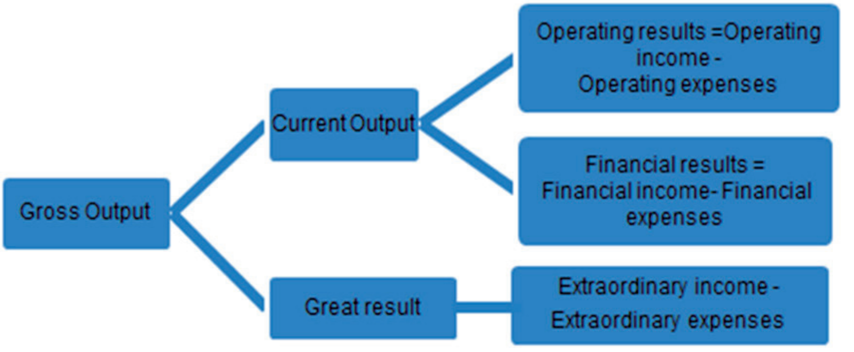


Figure 5. The formation of the result of activity performed broken down by components

The operating result is decisively influenced by the level of income and expenditure obtained. As a result, the manager must seek firstly the dynamics and the structure of sales. The risk of reducing them may be caused by:

- emergence of new competitors;
- reduced consumption due to the manifestation of strong economic and financial crisis;
- insufficient promotion of products;
- emergence of similar products at lower prices, on the basis of labor productivity growth at the level of competitors;
- punishment by consumers low quality level of products supplied;
- existence of a structure based mainly on marketed products which determine lower revenues.

At the same time, it is necessary to identify the impact of the costs upon the operating result. Thus, an increase in costs can be determined by:

- rising prices of raw materials supply, fuel and energy;
- record of high specific consumption per product unit;
- technological losses;
- low labor productivity;
- use of technologies worn out, etc..

But, a great importance is had not only by the determination of the indicator stated and by the analysis of its components, but also by the investigation process from the dynamic point of view because it can provide additional information on level of performance correlated with potential generating causes. In this respect, it is necessary to calculate the indexes that diminish the evolution of operating incomes and operating expenses and their comparison (Table 8).

We consider that the income index is calculated as ratio between the income from current period and the income obtained in the previous period. We use for comparison the index from total incomes and the index from operating activity.

Cost index	Method of calculation	Index compared with income	Effects	Causes
Index of total expenditure	$I_{CT} = \frac{CT_1}{CT_0} x100$	$I_{CT} > I_{VT}$	Economic inefficiency: the consumed loss-monetary unit Bankruptcy	• Imbalance in the components • Inappropriate management of the entire activity
Index of operating costs	$I_{CE} = \frac{CE_1}{CE_0} x100$	$I_{CE} > I_{VE}$	Diminishing profit	• Inefficient consumption in comparison with the efforts
Index material costs	$I_{CM} = \frac{CM_1}{CM_0} x100$	$I_{CM} > I_{QV}$	Profit reduction amid the inefficiency of materials consumption	• Wrong choice of suppliers • Weak report price/performance compared to other suppliers • Faulty delivery conditions • Improper management of material elements • Inappropriate quality of the materials • Losses in the technological process; • High specific consumptions; • Inappropriate storage of materials.
Index of personnel spending	$I_{CP} = \frac{CP_1}{CP_0} x100$	$I_{CP} > I_{CA}$	Ineffective personnel costs	• Remuneration uncorrelated with effort (inefficient relation between wages and productivity); • Low level of staff training compared with the requirements of the position held; • Increases of the expenses with insurances and social protection;
Index expenditure taxes	$I_{CIT} = \frac{CIT_1}{CIT_0} x100$	$I_{CIT} > I_{CA}$	Reducing the capacity of an auto financing in the future	• External capacity that cannot be controlled by the manager; • Increases of the taxes or establishing additional fiscal obligations

Table 8. Effects and causes of the evolution of revenues and operating expenses

Analysis can be extended to the other components. Thus, if the index exceeds the financial cost of financial income, there is an inefficiency of the funding policy (table 9).

Cost index	Method of calculation	Index compared with income	Effects	Causes
Index of financial expenses	$I_{CF} = \frac{CF_1}{CF_0} x100$	$I_{CF} > I_{VF}$	Reducing available resources	•Us of mostly borrowed resources in comparison with the own ones; •Inefficient use of the financial resources; •The increase of interest rate for the loans contracted; •Recording some unfavorable differences for the exchange rate; •Disposal of investment securities at a lower values than the book value; •
Index of extraordinary expenses	$I_{CEx} = \frac{CEx_1}{CEx} x100$	$I_{CEx} > I_{VEx}$	Reducing the gross profit of the year	•Natural disasters

Table 9. Effects and causes of the evolution of financial and extraordinary revenues and expenses

Causes may include the prevalence compared with their borrowed resources; increase the interest rate on loans, recording a difference of an adverse exchange rate, yielding investment securities at a lower value than the book, etc.

If the operating result and the financial one are positive, there is a state of, "financial health" in the company, which, combined with a shorter duration of the credit customer compared with the credit provider (the collection of receivables is achieved at a shorter interval than the payment of liabilities), determines a state performance. Situations that can occur in practice are presented in table 10.

In this connection, we draw from the research conducted a series of ideas to be considered of being in risk management:

- Determination of indicators is based on the information recorded at a time (the end of a financial year), without taking into account the dynamic changes recorded;
- Analyzing the performance indicators that do not allow forecasting famous risks of integrated business, respectively the dependence on suppliers and customers;
- Distortions in the functionality of the production, supply or sale may adversely affect the overall activity performed, although a past performance can mark a positive situation;

- Ensure the ability to pay and the cash required is determined by the system performance indicators, a high level of profit recorded in the profit and loss does not always correspond to the actual cash flow - so, large companies faced with situations such as the current way of calculating the profit is based on elements whose achievement is not always certain (for example, claims that cannot be collected due to customer bankruptcy or financial bottlenecks) or equivalent of non-specific economic flows of the financial year.

Results recorded	Generated financial situation	Measures imposed
Re>0 Rf>0	Financial performance	Continuing the activity in the same direction, in accordance with the action plan set.
Re>0 Rf<0 Re>|Rf|	Good financial situation	Reducing financial expenses(use of a financing structure which determines the lowest cost of obtaining resources)
Re<0 Rf<0	Difficult financial situation	Reorganizing the activity
Re<0 Rf>0	Strategic situation	Reducing at a minimum level the interval in which are recorded these ratios

Table 10. Situations determined by the evolution of the components of the current result of the year

6. Using the discriminating statistical models in determining of financial and the financial condition and the risk analysis

The financial condition of an economic agent can be considered a qualitative indicator that characterizes the overall activity of the economic agent, is subjected continuously to the influence of quantitative variables. This allows the use of statistical discriminant analysis to determine the financial status of an economic agent (in-depth presentation of this method and of the models developed on its basis was performed by Anghel [14]. The best known method from the specified category is the scores method. For variables taken into account are determined certain weights so that their sum reflects a global indicator called" The Z score ". This method involves the building of a database with information on a group of bankrupted companies and a group of healthy companies that allow calculation of financial ratios for each entity subject to the analysis and the determination of the best combination of rates that allow the differentiation of the companies analyzed. Basically, the indicator resulting (statistical discriminant model) reflect the overall situation of the company with manifest and predictability. One of the first score functions used in the analysis of default risk was developed by Altman in 1968, which allows the registration of a degree of predictability of 75% of bankruptcies two years prior to their production [15] . The model was further developed to be applied in all branches. Altman model is frequently used in financial practice and enable a correct the result in the classification of 70% [16].

It should be mentioned also other important models: the Conan – Holder model [17], the model of the Central Balances of the France Bank or the method of the French Commercial Credit. In the Romanian school, the possibility to determine this pattern was limited because the transition from the centralized economy to the market economy was quite large and the information required was not available or relevant. However, the significant concerns embodied in the formulation of the scoring models have many authors [18 -20, 14]. The application of scoring functions in the Romanian economy cannot yet be considered as a safe situation of risk management. We believe that the models that have demonstrated the applicability of the developed economies do not correspond to an economy characterized by high instability. In addition, the proper identification of the bankrupt companies, the long absence of the legislative provisions governing the bankruptcy, the lack of analysis of a wider range of firms before bankruptcy, focusing on financial variables without including non-financial indicators and trying to develop the models available in all branches reflect some of the aspects that require the completion of the risk management techniques based on a discriminant analysis with other possibilities.

7. Conclusions

The manifestation of risks in a higher or lower degree cannot be eliminated, regardless of the strategy used. However, in practice it has been shown that a potential risk with higher negative impact but which has been identified and controlled can cause fewer losses than a lower but uncontrolled risk. Therefore, a good manager should use a risk management strategy in order to enable the reduction or even avoidance of losses caused by their manifestation. A proactive risk management can generate durable benefits for the company, materialized wither in costs reduction or in improving or making internal processes efficient, or in canceling insurances expenses, which are not justified from economic point of view.

The problems presented in this work are due to identify the issues that highlight the importance of active risk management, to optimize the ratio of the risk level –the level of the profitability achieved, the need to monitor and update the information on risk and the possibility of realization, the role of risk management in identifying the opportunities of the business determines the highest level of profitability in accordance with the accepted risk level and the amount of available resources, the proactive risk management capacity generates sustainable benefits for the company, resulting either in reducing the costs or to improve and streamline of the internal processes or the elimination of the insurance costs, economically unjustified, the need of training the decider as a person capable to take the increased costs due to the running of a business, whether to obtain a corresponding reduction in risk and thus of the losses that would be determined by their expression ; the attraction of the approach risk management to empiricism in the application of techniques with clear and measurable objectives.

The application of techniques and methods of risk management must be a starting point for substantiating and applying the company's financial strategy. The course of action must be related to identifying the elements that may cause the manifestation of risks, sizing the probability of their manifestation, forecasting channels of influence and establishing means of protection.

It is not advisable to use only statistical models to determine the financial status (statistics entails the use of previously recorded data, and the maintaining of a constant level of indicators calculated on their basis is unlikely in an environment subjected continuously to profound transformations at short time intervals) but is required their completion with analysis reported to financial equilibrium, level of profitability, the ability to pay, the changes in the financial flows so as to enable an evaluation of the global risk and to apply an optimal management of it.

Taking into consideration the issues presented in the paper it resulted: the importance of active risk management, to optimize the degree of risk assumed - the level of profitability achieved; the need to monitor and update the information on risk and the possibility of realization; the role of risk management in identifying business opportunities that lead to the highest level of return, consistent with accepted risk level and the amount of available resources; the capacity of proactive risk management to generate sustainable business benefits, materialized either by cost reduction or improvement and streamline of internal processes or the elimination of expenses with insurances, economically unjustified; the necessity to train the decision maker as a person able to undertake increasing the costs of running a business, whether obtaining a corresponding reduction in risk and thus of the losses that would be determined by their expression; attracting risk management from the area of the empiricism into that of a clear approach and the application of techniques with clear measurable objectives.

Author details

Gabriela Dobrotă
Faculty of Economics and Business Administration,
Constantin Brâncuşi University of Târgu Jiu, Târgu Jiu,
Romania

8. References

[1] Stancu I (2002) Finance. Bucharest: Economic. 867-868 p.

[2] Dobrotă G, Paliu Popa L, Babucea A.G, (2010) Risk management at the level of Economic Agents, whitin the context of Sustainable Development, Recent advances in risk Management, Assessment and mitigation, In Proceedings of the International Conference, RIMA'10. Bucharest: WSEAS Press. pp 380-386.

[3] Andronic C.B (2000) Firm performance. Iași: Polirom. 157 p

[4] Giarini O, Stahel W.R (1966) Certainty limits. Bucharest: Edimpresss- Caruso. 255 p.

[5] Hofstede G (1996) The management of the multicultural structures. Bucharest: Economic.136 p

[6] Sonigo P (1994) Risk management. Risques: 17: 131-136

[7] Williamson E.O, Winter G.S (1997) The nature's company Timișoara: Sedona. 191 p.

[8] Cocriș V, Ișfan V (1997) The economics of business. Iași: Graphix.118 p.

[9] Dobrotă G (2009) Financial management of the company. Theory and applications. Craiova: Scrisul Românesc. 111 p.

[10] Baicu M (2010) Risk Management in Business Bucharest: Foundation România de Mâine. 178 p.

[11] Prunea P (2003) The risk in economic activity. Hypostasis. Factors. Modalities to reduce Bucharest. Economic. 110 p.

[12] Dobrotă G, Chirculescu F (2008) The economic and financial performance of the enterprises. Annals of the Constantin Brâncuși University of Târgu Jiu, Economiy series.2: 93-102.

[13] Avare Ph, Legros G, Ravary L, Lemonnier P (2006) Management and Financial Analysis Bucharest: Economic. 24 p.

[14] Anghel I (2002) The bankruptcy. Radiography and prediction. Bucharest: Economic. 61 p.

[15] Altman E.I. (1968) Financial ratios, Discriminant Analysis and the predictin of corporate Bankrupcy. Journal of Finance. 23(4): 589-609.

[16] Stancu I (2007) Finance. Bucharest: Economic. 787 p.

[17] Conan, Holder, (1979) Variables explicatives de performances et controle de gestion dans les P.M.I., Universite Paris Dauphine.

[18] Mânecuţă C, Nicolae M, (1996) Building and Using score function for the diagnosing the efficiency of economic agents. Finance, Credit and Accounting. 5, 47-54.

[19] Băileşteanu Gh, (1998) Diagnosis, risk and efficiency in business Timișoara: Mirton. 294 p.

[20] Ivoniciu P (1998) The analysis of bankruptcy risk by the method scores. Finance, Credit and Accounting. 4: 17-19.

Chapter 12

Integrated Risk Management System – Key Factor of the Management System of the Organization

Emilia Vasile and Ion Croitoru

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/47883

„Life itself is a risk"

1. Introduction

The management of any organization, whether working in the public sector, whether working in the private sector, aims in order to achieve its objectives to monitor and reduce risks. Risk control is achieved by managing them effectively, namely by implementing an adequate risk management system.

Risk management is an important concept related to safety and financial integrity of an organization, and risk assessment is an important part of its strategic development.

The strategy of an organization on risk management should be that all the risks it faces must be identified, assessed, monitored and managed so that they are maintained in a certain limit, accepted by the entity's management.

2. Risk management – Defining function within the organization

Risk management is the process of identifying, analyzing and responding to the risks the organization faces and is exposed to. The costs of implementing this system depend on the methods used to manage unexpected events.

Risk management process is an ongoing one and the results are embodied in the decisions taken on accepting, reducing or eliminating risks that affect the achievement of objectives. The aim is to optimize the organization's exposure to risk in order to prevent losses, avoid threats and exploit opportunities.

2.1. Conceptual approaches for risk

In general terms, risk is part of any human effort. Once we leave to go back home, we are exposed to risks of different levels and degrees. It is significant that some new risks are completely voluntary, and some are created by us through the nature of activities.

The word *"risk"* derives from the Italian word *„risicare"*, which means *"to dare"*. In this sense, *the risk is a choice, not fate*[1]. From this definition it follows that the risk is not an option, but we are permanently exposed to risk in everyday life, what is really important is that each time, to gain control over it.

Nowadays there is no unanimously accepted definition of the concept of risk by all specialists in the field. Among the most commonly used definitions, we present the following:

> *"Risk is the possibility of obtaining favorable or unfavorable results in a future action expressed in terms of probabilities."*

or

> *"Risk is a possible future event whose production could cause some losses."*

or

> *"Risk is the threat that an event or action to affect in a negatve manner the capacity of an organization to achieve its planned goals.*[2] *"*

The analysis of these definitions of risk gives rise to the following conclusions:

a. Probability versus consequences. While some definitions given to risk focus only on the probability of the occurrence of an event, other definitions are more comprehensive, including both the probability of risk manifestation and the consequences of the event.
b. Risk and threat. In defining the concept, some experts have put an equal sign between risk and threat. We specify that a threat is an event with a low probability of manifestation, but with high negative consequences, since the probability of manifestation is difficult to assess in these cases. A risk is an event with a higher probability of occurrence, for which there is sufficient information to rate the probability and consequences.
c. Comparing only negative results. Some concepts about risk are focused only on negative events, while others take into account all variables, both threats and opportunities.
d. Risk is related to profitability and loss. Achieving the expected result of an activity is under the influence of random factors that accompany it in all stages of its development, regardless of the domain of activity.

In conclusion, the risk can be defined as a problem (situation, event etc.) which has not yet occurred, but can occur in the future, threatening the achievement of agreed outcomes. Viewed in this context, risk is the uncertainty in obtaining expected results and should be treated as a combination of probability and impact.

The probability of risk occurrence is the possibility that the risk materializes and it can be appreciated or determined by measurement, when the nature of risk and available information permit such evaluation.

The risk impact is a consequence of the results (objectives) when risk materializes. If the risk represents a threat, the consequence upon the results is negative and if the risk represents an opportunity, the consequence is positive.

The probability of risk occurrence and its impact on the results contribute to establish the risk value.

Based on concepts presented above, in our opinion, the risk is a permanent reality, an inherent phenomenon that accompanies all activities and actions of an organization and that occurs or not, depending on the conditions created for it. This could cause negative effects by deteriorating the quality of management decisions, reducing profit volume and affecting the organization's functionality, with consequences even in blocking the implementation of activities.

In the literature, but also in practice, besides the concept of risk other concepts are used, respectively:

Inherent risk is the risk that exists naturally in any activity and is defined as *"the risk existing before* the *implementation of internal control measures to reduce it"* or *"all risks that threat* the *entity/organization and may be internal or external risks*, *measurable or immeasurable".*

Residual risk is the risk remaining after implementation of internal control measures. Applying these measures should have as effect the limitation of inherent risk to a level accepted by the organization. The residual risk should be monitored in order to maintain it at accepted levels.

Risk appetite is the level of exposure that the organization is prepared to accept, namely the risk tolerated by the organization.

Practitioners recommend to organizations' management to bear in mind that risks can not be avoided and under these conditions to be concerned by their evaluation to keep them "under control" at levels considered acceptable, tolerated by the organization, and not to seek the total elimination of them, as this can lead to other unexpected and uncontrolled risks.

2.2. Risk – Threat and opportunity

Internal and external environment in which the organization operates generate risks. In these circumstances the organization should identify its weaknesses and threats it faces, in order to manage and minimize them. Also, strengths must capitalize and exploit opportunities.

In this respect, designing and implementing a risk management process at corporate level, is appropriate and necessary due to uncertainties of threats in achieving organizational goals.

The implementation of this concept leads to certain changes within the organization, whose effects should be materialized through a better use of available funds and obtaining levels of profitability planned, namely:

- *risk management requires modifications in leadership style*, the organization's management would be forced besides the consequences treatment measures of events that occurred, to devise and implement adequate internal control devices to limit or eliminate the possibility of risks manifestation. Implementing these control devices should enable the organization to master, within acceptable limits, risks and to achieve the objectives.
- *risk management ensures the efficient and effective achievement of objectives*, mastering threats the organization deals with, allows to hierarchy risks based on materialization probability, of impact magnitude and costs posed by mitigating or limiting unwanted effects.
- *risk management requires a healthy internal control system*, designing and implementing adequate internal controls and ensuring their operation require a reasonable assurance that objectives will be achieved. Enhancing and strengthening the internal/management control system is indispensable without designing and implementing appropriate risk management.

Risk management is characterized by the establishment and implementation of concrete activities and actions of identification and risk assessment leading to determine the risk level and by this act to implement adequate internal control devices to limit the probability of the risk occurring or the consequences if the risk materializes. The process must be coherent, integrated to the objectives, activities and operations carried out within the organization.

The staff within the organization, regardless of the current hierarchical level, should be aware of the importance of risk management to achieve planned results and to form necessary skills in order to perform monitoring and control based on principles of efficiency and effectiveness.

The functional structures responsibles within the organization have the task to identify and analyze regularly the risks related to their activities, to propose and substantiate appropriate measures in order to limit the possible consequences of risks and ensure approval by decision makers within the organization.

Practice[3] recommends that any organization needs to manage its risks, because in many cases the occurrence of risks can have serious consequences upon the activities, sometimes these consequences jeopardizing the very existence of the organization[4].

The complexity of risks and their increase has led organizations' management to understand that it is better to manage a risk than to cover a loss. Based on this requirement, many organizations have proceeded to implement risk management, developing specific strategies that have defined the organization's behavior towards risk and risk management arrangements.

2.3. The importance of risk management organization

Risk management is a preventive attitude on the elimination or limitation of damages, if any possibility of a risk materializing, namely a process of identifying, analyzing and responding to potential risks of an organization.

In these conditions, the role of risk management is to help understand the risks the organization is exposed to, so that they can be managed. This role varies depending on when the analysis is done, as follows:

- if the risk assessment is conducted before the risk materialization, the goal is to avoid the occurrence of this event;
- if the risk assessment is carried out after the risk has materialized, the goal is to ensure the development of the activities and the organization's activities continuity.

The advantage of implementing the risk management system within the organization is to ensure economic efficiency. To achieve this requirement, the organization's management has the responsibility to make known the risks they face and manage them properly, in order to avoid consequences for their materialization.

2.4. Responsibility for risc management

Risk management is the responsibility of the organization's management, and the central objective of this process aims the risks management so that resources to be used efficiently and effectively in order to maximize profit and minimize threats, while safeguarding the interests of employees and customers.

In this respect, the entity's management must act in the following directions:

- establishing the definition of risk that is widely accepted and understood across the organization and also the types of risk;
- assessing current risks and monitoring potential sources of internal and external risks;
- establishing clear responsibilities on each hierarchical level and per employee concerning the implementation of risk management process;
- developing an adequate information system for the management on risks and risk assessment system;
- setting tolerance in taking risks and limits of exposure in accordance with it;
- permanent analysis of achievements and poor results in risk management and continuous improvement of risk management process;
- ensuring an adequate level of knowledge and skills of employees in accordance with the requirements followed by the implementation of the risk management process.

To ensure an efficient risk management is necessary to create certain organizational structures appropriate for the policies and strategies of the organization. In this respect, the organization should adopt appropriate policies regarding the organization plan, in order to effectively monitor each risk or category of risk and in an integrated manner, the whole risks system accompanying activities.

Policies and strategies that may be adopted regarding the organization plan are related to:

- establishing and developing its own system of rules and procedures that, implemented, to ensure avoiding or minimizing risks;
- establishing appropriate functional structure based on a clear concept, which should ensure appropriate departments in order to contribute at identifying and monitoring risks.

Given that risk can be identified, evaluated and limited, but never completely eliminated, the organization must develop both general policies and specific policies to limit exposure.

2.5. Effectiveness of risk management

The activity of an organization is characterized by all processes, procedures, inputs, outputs, resources (financial, material, human and informational) and technical means for recording, processing, transmitting and storing data and information on activities and environment where the system is operating.

By internal/management control programs prepared each functional structure should identify the risks they face, and by using procedures and risk management policies to ensure their maintenance at acceptable levels.

Risk management is an ongoing, structured process, that allows identifying and assessing risks and reporting on opportunities and threats affecting the achievement of its objectives. The benefits of implementing the risk management process include:

- higher probability of achieving the entity's goals;
- improving the understanding of risks and their implications;
- increased attention to major issues;
- limitation to the consequences by implementing adequate internal controls;
- assuming a certain tolerance to risk is acceptable;
- broader information for adequate decision making in terms of risks.

The organization's management and staff perform risk management activities in order to identify, assess, manage and control all types of events or situations that may affect its activities.

In the world today has become increasingly more imperative for corporate managers to monitor and manage risk[5] in all aspects. A good risk management means avoiding or minimizing loss, and also treating opportunities in a favorable manner.

Risk management is necessary because organizations face uncertainty and the biggest challenge of the leadership is to determine what level of risk it is prepared to accept to achieve its mission, in order to add value to activities and to achieve planned goals.

Risk management is an essential component of the organization's success and must become an intrinsic part of its functioning. It must be closely related to corporate governance and internal control, but also connected with performance management[6].

3. Integrated approach to risk

Integrated risk management process is designed and set by the management and implemented by the whole staff within the organization. This process is not linear, a risk management may have impact also on other risks, and control devices identified as being effective in limiting a risk and keeping it within acceptable limits, may prove beneficial in controlling other risks.

Risk management currently knows an appreciation and recognition increasingly large, both in theory and practice, which means, on the one hand the increase of number of specialists in the field, and on the other hand the interest of managers within organizations to design and implement effective risk management systems to meet the objectives.

Mastering risk determines organizational development, performance growth, both generally, of the whole organization and also of individual activities.

3.1. COSO and integrated risk management

Referring to risk management, COSO presented an initial framework methodology for implementing internal controls, built-in policies, rules, procedures and regulations that have been used by various organizations to secure control over how to run the plan and meet objectives.

Later, after the appearance of great scandals of fraud and the need to improve corporate governance processes, large corporations talked about and set up risk management departments to help implement procedures regarding the identification, assessment and risk control.

Following the emergence of these needs, Treadway Commission, COSO model promoter, initiated a program in order to develop a general methodology that can be used by organizations' management to improve risk management.

Risk management within the organizations was created on the concept of internal controls, but the focus was particularly on risk management. This was not intended to replace internal controls, but incorporating basic concepts of internal control in this process.

Thus, between risk management and internal control was preserved a strong connection interrelated with common concepts and elements.

3.1.1. Risk management and internal control

The main objectives of internal control/management system are to ensure the efficiency and effectiveness of activities, the reality of reporting and regulations compliance in the field. The internal control/management system is developed and monitored in order to implement by the organization's management, which is responsible for designing adequate internal control devices in order to ensure limitation of significant risks and keeping them within acceptable limits, aiming to give the security that the organization's objectives will be met.

Risk management system was structured on components of internal control/management, structured according to COSO model, namely on five elements, whose implementation ensures that the tools/internal control devices exist and function as intended.

These components were defined as:

- *the control environment* specific to the organization is the one that sets the foundations of internal controls system, influencing the control awareness of employees and represents the basis for other components;
- *risk assessment* is carried out by management, is performed at both corporate and activity level and includes identifying and analyzing risks that affect the achievement of objectives. In general, risk assessment involves determining the level of importance of the risk, assessing the probability that the risk to occur and determining the way to manage it;
- *control activities* are policies and procedures to ensure that management's provisions are respected. By this, it is ensured that all necessary measures are taken in order to manage risks and achieve the objectives set by management;
- *information and communication* helps other components through proper communication to employees of their responsibilities with regard to internal control and provision of relevant, reliable, comparable and understandable information so that they could perform their duties and tasks;
- *monitoring* implies the verification made by the management of the implementation means of internal controls it demanded, or by responsibles pursuing if internal controls imposed by it work and if they are sufficient so that activities or actions to take place as planned.

3.1.2. Objective of risk management system

COSO defines integrated risk management as *"the process conducted by the Board, management and others, applied in setting strategy and across the organization, designed to identify potential events that may affect the entity and to manage risk within the risk appetite to provide a reasonable assurance regarding the achievement of organizational objectives"*[7].

From the content of this definition it follows some essential elements, characteristic to the integrated risk management, as follows:

- the process is conducted permanently throughout the organization, being circumscribed to other activities;
- the purpose is to manage risks associated with objectives and to secure expected results through their implementation;
- within the process is involved the whole staff, regardless of the hierarchical level;
- the approach starts from the strategic goals rather than from operational objectives;
- the process is applied to the entire organization and not functional structures.

The general objective of integrated risk management is to effectively manage uncertainties, risks and opportunities. The need for risk management stems from the fact that uncertainty is a reality and the reaction to uncertainty is a constant concern.

Risk management involves establishing actions to respond to risk and to implement adequate internal control devices, with which to limit the possibility of occurrence or consequences of risk, if it would materialize. In order to ensure efficiency in achieving objectives, the process must be coherent and convergent, integrated to objectives, activities and operations carried out within the organization.

Also, regardless of the staff's hierarchical level, it should be aware of the importance of risk management has in achieving its own objectives and thus to form the necessary skills to perform monitoring and control based on principles of efficiency and effectiveness.

In order to ensure the success of this approach and to achieve an effective risk management, within the organization it needs to create a culture of risk, namely developing a risk management philosophy specific to the organization and management, and awareness of risk's negative effects at all levels of the organization.

From the above it is found that the need for internal control/management is determined by the existence of threats or opportunities in carrying out planned activities or actions with negative consequences in the organization. This requires the establishment and implementation of certain internal control devices in order to prevent or limit the risks.

Also, the need for risk management stems from the fact that risk is everywhere, in everything we want to achieve. It can not be removed; any action to eliminate risk can lead to the emergence of new risks, uncontrolled, which may affect to much greater extent the organization. In these conditions, the risk needs to be minimized, process that can be achieved by establishing and implementing adequate internal controls.

3.2. The role of integrated risk management system

Risk management process is considered to be a set of activities and actions carried out in a certain manner and order to prevent or reduce exposure to risk, resulting from an operation or several operations.

In practice, most commonly applied concept of risk management is that managing risks should be carried out separately within departments independently organized in the organization's functional structure. This method provides simplicity and efficiency form in making decisions on risk management, but leads to actions and multiple records of the same exposure to risk and does not address correlations between different exposures.

There are other practices too, which considers that each employee must be responsible for the risk management, having the competence to identify risks and implement appropriate internal controls to mitigate the probability of their manifestations. This mean of managing risks does not lead to results and does not ensure the guarantee of conducting activities given that they were planned, because it does not ensure the requirements for exposure on the same activities, and the process is influenced by knowledge and understanding by employees of the risk management system implemented within the organization.

These traditional risk management processes are usually fragmented, meaning they are found implemented at the operation or transaction level and are aimed at preventing losses.

Managing risks in these cases *"does not consider the fact that risks are a source of competitive advantage".*

Recent research on models and risk management strategies focus on competitive advantages of risks if they are approached as a whole or at system level. In this case the system is considered to be composed of all processes and activities necessary to achieve the objectives.

This approach requires that all relevant functions within the organization (personnel, finance and accounting, manufacturing, commercial, procurement, IT, legal, internal control, internal audit, strategic development, marketing etc.) to participate in risk management process.

For implementing the integrated risk management is necessary that the organization to be viewed from the standpoint of system, both as the link of the industry in which it operates and as part of it, acting in accordance with certain principles, features being: the complexity, limitation of resources, factors that influence its activity, the nature of events, the possibilities for development.

In this view, it is considered that the risks should be managed in an integrated way, to eliminate multiple records on the same risk exposure and to analyze correlations between different exposures. This risk management approach is complex; it requires a large volume of information necessary for decision making and higher costs of administration. At the same time, making wrong decision can have a high impact on the business, or even on the organization.

The integrated risk management system, based on this concept, must be interdependent with the organization's development needs and to include the processes of development and establishment of elements concerning assessment, monitoring and risk management. At the same time, integrated risk management must be also approached in correlation with all types of risk management for each functional structure of the organization.

Integrated risk management system operates with broad categories of risk (personnel risk, financial risk, legal risk etc.), with different risks attached to various activities, risks associated with different operations or transactions, and also with external risks that may affect the development of the overall organization (risks related to legislative changes) or making one or more activities carried out within the organization.

In these conditions, implementing the concept of integrated risk management within the organization is more than necessary because the risk management process should be approached by all types of risk that are found and affect all functional structures of the organization.

The approach in this unitary manner, of the exposures, respectively as a righteous and coherent system of exposure to various risks, of connections and mutual conditioning between them, will enable effective management of risks that may affect achieving the objectives and will contribute to improve activities and performance growth within the organization.

The integrated risk management system can identify all risks that affect the implementation of processes and activities attached to an organizational goal; it can assess the overall consequences and adopt measures depending on the level of uncertainty and the existing inherent risk that affects achieving objectives set.

Also, integrated risk management allows the foundation and decision making to lower hierarchical levels of the organization and also at the top level and ensures co-ordination of activities in order to solve current problems between certain functional structures. It helps to increase efficiency within the organization also by others administrative or managerial ways, such as better allocation of resources.

The implementation of integrated risk management within the organization will provide to shareholders and potential investors, more concrete and reliable information on the risks to which it is exposed, which will allow them to base their decisions in more optimal conditions.

Once with the development of organization's activities, the old risk management systems become inadequate and risk exposures, especially the risk of fraud and error increases significantly. Implementing the integrated risk management system involves the design of evaluation criteria capable of measuring all activities related risks, by considering the relationships and connections between them and thus, to determine the exposure to any organization's risk factor or its functional structures at any time.

This risk management process, characterized by the development of integrated risk management methodology, shall include as steps: establishing the organizational context and risk management, identifying, analyzing and assessing risk, risk treatment, risk control, communication and monitoring the risk management plan.

The process should not be a linear, the risk management may impact on other risks, and measures identified as being effective in limiting a risk and keeping it within acceptable limits may prove beneficial in controlling other risks.

3.3. Integrated risk management system functions

The effectiveness of implementing an integrated risk management system, compared with traditional risk management, is determined by the fact that it reflects the integration of all activities related to risk and risk management in a single system. This system is operated and controlled from a single management level, thus eliminating duplication and disruption of communication and action that can occur within a classical system.

The functions that the integrated risk management system meet within the organization's management system can be classified as follows:

a. *defining goals and setting objectives of the organization on risk.* Setting goals represents a defining requirement for the identification, assessment and risk response planning. The organization must define properly its objectives, so to be understood and carried out by people who were assigned to.

The basic role of integrated risk management is to provide to the management and organization's board a reasonable assurance regarding the achievement of objectives. In this respect, COSO[8] states that in order to identify associated risks it should be established in advance the organization's objectives, which shall be grouped into four categories as follows:

- strategic objectives, that define the mission and long term development directions;
- operational objectives, that refers to the effective and efficient use of available resources;
- reporting objectives, that refers to reporting reality;
- objectives of compliance, that refers to comply with the regulations, standards, rules or regulations applicable to the organization.

In order to define the objectives, the key is that, first, to define strategic objectives, and then, of these, to derive other types of goals: operational, reporting and compliance.

Also, for each goal it is necessary to establish risk tolerance, accepted materiality concerning the degree of achievement of identified indicators attached to the objectives in order to be considered achieved.

b. *determining courses of action to manage risk.* To achieve risk management within the organization, the lines of action of the integrated risk management are:

- defining the organization's strategy on risk;
- setting activities to be achieved if the risk occurs;
- evaluating results and measuring performances;
- risk monitoring at corporate level;
- reviewing corporate strategy on risk.

The strategy on risk must be coherent, contain how to recover losses caused by an adverse event and to integrate risk response measures.

Activities to be carried out if the risk materializes deal with the settlement of measures to address the consequences of risk, recover losses and identifying and implementing appropriate control devices to eliminate the causes that led to the risk occurrence.

To apply vigorously decisions taken in order to ensure effective functioning of integrated risk management will ensure continued operations and obtaining the expected results.

Monitoring risk at corporate level refers to observing the functioning of integrated risk management system, identifying and reporting existant weaknesses to adopt necessary remedial measures.

Updating the strategy on risk is necessary to be made whenever the organization changes its development strategy or strategic objectives, and also when management's risk policy changes.

Also, periodic review of risks involves the redistribution and concentration of resources in areas of interest.

c. *determining relations between integrated risk management system and other subsystems of the organization.* The organization's management must permanently ensure the interdependence between the objectives of the organization, its functional departments and risk management.

Risk management process aims to identify and assess risks that can affect the objectives' achievement and to establish risk response measures. It should *"become part of the organization's functioning as the base of management approaches*[9]*".*

Considering that the objectives concern all levels of the organization, strategic, general and operational, being defined at strategy level, functional departments and even individual level, in a post, it is required that risk management to be aware of all the relationships that occur or develops between them or within them.

The incomplete determination of the relationship between risk management system and other subsystems of the organization, will lead to an inadequate identification and management of risks associated to the objectives with major negative consequences on the organization.

d. *setting activities, responsibilities on risk.* Seeks to identify all activities in progress within integrated risk management process and establish responsibilities for implementing each activity. Since the process involves all functions and functional departments of the organization, it is required that the activities and responsibilities on risks, defined and agreed at their level, to be communicated to employees involved in carrying out the activities.
e. *defining performance indicators.* For each strategic objective, operational, reporting or of compliance defined at corporate level, must establish performance indicators by which to ensure measurement of the degree of achieving goals. Also, setting goals to achieve within each indicator, will allow establishing performance resulting from the risk measures imposed within each goal.
f. *allocating resources necessary to carry out activities and training the staff involved.* For each activity planned to be conducted, it must be identified the necessary resources for their achievement, respectively financial, human, material and information resources. Resources necessary in order to accomplish the activities must be available and approved in budgets.
g. *communication and consultation on the results*, performance evaluation *related to risk compared to objectives planned.* Communication involves on time and clear transmission of necessary information about risk, as follows:

- the responsibles for risk management communicate information about the process content and also on management decisions relating to any measure on risk;
- the responsibles for risk of functional structures communicate information on risks associated to objectives established, and on how risks are managed.
- the entire staff reports information on identified risks and whose management needs to be achieved.

The consultation on the results aims to provide information on risk exposure, after their evaluation and the implementation of control measures. The role is to establish the effectiveness of control measures applied.

Performance evaluation of risk aims to determine performance obtained due to the risk response compared to the costs involved for implementing control measures taken to reduce risk and maintain its level within the risk appetite.

h. *monitoring effects and reviewing formulated strategy.* It involves evaluating the efficiency and effectiveness of risk management process within the organization and conducted according to the results obtained to carry out the appropriate review of the risk strategy, in order to ensure the minimization of adverse events and appropriate integration of measures to respond to risk.

In our opinion, we believe that the implementation and operation of an integrated risk management is neccesary, it can be done through ongoing monitoring of risk and integration risk response measures, based on risk strategies, which ensure the objectives achievement and deliver the expected results, in case of an event causing loss.

The firm implementation of decision taken, as the effect of the effective operation of integrated risk management system, gives premises for further activities and obtaining performance across the organization.

Knowing threats that affect the achievement of the goals will allow their classification according to the level of materialization, the extent of impact on the objectives and costs involved for the measures necessary in order to minimize risk effects. Establishing a hierarchy of threats will lead to establish an order of priorities in resource allocation.

4. Integrating risk management into the management sistem

The conception, implementation and operation of an integrated risk management system must ensure ongoing monitoring of risk and the integration of the risk response measures in a coherent risk strategy.

Risk strategy should contain clear objectives on risk policy promoted and applied within the organization, to define exposure levels and response to risk in all circumstances where it is analyzed and evaluated. Also it should be set the terms and conditions for recovery of losses whenever the risk is manifested and had or will have financial consequences.

4.1. Integrated risk management system - Part of the organization's management system

Implementing an integrated risk management within the organization will allow the organization's management to focus its resources on those risks that affect the objectives achievement, in order to protect assets, ensure continuity of organization's activities and adopting the effective decisions.

Risk management function must be a defining function within the organization and provide a complete and coherent set of activities and actions that define decision-making of the organization if the risk materializes and to guide staff in risk management.

An effectively integrated risk management system must ensure the recovery of the organization in case of interruption in activity, by maintaining its essential functions, at least of minimal levels from event appearance until its remediation.

The decisive part in the functioning of an integrated risk management system is the plannification in order to ensure business continuity, because it contains measures of recovery for activities under risk event.

The approach, implementation and functioning of an integrated risk management system in the organization is achieved depending on the processes undertaken, the organization situation and leadership style. However, to ensure process efficiency it needs to be taken into account primarily the following:

- COSO[10] principles on the integrated risk management, whose compliance involves designing and implementing an efficient risk management, which contributes to further objectives and efficient use of resources;
- risk approach within the integrated risk management, starting from strategic risks and then the operational, reporting and of compliance;
- analysis and risk assessment must be done in terms of relevant factors, materiality, impact, probability;
- preparing reports on risk management, having practical value that can be used by management in making decisions.

The role of integrated risk management system is to ensure the implementation of risk management function within the organization's management system. Its functions are activated while the organization's management system signals the existence of threat in achieving its objectives and deliver the expected results because of their activities.

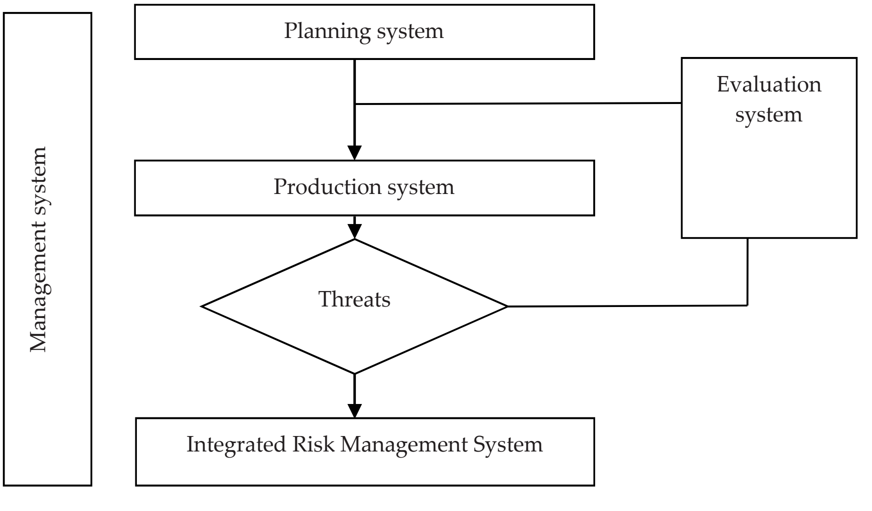


Figure 1. The management system of an organization

From the scheme presented above it can be seen that developing and implementing an integrated risk management enables entity's management to focus efforts on the risks affecting the achievement of the objectives.

Also, the integrated risk management system reflects the integration of all activities and actions related to risk and risk management in a single system so that it can act upon them at one level. By it, the parallelism and dysfunction of action and communication are eliminated, occuring within organized systems operating independently of each other.

Implementing an integrated risk management system within the organization leads to the following:

- strategic risk analysis, operational, reporting and of compliance that may affect the achievement of organizational objectives;
- definition and prioritization, according to risk level and costs required, of control devices required to eliminate the consequences if the risk has manifested or to limit the risk, if it constitutes a threat to the organization;
- identification and evaluation of internal controls related to activities and actions attached to the objectives, both in terms of their existence and of those expected to exist and the establishment of areas or zones that require implementing control measures, so that targets set to objectives to be achieved within conditions planned;
- designing and implementing certain internal control measures to improve activities;
- providing conditions for the organization to comply in different situations;
- establishing data and critical information concerning the environment of the organization, that may be used in the analysis and decision-making strategy.

Exercising risk management function, as defining function within an organization, involves making through integrated risk management system a coherent set of processes, activities and operations, by which it is ensured an effective risk management and defined the decision-making process if risk occurs.

However, depending on the types of risks identified, on the response to risk determined according to risk appetite, on the costs involved and the levels at which risks may be maintained after their treatment, integrated risk management system can guide organization to improve work according to the benefits of good risk management.

4.2. Assessing and measuring risks – Component of integrated risk management system

In the integrated risk management process, the component on risk assessment is a major step aiming to:

- identify significant risks within the organization, associated to objectives;
- assess the capacity of the internal control/management system to prevent/manage risk effectively;

- determinate significant risks and uncontrolled adequately by the organization and that are going to be treated to reduce exposure levels;

Risk assessment depends on the probability of occurrence and severity of the consequences if the risk materializes, meaning the impact of risk and uses as tools the risk assessment criteria. These criteria should cover the purpose, in which risk was identified, in terms of compliance and performance.

By prioritizing are selected medium and large risks on which will conclude responses to the risk.

The risk assessment process includes the assessment of inherent risks existing before the implementation of control measures and residual risks, resulted after implementing control measures and have two phases, namely:

a. *Assessing probability* is a qualitative element and is carried out by evaluating the potential for risk occurrence, by considering qualitative factors specific to the context in which goals are defined and achieved. This can be expressed on a scale of values on three levels as follows: *low probability, medium probability and high probability*. Illustration:

PROBABILITY	**If**
LOW	Rare modifications in the regulatory framework, over 3 years Less complexity of activities and actions Experienced staff Objectives and targets are not changed Reliable, adequate and updated information Processes well designed, formal and conducted
MEDIUM	Legal framework is relatively new or experienced significant changes Average complexity of activities and actions Average level of employment and experience of staff Rare changes of objectives and targets Existing information from many sources, but insufficient Processes related to practice
HIGH	Very frequent modifications in the regulatory framework High complexity of activities and actions Inexperienced staff and newly employed Frequent changes of objectives and targets Poorly designed processes and lead Insufficient and outdated information

b. *Assessing impact* is a quantitative element and is carried out by evaluating the effects of risk if it would materialize, by considering quantitative factors specific to the financial nature of the context of achieving objectives. This can be expressed on a scale of values on three levels as follows: *low impact, moderate impact and high impact*. Illustration:

IMPACT	If
LOW	Low cost of implementation of activities/actions, under planning No losses of financial assets, employees nor materials Good image of the organization Competencies and responsibilities provided in decision making Good quality services Continuity of activities is ensured
MODERATE	Costs of implementing the activities/actions equal to planning Reduced losses of financial assets, employees and materials Moderate image of the organization Decisions made without assuming responsibilities Moderate quality of services provided Very rare interruptions in activity
HIGH	High costs in relation to implementation planning of activities/actions Poor image of the organization Decision making without ensuring the competence and responsibilities Poor quality services provided Significant break in activity

Risk analysis criteria are represented by the probability assessment of risk occurrence and the impact level assessment if the risk would materialize, as follows:

- the probability assessment is made based on the analysis and evaluation of quality factors specific to the context in which objectives are defined and met;
- assessing the level of impact is made based on the analysis and evaluation of quantity factors specific to financial nature of the context of achieving objectives.

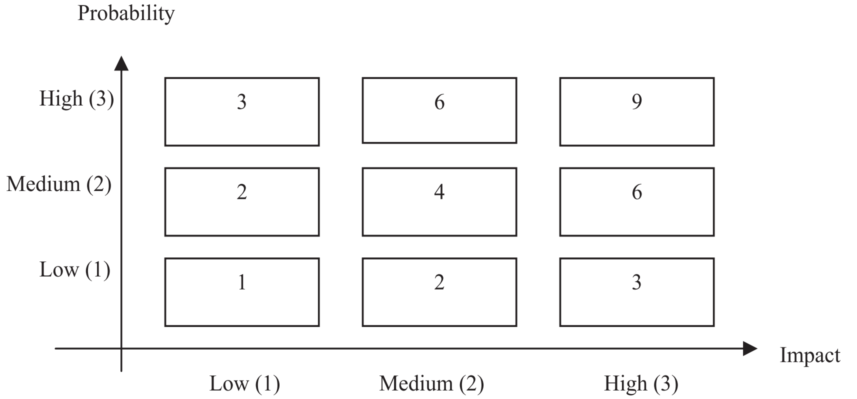


Figure 2. The level of risk depending on the probability and impact

Establishing the response to risk and pursuing if it falls into the risk appetite, agreed by the organization's management, is carried out by multiplying probability and risk impact, obtained from the formula:

$$PT = P \times I,$$

where: PT = total risk score
P = probability
I = impact

Depending on the outcome of the risk measurement process, applied to all risks the organization faces and that affects achieving objectives employment shall be: high risk, medium risk and low risk as follows:

- for PT = 1 or 2, low risk
- for PT = 3 or 4, medium risk
- for PT = 6 or 9, high risk

c. Assessing internal control

To assess the internal control are considered the risks associated with the objectives the organization faces and that were measured.

Internal control assessment process involves the identification and analysis of internal controls expected and existing, implemented by the entity to manage risks and aims to establish areas where it does not work or work improperly. This can be expressed on a scale of three levels as follows: compliant internal control, internal control partially compliant and non-compliant internal control. Illustration:

INTERNAL CONTROL	**If**
COMPLIANT	Implemented internal control system, prevent risk materializing. Regulatory framework of risk management and internal control known. Positive attitude towards internal control/management and risks. Internal control/management integrated into organization's activities and actions. Risk management ensures identification of significant risks, their assessment, establishing risk management measures and monitoring their effectiveness. Systematic reporting on activities development. Objectives met and appropriate remedies for violations.
PARTIALLY COMPLIANT	Internal control system is implemented, but does not prevent risk materializing. Neutral attitude towards internal control/management and risks. Internal control/management is partially integrated into the organization's activities and actions. Risk management process ensures the identification of risks, their assessment, but risk management measures are not always adequate and effective. Systematic reporting on activities development, but states objectives met.

NON-COMPLIANT	Internal control system not implemented. Regulatory framework of internal control/management is not known. Uncooperative or indifferent attitude towards internal control/management. Internal control/management perceived as a separate activity, conducted in parallel with the activities of the entity. Risk management does not provide identification of significant risks. Systematic reporting on activities development, but information is not reliable.

Risk response involves establishing and implementing possible actions, selecting those appropriate to the risk appetite and the costs required to implement risk management measures, by considering the following:

- for objectives whose risks were classified as low risk and for which internal controls have been assessed as compliant, the risk is residual, so the organization's exposure is below the accepted level. In these cases, the organization accepts the risk as such, without interfering for its treatment, but will provide ongoing permanent monitoring to ensure that the exposure level does not change.
- for objectives whose risks were classified as medium risk or high risk and for which internal controls have been assessed as partially compliant or non-compliant, the risk is inherent, so organization's exposure is above the accepted level. In these cases, the organization will proceed to treat, avoid or transfer risks.

5. The structure of integrated risk management

Achievement of the objectives of integrated risk management within an organization presupposes the meeting, in a logical sequence, of specific and required activities, as follows: setting the context, setting the objectives, risk identification, risk assessment, setting a risk response, implementation of control measures, information and communication and monitoring.

5.1. Integrated risk management process

Integrated risk management is structured on component elements of the COSO model, indicating that the control environment is defined by the internal environment and risk assessment consists of setting goals, identifying events, risk assessment and risk response.

5.1.1. The internal environment

It represents the theoretical and conceptual stage of risk management process, which presupposes an organizational culture on risks and knowledge of risk management operating concepts, and whether they are implemented and known at all levels within the organization.

This stage involves carrying out specific activities to implement risk management within the organization, as follows:

- *establishing an organizational context*, that analysis of objectives, operating structure, delineation of duties and responsibilities and the main conditions in which the organization operates. They also set requirements for future development of the organization and key risk exposures, including the characteristics and consequences;
- *setting the context of risk management*, the concept of the organization against the risks they face and the level of acceptance in relation to exposure to risk.

In relation to the means of establishing the context of implementation of risk management it is established and designed risk management policy, objectives and tasks of the implementation of risk management methods and methodologies for the identification, evaluation, treatment and control risk. At the same time, it is determined the structure responsible for risk management, the powers and responsibilities of it, taking into account the fact that *"management activity it means to commonly achieve the necessary objectives for the final of the organization*[11]*"*.

The characteristic of this work is the tone given by the organization on risk management and methodology they use in risk management and how are communicated the concepts of risk and the response of staff on risk management philosophy.

5.1.2. Objectives establishment

Implementing an integrated risk management system involves identifying and assessing the risks that are threatening to accomplishment of objectives.

This includes risks related to activities and actions of input and risks of actual processes undertaken within the organization, risks that prevent achieving the intended results and the risks about the impact of realized activities on organizational development.

Identification of the events that may affect achieving the expected results is only possible if objectives are set in advance and under each one were defined activities necessary to ensure their implementation which, therefore ensures, the delivery of the expected results.

If we consider the approach according to which performance is characterized as "achieving organizational objectives regardless of their nature and variety" [12], we believe that goals should be established to represent a challenge for management and employees.

Management by objectives has a beneficial effect for the organization, it facilitates the exercise of effective control over all activities, motivates employees to participate in the objectives and it creates a coherent organizational framework which stimulates the collaboration between all structures within the institution.

The control of meeting the objectives is considered necessary for the management of the organization and requires each manager to have established controls for each activity and objective for which he has responsibility. At the same time, it must be taken into account the impact of likely risks that may jeopardize the attainment of these objectives, so it is necessary to design and implement appropriate risk management systems.

5.1.3. Identification of events

To ensure achievement of activities as planned, it is necessary for the management to identify all events, internal and external, positively or negatively affect the objectives, and depending on the probability of event and type of consequences that can be produced in the organization they are divided into risks and opportunities.

Risk identification, depending on the time in which the process takes place, involves the following stages:

- the initial risk identification specific to newly established organizations and those who have not previously identified risks.
- the permanent identification of risk is specific for those organizations that have implemented a risk management and necessary for assessment of risks that have not previously shown to change their circumstances, and the limitation of the probability to manifest[13].

An effective risk management involves identifying risks at any level, where there is a threat on the goals and taking specific measures to limit the problems caused by these risks. Risks can be identified and defined only in relation to those objectives that are affected by their materialization.

Risk identification can be achieved in two ways:

- self-evaluation of risks carried by each employee involved in the objectives and activities, regardless of where they performed tasks hierarchically, by monitoring the risks they face daily,

or

- establishing a special department within the entity, that has responsibilities regarding the evaluation of operations and activities within the organization and on this basis the identification of risks that characterize the organization's objectives and individual goals set for thr employees.

Application of either of two ways to identify risks can have negative consequences for the entity because, first, each employee has a certain culture and training which leads to a different understanding of risk management, making monitoring, to identify risk differ from employee to employee. Also, some employees can be more involved in current tasks and pay less attention to their risk management.

Second, establishing a specialized department, with responsibilities in risk identification ensures not always effective risk management. However, as much the staff of this department is prepared, it is very difficult to know in detail how to achieve the activities and therefore to identify all threats that may affect achievement of objectives.

The practical and effective risk identification is the combination of the two forms presented. Thus, employees from all levels of the organization have responsibility for identifying and reporting threats to their achievement by the specialized compartment, and it has the responsibility to assess each reported event and if it finds that the event reported is a risk to do registration, evaluation and its treatment.

In identifying and defining risks should be considered the following rules[14]:

- *risk must be an uncertainty*, so it must be considered whether it is a possibility or about an existing situation which is an existing problem and not a risk;
- *difficult issues identified should be assessed*, as they can become repeat risk situations;
- *problems not occurring are not risks*, this means that the organization has control over them, and their analysis may lead to consumption of resources;
- *problems that are guaranteed to arise are certainties* and measures are to be taken as such, with certainty as a starting point;
- *risk should not be defined by its impact on the objectives*, as the impact is the result of the risk materialization;
- *risks are identified by correlation with the objectives*; the aim is to identify those threats that could lead to failure of objectives;
- *risks have a cause and effect*, the effect is the consequence of the materialization, and the cause is a circumstance which favors the appearance of risk;
- *making the distinction between inherent and residual risk*. Inherent risk is related to the objectives and the risk is there before intervening with internal control measures. The residual risk is the risk result after implementation of internal controls. Residual risk that results from the inherent risks cannot be controlled completely, whatever measures were taken, uncertainty remains.

On identifying opportunities, they are performed by employees within the organization regardless of where they are, and their recovery is the responsibility of management, to be used to increase efficiency and effectiveness of activities.

5.1.4. Risk assessment

Achieving this step involves assessing the likelihood of risks materializing and the impact of risk when it would occur, and classification of risk on 3 levels (high, medium or low) based on a risk analysis matrix.

After the risk assessment process is done, priorities are established so that high risks are considered by management to treatment.

The purpose of risk assessment is to establish a hierarchy of risks within the organization and to establish the most appropriate ways of dealing with risk.

Risk assessment process involves consideration of the following:

- *the probability of materialization of the risk* stems from the fact that, at some point in the progress of activities, there may be conditions that favor the emergence of risk. In these conditions, analysis of the causes which favored the emergence of risk can lead to an appreciation of its opportunities to materialize;

- *the impact of risk on the objectives* represents the consequence of risk materialization, and how risk affected the achievement of the objective;
- *risk exposure* represents the extent to which risk can be accepted by the organization, if it would materialize;
- *determination of the specific outcome*, involves risk assessment after deployment of control. The result may be a risk exposure exceeding the limits of acceptance, which means that risk is inherent, which involves the review of existing internal control mechanisms, or exposure below the limits of acceptance, which means that the risk is residual.

The risk assessment is performed to identify the likelihood and impact of risk and thus to determine how it can be managed.

Risk assessment must be the essential component and a constant concern of management organization, as the people change, regulations change, the objectives are reviewed or new ones established. All these contribute to the continuous changing of the map risks, namely the emergence of new risks, modification of existing risks and the level that the organization accepted the risks.

5.1.5. Reaction to risk

Information collected following the risk assessment is processed and measures to diminish risk exposure identified. To limit exposure the organization should identify opportunities to reduce risk, the probability of the event, or if this it is not possible, to establish measures to eliminate risk.

Also, the organization should develop appropriate criteria for risk management to reduce the likelihood of risk and risk consequences. If risks are not well managed or costs are high relative to benefits of the activities, the criteria should be directed to transfer the risk or eliminate the risk.

The management of the organization, based on the risk assessment, will determine the response to risk, as follows:

- *accept the risks as they are*, without mitigation measures, and without devices to establish and implement internal control. Acceptance or tolerance of risk as the risk response strategy is recommended for the risks inherent with low exposure, less than the risk tolerance.

After acceptance, the risk becomes residual and will be monitored regularly, aiming as it does not change the level of acceptance.

Setting the limit for the tolerance of risk is the responsibility of management and involves the establishment of the exposure that can be assumed, in conjunction with costs and control measures to be taken. If the risk exposure is a probabilistic measure on a sized scale (combination of probability and impact) then the risk tolerance must respect the same features.

- *treat risks,* and that will identify and implement appropriate control devices, to limit the probability of risk manifestation and keep it within acceptable limits.

In practice, for risk treatment the following categories of controls instruments are used:

- *preventive control tools* are used when it is intended for an undesirable outcome not to materialize or for limiting the risks that may materialize;
- *corrective control tools* are used when it seeks to correct undesirable results of risk materialized and is a way to recoup losses;
- *direct control tools* are used when seeking to ensure any particular outcome, that is when we intend to effect a particular risk that may materialize to be oriented towards a tolerable direction of the organization;
- *detective control tools* are used when aimed at identifying new situations arising from the risk materialized.

If the risks materialize, the cause is represented by the internal control that either has not been implemented or was implemented but they not functioned properly.

- *avoid risks*, risks that cannot be treated, and treatment costs are higher than assumed results, will be eliminated or kept within reasonable limits by reducing or abolishing their activities.
- *transfer risks*, risks that cannot be controlled will be transferred to other units or organizations. This option is especially beneficial for financial or economic risks. Transfer risk is a measure to help reduce exposure to a functional structure of the organization, but another functional structure or organization, which are capable or specialized in managing such risks, will take the risk exposure.

Diversity of internal control is considerable for all aspects of activities and can be classified as: objectives, resources, information systems, organization, procedures and supervision[15].

Objectives - grouping tools/internal control devices implemented through measures aimed at: their clear defining, their decomposition into a pyramid up to the job, convergence, measurability, association of measurable outcome indicators and monitoring information system.

Means - is the group of devices/tools of internal control implemented through measures of adequacy of resources against objectives.

Information system – it groups devices/an internal control instrument operationalized and aims to achieve a complete information system and steering, reliable, comprehensive and appropriate.

Organization - grouping devices/internal controls instruments resulting from application of measures aimed at correcting anomalies detected in the procedural and structural organization and that are circumstances favored for the manifestation of risk.

Procedures - are tools / internal control mechanisms which control the risks arising from lack of processes and rules to be observed while activities are taking place.

Supervision - grouping instruments/devices of internal controls designed to control risks arising from abnormal exercise hierarchical control. Such internal control tools are aimed at the management style of the makers of different levels.

5.1.6. Risk control

Represents policies, procedures, controls and other management practices established by the organization to make a prudent management of risks, and ensure the implementation of activities as intended. Also, to control risks is to ensure that objectives are met and significant risks are properly managed.

To prevent conflicts it is recommended to ensure independence of risk control to functional structures of the organization that runs the identified risk. Any measure taken to control risks should be placed in the famous "internal control system", which is responsible for directing the implementation.

Risk control requires that the functional structure where there is a risk, carry out continuous monitoring of risks and appropriate mitigation of the manifestation probability or risk impact. Otherwise, the risks are uncontrollable and there are no means of intervention to limit the probability and risk impact.

5.1.7. Information and communication in the supervision of risk

Activities are initiated by the management entity for transmission to employees of their responsibilities regarding the identification and monitoring of risks.

At the same time for employees to ensure proper risk monitoring in accordance with the requirements of established risk management process within the organization, it is necessary for the management to provide appropriate and timely information for them to accomplish the tasks set.

5.1.8. Risk monitoring and supervision

Risk monitoring involves reviewing and monitoring whether their risk profile changes following the implementation of internal controls. Review processes are implemented to assess whether: risks persist, new risks have emerged, the impact and likelihood of risks have changed, internal controls are effectively put into practice or risks should be redefined.

Risk monitoring involves tracking the knowledge of strategies applied to risk management, of their implementation and the evaluation of performance after implementation. Risk-sensitive areas are monitored continuously, and the results are sent in the initial stage for reconsideration, identification and implementation of adequate internal control tools or application of other ways to reduce exposure to risk.

The management of risk register, which contains summary information and decisions in risk analysis, attests that the organization has introduced a risk management system and that it works.

The process of identification, assessment and risk treatment must ensure that risk analysis is carried out periodically and are established mechanisms for information management on new or emerging risks of changes in already identified risks so that these changes to be addressed properly.

Risk monitoring is necessary to monitor progress of risk profiles and to ensure that risk management is appropriate and is obtained by revision of the risks.

Risk monitoring is done through internal control, which must be flexible, and develop appropriate control tools in areas where the risk is not sufficiently controlled or reduce those instruments where excessive risks are controlled.

Risk management must consider internal control system implemented in the organization, and the expected internal controls and internal controls existing, and considering their sufficiency identifies the risks, makes them subject to the evaluation and based on results establishes the internal control necessary to be implemented in order to limit exposure.

5.2. Internal and external environment and its influence over the integrated risk management

The implementation of a risk management system within the organization should impose establishing relationships both within the organization and beyond. Also, the ones responsible for implementing integrated risk management have relationships with the entity's management and staff of the entity's functional structures.

The management of the entity shall decide on the risk management strategy adopted in the organization and approve any measure relating to the risks. In this regard, is regularly informed of the results of risk management and carry out in order to establish ways in which the risk management is done.

The ones responsible for risk management in the organization are communicating and realizing the risk strategy and policy promoted to all the employees, and any decision taken by management on risks. Receive from the structures any information on the risks, analyst, process, and make proposals for the management on appropriate measures to be taken depending on the nature of managerial implement these measures.

Risk communication and how they are required to be managed is based starting on the management level to the level of execution and shall ensure that:

- risks related strategy and all associated risks to the objectives are known by all the staff involved in achieving the objectives;
- staff in the organization is aware of the risks they assume and their monitoring system.

The nature of relationships established in risk management process is a functional one, respectively; the ones responsible with the risks have the authority to charge risk of transmitting to the functional structures of the entity information on risk strategy and information related to risk management process implemented. At the same time, they require the information about the identification and management of risks.

The increase of confidence in the risk management system promoted and implemented at the organizational level is achieved by:

- developing a clear risk management strategy;
- sufficient support at all levels to ensure risk management;
- development of simple risk management systems;
- communication with all parties involved in risk management;
- communication and equal relations and cooperation between different functional structures of the organization;
- improving the risk assessment.

Entity's activities are influenced by several external factors, the nature of threats that affect achievement. Integrated risk management system must identify the nature of the risks these threats, to analyze, evaluate and determine the response to risk. In some cases, establishing a risk response does not ensure acceptable risk appetite as risk reduction measures are dependent on the activities and objectives of the organization.

To ensure acceptable levels of risk there should be a system of relationships established with various external factors, which, put in place, to ensure reduction of exposure.

Building and implementing an integrated risk management system helps to direct resources to risk which particularly affect the activities and support the organization in achieving its objectives.

6. Impact of integrated risk management on the organization

To ensure good risk management it is important to provide assurance that each employee understands properly the risk management process within the organization and knows his role and responsibilities in this process.

Risk management process does not require identification and elimination of negative events that may affect the carrying out, if the risk occurs, but also aims to analyze and evaluate risk and risk appetite according to design and implement control devices to limit the probability of risk. It provides the management with a *"framework approach to effective risk management and its possibilities".*

Risk management objective is to identify risks, causes that generated them and establish appropriate control device to reduce its level, but using the lowest cost.

By implementing an integrated risk management system shall ensure:

- strategy development, objective setting and risk management mechanisms considering the risk appetite. The organization will define its development strategy to the risks they face and how to manage, taking into account the limit of the appetite to which it may be exposed. The objectives are dependent on the planned development requirements and performance levels established, but should be considered the risks to the objectives and the costs necessary to manage these risks.

- development of a framework for the level of response to risk. This involves performing analysis and diagnostics, in order to determine the level of risk to which the organization can be exposed and considering the results obtained, to proceed with the acceptance, treatment, and avoiding or risk transfer.
- improving the expertise to identify events that threaten the organization and establishing decisions with efficiency and effectiveness. Applying an integrated risk management process will allow evaluation of the risks, by providing a link between the objectives, functional departments of the organization and components of risk assessment. Making this process will help increase the expertise in knowing events facing the organization, the nature of the risks threatening the objectives and nature of opportunities.
- identifying and managing risks that affect the achievement of objectives and the set planned results and not risks of every operation or activity achieved. Integrated risk management system, is not fragmented, to ensure identification and risk assessment, in isolation, only at the operation or action, but is a system for identifying and addressing risks to the target integrated. This ensures that by implementing a single control measures to be managed more risks. It also allows knowledge of risks affecting achievement, which ensures that decisions are based and to consider the risk exposures.
- identifying opportunities following monitoring events and their capitalization with benefits in increasing efficiency and effectiveness of the activities. Integrated risk management system takes into account the analysis and evaluation, events that may affect achievement of objectives. These can be negative events that are risks and positive events that are opportunities.
- appropriate use of capital. Knowledge of risks the organization is facing in achieving objectives, allowing management to guide decisions to those activities where the risks are well managed, thus ensuring better use of available resources.

Integrated risk management model has some limitations due to errors, avoiding checks, and human judgment in making decisions that can sometimes be wrong. These limitations make it impossible to issue an insurance of the need to achieve objectives.

At the same time, responsibility for designing and implementing appropriate risk management is the organization's management and other staff to support risk management philosophy and apply established rules on risk management, each in their area of responsibility.

The classic risk management process, which was joined and implemented by most of organizations, is a fragmented on, in the sense that functional structures within an organization manages its own risks independently. Thus, each compartment, according to the procedures and methodologies developed, shall identify and manage risks associated to objectives independently, without a coordinated approach and without taking into account the interdependencies of risk within the entity.

7. The advantages of implementing integrated risk management

Integrated risk management has mechanisms to help the recovery of the organization in the situation of work stoppage, major incident or disaster, by maintaining minimum levels of business critical functions.

The main feature of integrated risk management system is that it integrates risk monitoring mechanisms of the functional departments of the organization and its culture, with a focus on the risks associated with strategic objectives. Also, the emphasis is on monitoring and controlling risk, and minimizing it.

Advantages of an integrated risk management can result as follows:

- insurance of correlation with the risk appetite of the organization strategy. Risk appetite is a limit to which risk can be accepted and to which the organization may be exposed. The management has the opportunities available to them to achieve goals and select the most advantageous option in conjunction with the profile of risk;
- helping to improve decisions about risk treatment. The management options to limit risk, assess their correlation with risk appetite and costs and determine appropriate measures for risk management.
- integrated approach of risk allows that, by establishing a single internal control measure to be able to handle more risk, or one risk, but which is found in several functional structures of the organization;
- the capitalization of opportunities. Integrated risk management takes into consideration events outside the negative and positive nature of the risks, the nature of opportunities.
- improvement of management decision. Knowledge of risks the organization is facing and the level of risk exposure contribute to a more realistic analysis and substantiation of managerial decisions. Substantiate decisions can be made by considering the following requirements: the existence of one or more objectives to be achieved; existence of several alternatives, including economic factors in decision-making plan, making the decision, decision and action unit, clear and optimal fit between

Between components of integrated risk management and objectives of the organization must be a direct relationship. The analysis and risk assessment by following the eight components of integrated risk management, namely internal environment, identification, analysis and risk assessment, risk treatment, risk control, information and communication and monitoring of risks is done for each structure and functional organization for each objective.

By applying this method it is showed that risks are assessed and treated for all purposes of the organization, regardless of their definition (strategic, operational, reporting and compliance) and regardless of the compartment or structure that are defined.

Meanwhile, the integrated risk management process represents an instrument that allows a coordinated approach across the organization to identify and analyze the mechanisms of risk whose initial starting point is the strategic dimension.

Integrated risk management is a powerful tool that enables the management of the organization to have a picture of the risks affecting the achievement of strategic and operational objectives, and provides at the same time, leverage for the foundation and management decision making.

The process of identification, analysis and assessment takes into account the events of the organization, which can take negative shape and are associated with risks or positive shape and are associated with opportunities.

Author details

Emilia Vasile and Ion Croitoru
Athenaeum University Bucharest, Romania

8. References

[1] K. H. Spencer Pickett -The Internal Auditing Handbook, Second Edition, John Wiley & Sons Pub. House, III, River Street, Hoboken, NY 07030, USA, 2006.
[2] Vasile Emilia, Marcel Ghiţă, Marin Popescu, Control şi audit intern [Control and internal audit], Bren Pub. House, 2004, Bucharest
[3] Risk management. Sydney: Standards Australia International, 2004;
[4] Marcel Ghiţă, Ion Croitoru, Dan Ţogoe, Guvernanţa corporativă şi audit intern [Corporate governance and internal audit], Europlus Pub. House, Galaţi, 2010
[5] Michel Crouhy, Dan Galai, Robert Mark, Risk Management, McGrow-Hill, 2001
[6] Marcel Ghiţă, Ion Croitoru, Marin Popescu, Dan Vilaia, Maria Popescu, Guvernanţa corporativă şi auditul intern [Corporate governance and internal audit], Tehnica-Plus Pub. House, Chişinău, 2009
[7] Marius Tomoială, Emil Mare, Controlul intern-managerial [Managerial-internal control], Ştef Pub. House, Iaşi, 2012
[8] Integrated frame of COSO on Internal Control
[9] Treasury H. M., Cartea Portocalie – Gestionarea riscurilor, principii şi concepte [Orange Book - Risk management, principles and concepts], 2004
[10] Integrated risk management – Integrated framework, COSO II, 2004
[11] Vasile Emilia, Human Resource Management, Universitară Pub. House, Bucharest, 2004
[12] G. Lavalette, M. Niculescu - Stratégies de croissance [Growth strategies], d'Organisation Pub. House, Paris, 1999
[13] Popescu M., Croitoru I., Mangementul riscului, Internaţional Conferince an Risk Management Assessment and Mitigation, 2010

[14] Ministry of Public Finance, Methodology for implementing the internal control standard "Risk Management"

[15] Croitoru Ion, Marcel Ghiţă, Managementul Riscului [Risk Management], Magazine, Rep. Moldova, 2011

Chapter 13

Integrated and Personalised Risk Management in the Sensing Enterprise

Óscar Lázaro, Agustín Moyano, Mikel Uriarte, Alicia González, Teresa Meneu, Juan Carlos Fernández-Llatas, Vicente Traver, Benjamín Molina, Carlos Palau, Óscar López, Etxahun Sánchez, Saioa Ros, Antonio Moreno, María González, José Antonio Palazón, Miguel Sepulcre, Javier Gozálvez, Luis Collantes and Gonzalo Prieto

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50275

1. Introduction

Due to its impact on economy, resources, environment and society, manufacturing is of strategic value to Europe. European manufacturing has to embrace a new logic of global socio-economic sustainability, in which it addresses not only the welfare of its population, but also of emerging economies, contributing at the same time to the preservation of the environment and the resources. Megatrends that have a considerable impact on European manufacturing are:

- Ageing,
- Individualism,
- Advanced and emerging technologies / knowledge,
- Globalization,
- Urbanization,
- Sustainability,
- Finance and Public debt.

Under the influence of these megatrends, manufacturing sectors are undergoing structural changes in view of increasing their competitiveness through intelligent and sustainable solutions. The move from eco efficiency to resource efficiency is related to the need for building "citizen centred systems". This will require further improving the socio economic dimension of future metropolitan areas and factories by addressing the quality of life of the citizens living and working there.

This new perception of the worker in the manufacturing environments requires that a new approach is made available to manage risks and hazards in the manufacturing environment.

Worker's safety and health is observed jointly by legislation, knowledge generation and technology development to enable a full risk management focused on the employee (paradigm "factory worker first"). When the personalized risk management will be achieved, the European Strategy for Safety and Health at Work 2007 – 2012 will be able to reach the objective of 25% reduction in workplace accidents and then a yearly reduction of 5% to finally achieve the ambitious objective of zero-accidents.

So far, most efforts in work safety have been focused on improving work equipment features and definition of more secure tasks. Machine manufacturers have worked hard to provide security devices to eliminate or mitigate the risk, but success lies in considering security by design. Big gaps are detected in the process of establishment of security systems in industrial environments focused on the worker. The worker needs to be introduced as an active element in the risk management equation and proactive measures need to be facilitated to increase the effectiveness of the solutions in place.

All working environment variables and conditions in risk management require the challenge of finding technologies to monitor and manage the human factor in manufacturing processes. The reason is that the human factor is the main responsible for incidents and accidents in factories nowadays. The expected risk management system must incorporate proactive capabilities understood as the ability to detect the confluence of several risk factors with potential likelihood to cause an accident.

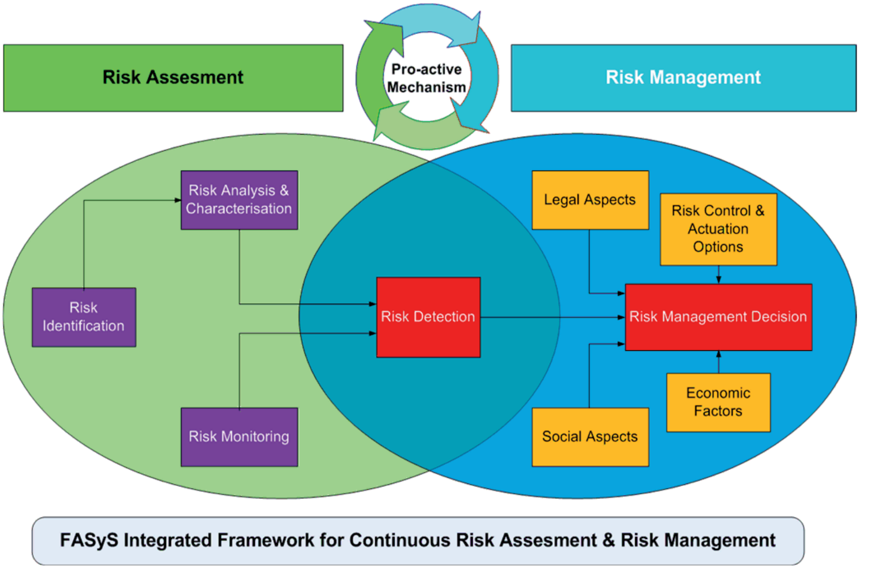


Figure 1. FASyS proactive risk management reference framework.

The best starting point is the general framework of proactive risk management provided by the ISO 31000:2009: Risk management – Principles and guidelines, see Figure 1. This standard has been suggested by the European Technology Platform on Industrial Safety

(ETPIS) as the framework for managing future manufacturing environments. To meet these challenges, technical, organizational and human resources are considered, in order to identify, detect, monitor and manage, on a continuous and effective manner, risks related to health and safety throughout the complete life cycle of the factory. For such solution to become fully effective, risk management system should be developed holistically taking into account an integrated view from sensing devices to reasoning mechanisms and intelligence, capable of reacting to extremely dynamic conditions.

The chapter is organized as follows. First a brief overview on the concept of Sensing Enterprise and Future Internet technologies, where the FASyS system will operate is provided. Then, Section 3 presents the FASyS model for proactive health and safety risk management. Subsequently, a more in depth discussion on key technological foundations of the FASyS model is presented in Section 4 to Section 6. Finally, the main conclusions and observations are summarized in Section 7.

2. The sensing enterprise concept

The Sensing Enterprise is a concept created by the FInES community in the context of the advent of the Augmented Internet. It refers to an enterprise anticipating future decisions by using multi-dimensional information captured through physical and virtual objects and providing added value information to enhance its global context awareness [1]. The enterprise will no longer be composed of and defined solely by atoms, but also by bits and kilobits.

The Sensing Enterprise concept is shifting boundaries – towards a borderless enterprise, where collaboration and continuous interactions among smart objects are central to the new scenario. Beyond the push and pull model, the sensing enterprise concept goes further to a direct presence, « sensing » data and transforming it into knowledge for business operation. The concept of sensing enterprise shifts the focus on the interaction among objects and systems.

The Sensing enterprise concept supports the notion of smart dust in the clouds as a new form and evolution of current state of the art computing systems. Thus, decentralised and delocalised computing and data storage resources provide dynamically scalable capacities to exploit linked open data that facilitate the exploitation of internal and external data systems. This highly flexible computing and sensing environment is the basis for a new generation of cross-cutting horizontal enterprise application areas. The Sensing enterprise concept leverages the power of sensor networks and decentralised intelligence to perform analysis and decision making both in synchronised real and virtual worlds.

3. The absolutely health and safety factory (fasys) model

The development of any excellence model should be based on a thorough analysis of the risks that can be faced in a particular working environment. However, the development of a proactive model demands that the very same approach can be used to completely manage in an integrated, proactive and continuous manner well-known as well as emerging hazards.

Figure 2. FASyS 13 prevalent hazards

This is the reason why the FASyS model has consolidated the vast diversity of incidents and accidents that can potentially take place in handling, machining and assembly factories into 13 prevalent hazards; such as trapping, falls on a level, awkward postures or repetitive and forceful movements, as illustrated byFig. 2. These 13 hazards have been used as reference in the development of the FASyS excellence model.

In the current regulatory framework, both legal and technical and having health damage prevention as the ultimate goal, the employer must "ensure" the maintenance and improvement of health, supported prevention services, which, through performances in R & D have to improve and evolve the performance and services provided. FASyS provides an integrated model for Continuous Risk Assessment, Monitoring and Management, that has to exhibit the following unique features:

- Integrated medical and technical risk management disciplines.
- Act as a single health model (mixed and integral)
- It is based and actions scientific knowledge for active risk prevention – technical
- Provides a "uniform" and universal framework for data and information management
- It can be "embedded" within the company's control and management system

3.1. FASyS excellence model

FASyS is the first integrated solution providing a coherent view to the 4 main dimensions that drive risk management i.e. methodology, technology, functionality and normative.

- **Methodological dimension** suggests that the risk model should be taken into account in the factory of the future; this model must establish the worker as the central point of health and safety management, thereby providing the missing link between occupational health, hygiene, ergonomics and psychosocial risks that current practices exhibit. This dimension of the model advocates for 4 different approaches to risk management so that proactive measures can be supported. FASyS methodological dimension combines safety performance based, risk based, incident and resilience based combined approaches as the means to address effective risk management. Moreover, this dimension increases traditional risk modelling functions on the assumption that risk should and will be monitored and therefore a suitable description should be made available.

- **Technological dimension** establishes the technological fabric needed to support the functional requirements of the risk management model. The technological dimension provides the technology blocks leveraging the concept of sensing enterprise. The FASyS approach to technology lies increasing interest and a prevailing role of ICT in the context of factory environment. In parallel with increased sensing and actuating capabilities, the improvement in backhaul communications present a new factory scenario where more autonomous intelligent reasoning mechanisms could be envisaged. The Internet of Things (IoT) scenario that needs to be handled is characterized by highly variable spatial and temporal contexts that should be effectively managed. FASyS combines the concept of autonomous systems with sensing and actuating capabilities with semantic-based distributed reasoning approaches to complex system operation. The technological dimension defines the reference architecture, where the risk management system will be integrated.

Figure 3. Three-level FASyS risk management reference architecture for sensing enterprises.

- **Functional dimension** includes all accesses to functional requirements needed to assure that the risk management cycle is complete and effective. The FASyS model defines 10 different modules that encapsulate the required functionalities to leverage ISO 3100:2009. The modules mainly deal with risk modelling and risk management strategy configuration, business impact analysis and system configuration, health and safety monitoring, autonomous actuation, decision support, personalised information and augmented training functions. Such scheme deals with a holistic and adaptive, evolving view on risk management.
- **Normative dimension** is based on the ISO 31000:2009 and establishes the five stages in risk management life cycle, which should include the previously described features to meet the requirements of each stage: Context, Organization, Monitoring, Intervention and Communication.

3.2. FASyS technological dimension

The implementation and demonstration of the FASyS model demands that coordinated progress is made in particular technology fields. FASyS has identified and prototyped technology at communication, complex event processing, human behaviour analysis, activity monitoring and reasoning level.

As depicted by the Figure below, FASyS technological fabric relies in 5 major components :

- **IoT Networking.** (a) Wireless sensor networks and activity monitoring (b) Communication security and privacy systems (c) Wireless communications in industrial environments.
- **Mixed Virtual and Physical World Detection and Evaluation:** (a) Industrial safety ontologies and reasoning engines. (b) Smart ergonomic characterisation solutions (c) Functional workplace adaptation models (d) Human error identification systems (e) Models for detecting psycho-social indicators and profiles (f) Adapted learning solutions (g) Chemical sensors for pollutant detection
- **Personal Health Systems:** (a) Applications for intelligent video analysis (b) Real-time risk detection tools (c) Automatic medical alert notification system (d) Occupational pathology assessment/diagnosis protocols
- **Machine Tool Active Security Systems:** (a) Part manipulation and part feeding systems (b) Volumetric protection systems (c) Auto-calibration and auto-compensation systems for large units (d) Intelligent part movement guiding systems (e) Visualisation systems for part/tool referencing
- **Comprehensive Real-Time Risk Management Systems: (a)** Personalized risk prevention strategies (b) Personalized decision support systems (c) Semantic solutions for services coordination (d) Emotional interfaces for effective risk communication.

All 5 technologic enablers are supported by a specific information distribution (acquisition and delivery) platform depicted by the Figure below. The FASyS services, provided by the functional dimension of the model, are leveraged in the FASyS plane based on the smart object system plane information and external information sources linked through the FASyS information gateway.

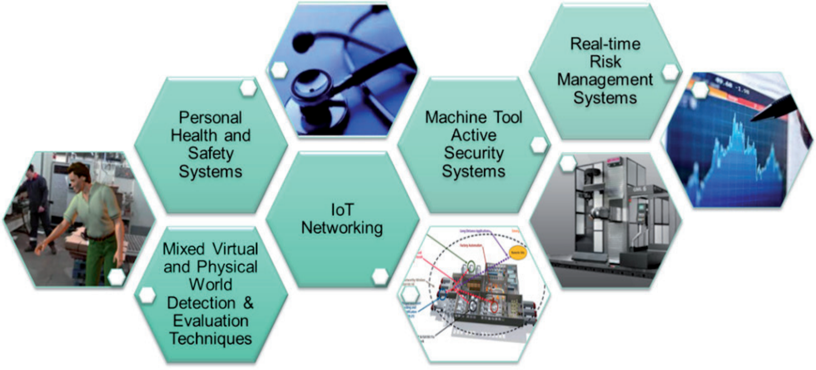


Figure 4. FASyS technological components

3.3. FASyS functional dimension

Ten modules are included in the functional dimension that are directly related to the health and safety services leveraged by FASyS:

1. **Risk assessment**. This module provides to safety managers many tools to access and manage the identified risks, their factors, values and relations. The management of these risks demands worker information, machine and device data and environment values related to each risk that could be monitored. In addition, this module interoperates with factory data stores.
2. **Preventive measures design**. In this module, the prevention responsible is able to design and establish the prevention measures catalogue to be used for each personalized risk identified. The preventive measure design use technical measures, medical protocols, data collection, affected users, execution managers and assessment agents.
3. **Economic impact evaluation**. In this module infrastructure and equipment implementation costs related to preventive measures are quantified. In addition, it is able to estimate the costs of non-prevention, in order to provide quantitative information about the integral risk management.
4. **Preventive measures configuration and management.** This module is used by safety and health managers. They use the complete prevention plan designed in the second module and associates devices (smart objects) to each action. In addition, this module monitors the correct operation of the devices and it alerts from any malfunction.
5. **Environment description module**. This module is used by safety managers and it monitors the real-time factory situation and its related actors, using visual tools to adapt the risk visualization through many filters.
6. **Personal health module**. This module is able to show to the health responsible a real-time monitoring of conducts, indicators and benchmarks to determine the evolution of the worker's health status and to provide early alarms related to health.

7. **Intelligent and automatic remote operation module**. This module provides tools to automate critical functions in manufacturing equipments. So, the equipment is able to adapt its operating parameters to the factory conditions, task values and worker status.
8. **Decision making automatic assistant module**. In this module, the safety responsible interacts with the decision support model, performing informed decision in multiple choice situations or in contradictory situations. In addition, it could be preventive strategies that need some personal interaction to validate a non-automatic event or decision.
9. **Emotional communication module.** This module empowers the prevention responsible to coordinate communication strategies on health and safety messages and to deliver them effectively, individually or massively. The messages could be sent on automatic or supervised manner. The module will also modulate the content of messages adapted to the emotional and psychosocial profile of the receiver.
10. **Continuous training module**. This module allows workers to be trained on prevention through "training pills", that are sent at the right time through the best channel. The safety responsible designs the protocols to assign the most suitable pills to a particular prevention strategy. The pills are training actions or mass reminders to one or many employees and with a predifined periodicity that respects emotional and congnitive constrains.

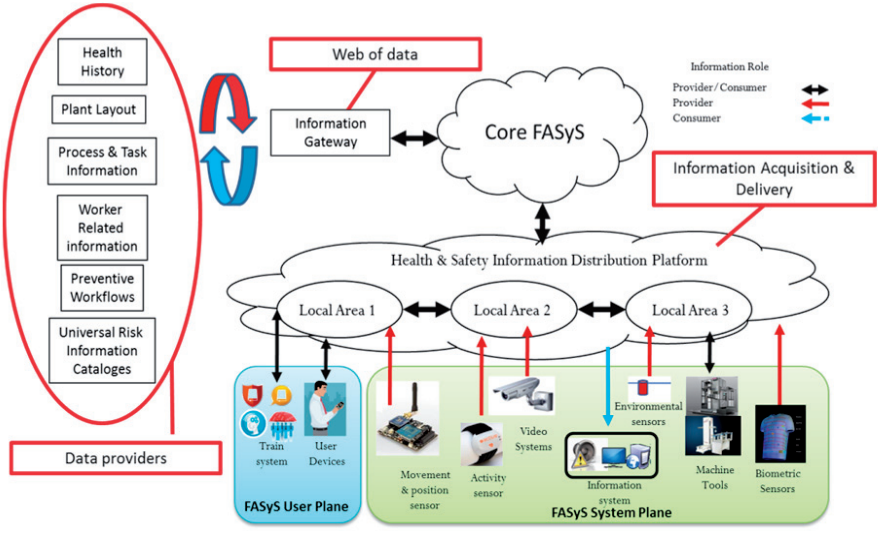


Figure 5. FASyS information data distribution architecture for the sensing enterprise.

4. Industrial wireless communications and Internet of Things (IoT) networking

The concept of sensing enterprise capabilities provided by FASYS are built at the smart object level leveraging on one hand enhanced industrial wireless sensing & communications and on the other hand, facilitating a semantic plane representation of the sensor information provided by smart objects. From a health & safety perspective, the major challenges faced by the wireless communication and IoT networking modules in the implementation of the sensing enterprise concept relate to (a) provision of a common architecture for secure heterogeneous communications (b) reliable wireless sensor network connectivity (c) universal object & sensor semantic representation. This section is devoted to describe how FASyS has addressed each of those fundamental challenges.

FASyS activities with respect to sensor and communication technologies are aimed at guaranteeing the desired levels of safety and health at work with the use of non-intrusive sensor systems (both personalized bio-medical and image sensors) to monitor the worker's physical conditions and the working environment, including environmental conditions and the state of the machinery interacting with the worker. The introduction of wireless communications in the factory of the future will also facilitate the deployment of distributed and mobile sensing applications to improve the factory's productivity and worker's health and safety.

4.1. Industrial wireless sensing and communications architecture

The deployment of heterogeneous wireless communications in industrial environments presents significant challenges [2]. On one hand, industrial environments are usually characterized by challenging propagation conditions (obstructions, interferences, etc.) that difficult the establishment of robust wireless links [3]. On the other hand, hybrid network architectures pose significant challenges to design a system platform efficiently managing data, in particular when real-time connectivity needs to be ensured across multiple wireless technologies [4] to support the reliable risk management. However, ubiquitously monitoring the worker's conditions requires a reliable mobile sensing and communications platform that ensures the wireless connectivity among the Wireless Sensor Network (WSN) nodes. FASyS has designed an end-to-end heterogeneous wireless solution that enables the continuous sensing of the working environment and the worker's health and physiological conditions in order to detect in advance any potential risks. Fig. 6. depicts FASyS's heterogeneous communications architecture for industrial environments.

To transmit the sensed data to a control centre, a wireless backhaul including medium range technologies for communications within the factory and long range technologies for the transfer of the aggregated data to the control centre has been proposed. The medium range technologies (IEEE 802.11/WiFi and IEEE 802.16/WiMAX) transmit locally sensed data (including video) from different areas of the factory towards a factory's gateway. The gateway can then transmit the received data using WiMAX and/or cellular HSDPA to a remote control centre. This architecture efficiently and reliably satisfies the requirements imposed by the industrial environment in general, and by the identified FASyS hazards in particular. The proposed architecture takes into account, not only radio propagation and communication aspects, but also semantics and security planes.

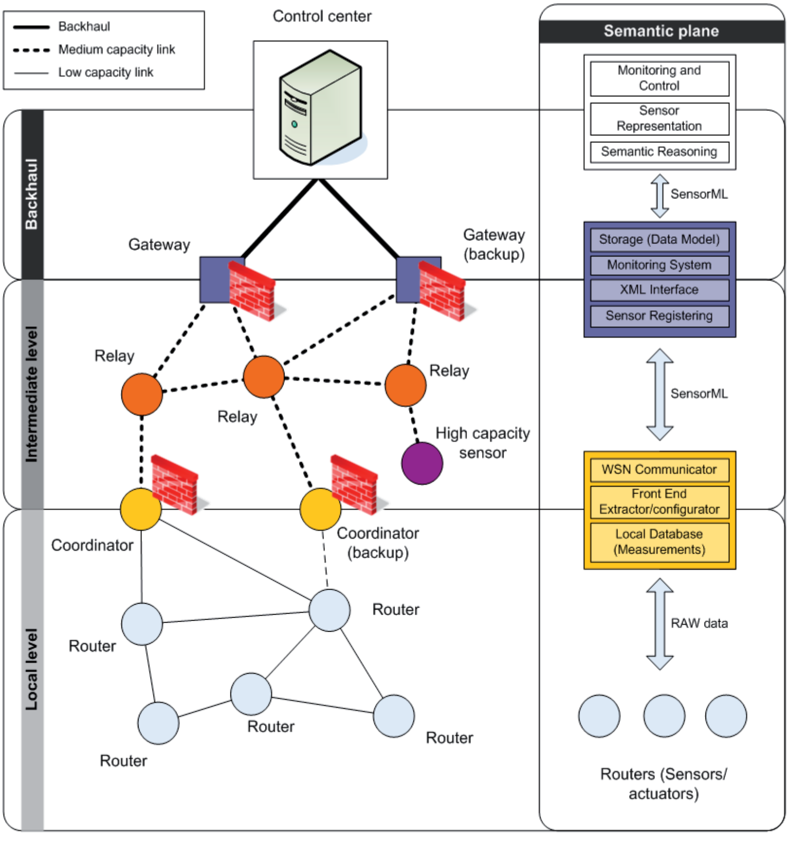
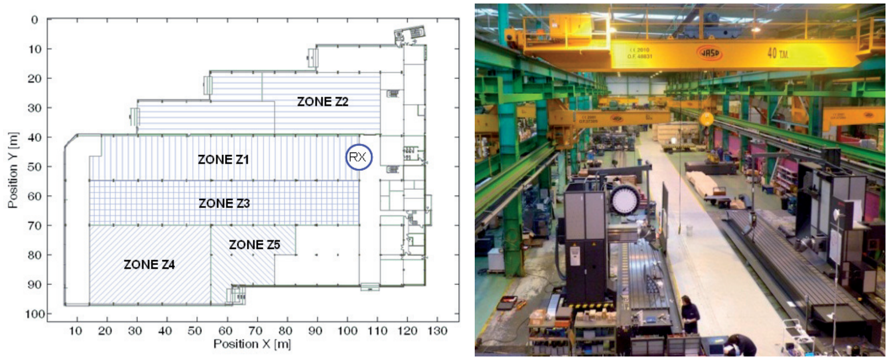


Figure 6. FASyS heterogeneous communications architecture

4.2. Industrial wireless communications and sensing connectivity

To evaluate the performance and connectivity levels of mobile IEEE 802.15.4/ZigBee [3] sensing communications, as well as the quality of service that IEEE 802.11/WiFi, IEEE 802.16/WiMAX and HSDPA technologies [4] can provide in industrial environments, a large field testing campaign has been conducted. This field testing campaign was conducted in GORATU covering a surface area of more than 10.000m² – see Figure 7a. As illustrated in Figure 7b, the plant is characterized by the presence of a large number of potential metallic obstacles that influence the radio propagation and thereby the wireless connectivity.



(a) Plan of GORATU's main factory (axis in meters) (b) View of one of the factory's corridors (Z1)

Figure 7. GORATU's main factory of machine tools.

An example, the results of one of the experiments conducted to analyse the connectivity between a TX mobile sensing mote (e.g. a mote attached to a worker or industrial vehicle) and a stationary RX base station using the MEMSIC Iris WSN motes are presented in Figure 7. In this experiment, the base station was strategically deployed with relatively good propagation conditions with the different areas of the factory (the RX base station was located at position RX in Figure 7.with an antenna height of h_{RX}=5m). During the experiments, the TX mobile node (antenna height of h_{TX}=1.2m) moved across different areas of the factory at pedestrian speed. This node was configured to periodically transmit a data packet every T seconds with a payload of 50bytes excluding headers, emulating the data transmissions of a body sensor device. Along its path, the TX mobile node experienced different propagation conditions with the RX fixed node: LOS (Line of Sight) with reduced obstructions (Z1); partial NLOS (Non LOS) due to cranes, pillars, and machinery (Z2 and Z3); NLOS due to multiple obstructing elements and high distance (Z4); and NLOS and heavy obstruction (Z5, the warehouse). The Figure below depicts the PER (Packet Error Rate) levels measured as the mobile TX mote moves around the factory. The figure shows the average PER levels experienced during time intervals of T_p=5s and the distance between the TX and RX nodes along the path; this figure differentiates the different zones of the factory (Z1, Z2,…, Z5). Additional experiments were conducted with different transceivers, antenna heights and transmission powers. The obtained results show that IEEE 802.15.4/Zigbee can provide the connectivity requirements of industrial applications, even for mobile applications. However, the transceiver, deployment conditions and locations must be carefully selected.

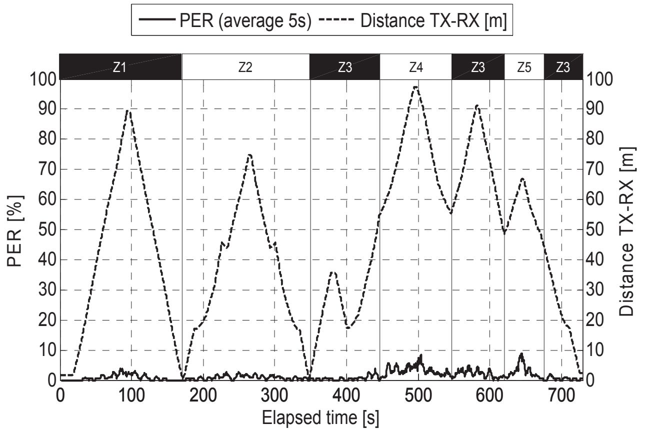


Figure 8. PER performance as a Memsic IRIS WSN mote with Pt=3dBm moves around the factory (h_{TX}=1.2m, h_{RX}=5m, T=200ms, payload=50Bytes).

4.3. Smart object semantic representation

Regarding semantics, a traversal control plane has been included in the architecture. Semantics are based in the Semantic Sensor Web paradigm, [5] and on a specific abstraction for virtual objects. Semantics, interoperability and exchange of relevant sensor configuration and information are based on Service Oriented Architecture. The key components of the FASYS semantic sensor environment are the Sensor Observation Server (SOS), a standard from OGC [6], and the HMI located in the command and control location of the risk management architecture.

The concept of semantic sensor network is used to organize, manage, interrogate, understand and control the different components of the data gathering process (i.e. network, sensors and the resulting data using high-level specifications). If semantics are introduced in the reasoning process of a FASyS subsystem, it is important to design properly the various steps of communication and interfaces if sensors and sensor networks are involved, as they impose various kinds of restrictions and limitations, such as power constraints, finite and limited memory, unreliable communication network and the quality and variability of data received.

The Semantic Sensor Web (SSW) or the Semantic Sensor Networks (SSN) base their operation on the existence of a sensor network that implements a physical layer (PHY), a sub-level medium access (MAC) and network layer (NET), usually implemented by standard protocols (e.g. Zigbee and 6LowPAN), but considering mechanisms and proprietary systems. The contribution of this type of mechanism is the addition to the data measured / generated by sensors in the form of metadata annotations of semantic information of a temporal, spatial and thematic, accessible through a Service Oriented Architecture.

The technology used in FASYS has been standardized by the OGC [6] and has been extended and specially applied to factory automation by the research team [10][7]. The concept of SSW, is based on the use of a special type of information infrastructure for web-centric collection, modelling, storage, subsequent withdrawal, sharing, manipulation, analysis and visualization of information on sensors and observation of phenomena from them. The definition of SSW by OGC is: "*Networks of sensors and sensor data storage accessible via the web, which can be discovered and accessed using protocols and application interfaces standards*"[6]. The standard and components of the SWE SOA are: Observations & Measurements (O&M); Sensor Model Language (SensorML); Transducer Model Language (TransducerML or TML); Sensor Observation Service (SOS); Sensor Planning Service (SPS); Sensor Alert Service (SAS) and Web Notification Services (WNS) [8]. FASYS considers the use of different SOS located in strategic points in the Communication Architecture in order to provide homogeneous access in the heterogeneous network to the different control applications.

FASYS semantic sensor system is based in the use of SOS and Sensor ML, it provides access to the data generated by the sensors so as the metadata to configure and customize each individual component (sensor) of the network. The main benefits of using semantic sensor networks in FASYS are: (i) Platform independence as practically any sensor or modelling system can be supported (even simulated sensors); (ii) easy development of services allowing dynamic connectivity between resources; (iii) Liaison with semantic environments, adding semantic information to the basic SWE paradigm; (iv) Traceability and support to the implementation and management of real-time measurements; (v) Flexibility in implementation: container capacity and existing sensors, implementing and processing services; and (vi) Scalability from a single sensor to a collection, individual, group or cluster of sensors. [9]

Regarding the design architecture, it is important to consider the location of the SOS within the network. Basically, there are two main approaches. The first approach considers locating the SOS in the coordinator node, as near as possible to the physical sensor. The second approach considers locating the SOS in the gateway node, as near as possible to the control center. In order to determine the most appropriate place for the SOS, it is important to consider the data flows that are envisioned between sensors and the SOS and between applications and the SOS in order to minimize data traffic. As the FASyS system uses a Complex Event Processing (CEP) system as key component of the Control Center that continuously issues requests to the SOS, the data flow is considered to be significantly higher than the data between sensors and the SOS. Therefore, the second approach has been selected in FASyS. Once the SOS has been set up in the gateway node, all sensors have to register with the SOS from each gateway to have controlled all data sources [10][11].

Regarding interoperability of systems and applications, the use of SOS provides syntactic and semantic interoperability. Interoperability is a property referring to the ability of diverse systems and organizations to work together (inter-operate). The term is often used in a technical systems engineering sense, or alternatively in a broad sense, taking into account social, political, and organizational factors that impact system to system performance. Interoperability may also be understood as the ability of two or more systems or components to exchange information and to use the information that has been exchanged. [12].

Regarding Interoperability we can distinguish two different possibilities syntactical and semantics interoperability [13]. FASYS provides both kinds of interoperability starting from the correct use of a SOS as support for merging and concentrating the information generated by sensors and distributed devices. Though strictly speaking the SOS provides syntactical interoperability, it is relatively easy to incorporate simple semantic support as temporal, spatial and thematic filtering are natively supported by SensorML and O&M, the two standard interfaces used by the SOS. Additionally, SensorML supports extensibility through annotations. If such annotations are part of a semantic vocabulary, then more complex semantic operations can be supported. SOS has been extended with a database based on a specific FASYS data model that includes some specific features not included in the OGC standard and are required to integrate the SSN in the FASYS HMI.

4.4. Mobile sensing applications

FASyS advanced IoT networking paves the ground for advanced monitoring functions that can be exploited by other parts of the FASyS risk management system to perform advanced health and safety prevention. Mobile sensing applications facilitated by FASyS include among other collision avoidance and continuous physiological monitoring.

4.4.1. Collision avoidance

Collisions between workers and fork-lift trucks, or between any type of vehicles, have been identified as one of the most common accidents in factories. Such collisions could be prevented if workers and vehicles would be equipped with WSN (Wireless Sensor Network) motes so that they can dynamically exchange information about their position and speed in real time. With this information, they could be able to detect in advance, and avoid, potential dangerous situations, such as the intersection shown in the Figure below. Robust and reliable wireless communication links should be established between any two nodes with a risk of collision, despite the potentially challenging propagation conditions, represented in the intersection by a wall, and large metallic machinery and obstructing elements placed within a large wood container at the intersection.

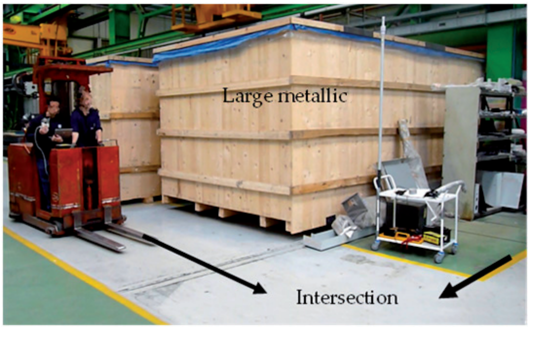


Figure 9. Collision avoidance use case: testing intersection

4.4.2. Continuous physiological monitoring module

A second example of mobile sensing application leveraged by FASyS is an innovative monitoring platform built up by up to six sensors following the HealthAlliance interface that arranges the worker physiological follow up in two strategies: (a) **intensive monitoring**, for those workers on health risk approaching minimum invasiveness as well as maximum ergonomics. (b) **preventive monitoring**, to check periodically main health indicators. FASyS is capable of integrating data coming from different sensors, working with different communication protocols and interfaces. This is done implementing Health Alliance 11073 standard logic communication and manufacturer protocol for the sensing devices. After the acknowledgment, the values are packed into HL7 standard messages and transmitted through secure pathways to the factory DPC (data processing centre), where the personalized health management module can access to it – see next Section on Occupational Personal Health Systems.

5. Occupational Personal Health Systems (O-PHS)

Most developed countries include within their basic welfare policies the right of citiziens, as workers, to be protected from sickness, disease and injury arising from their employment. Despite this intention, the International Labour Organization (ILO) estimates that 160 million workers are victims of occupational accidents and diseases every year [14] and over two million of people lose their lives from work-related accidents and diseases. The standards on occupational safety and health provide necessary tools for governments, employers, and workers to establish such practices and to provide for full safety at work. In 2003, ILO assumed a global strategy to improve occupational safety and health, which included the introduction of a preventive safety and health culture, the promotion and development of relevant instruments, and technical assistance [14].

One of the European objectives set for 2020 is the 25% reduction in the number of industrial accidents [15][16]. In order to reduce accidents it is essential to pay attention to the workers, their single workplaces and to their working conditions. In addition, favourable environments make workers feel more comfortable while they are in the factories, and thus the efficiency is increased. As a consequence, it is possible to obtain the maximum efficiency in the factory as a whole, which also produces economic benefit for the company. From a healthcare point of view, factories lack normally in an amount of enough information to allow a holistic care of the worker. Health data stored by companies are only a small amount of data, usually stored once a year, and referred to the physical condition of a person just in a particular moment [17].

5.1. Proactive paradigm for Occupational Health Systems

For these reason, future factories and enterprises need to do an effort in focusing resources and strategic planning towards making the workplace safer, healthier and to significantly reduce the number of accidents and the work related diseases in their population. In order for this to happen, it's necessary to anticipate and predict the occurrence of risk scenarios that can lead to a damaging situation, either by accident or health threat. This need generates a change of paradigm, moving from a reactive system providing management solutions to problems that have already happened and basic preventive measures, to a much more proactive model, where risk management is understood as a mostly preventive tool. To achieve this, it's important to collect, measure and analyze data during a continuous period of time, in order to evaluate the risks and their evolution.

In this line, future enterprise risk management solution should turn punctual monitoring into a more frequent and personalized vigilance, including individual and collective data. However, collecting information of many people during a long period of time requires collecting a big amount of data. People are not able to process so much information, so intelligent systems for massive data processing are needed. These intelligent systems classify data and generate alarms associated to the worker. Thanks to these alerts and all the other environmental and personal data stored, it is possible to predict health threats. Thus, it is possible to act in the most appropriate way for each worker in particular.

Nowadays, the number of sensors for monitoring personal health data is increasing. In addition, sensors that collect environmental parameters in industrial factories are being introduced more and more. The problem encountered so far, besides a reduced frequency of monitoring, is that these data are usually not thoroughly connected. The information is only collected in order to produce isolated diagnosis or identify single risks, but not common results, and the collected data become less relevant if they are not treated together. The final decision, in a dynamic environment like a factory, could be more precise if results came from a comprehensive study of a diverse set of parameters. However, once the data is collected and stored, a significant effort must be done in processing it in order to identify and highlight those elements that are directly related to present or future risks. This identification becomes more and more difficult as the amount of data analyzed increases and rage of risks augments. Therefore, it seems logical to state that, in parallel to the data collection efforts, new solutions for clustering, prioritizing and filtering information need to be put in place, to generate the most appropriate alarms in the right moment and in the right place. Risks nature can vary from emergency situations to predictive probabilities and Risk Management systems have to be able to discriminate between the two (and the whole range in between) and provide adequate communication of the contextual information so that the reaction to the risk matches the risk characteristics. Finally, once the information has been monitored and classified, the next point is focused on the intervention. With the aim of representing prevention protocols for this intervention, workflows are developed. Given the workers singularity, the adaptation of the prevention protocols is needed for each one of them. In this way, the elimination of the occupational hazard is much more effective

The Health model established in FASyS, is based on the "Ecological Concept of Disease", in which, the environment (Physical, Social, Economic and Biology, among others) is a set of external conditions and influences affecting the life and development of an organism, human behaviour or society, acting on the balance between the so-called "disease agents" and "human host", even capable of altering it and causing a potential disease situation where such agents exceed the capabilities of the host response and adaptation. This allows to establish, in accordance with the main risk factors (physical, chemical, biological, social and psychological), a pre pathogenic state or period, which we have called susceptibility stage, on which a primary prevention action is required, and other state or pathogenic period, which we called the disease, which holds the secondary and tertiary preventive actions and care actions where necessary.

From here, setting up standard profiles (comparable to health profiles), and pathogenic and pre-pathogenic profiles with abnormality parameters, becomes much more simple and affordable and allows to introduce these profiles in the system. The development and implementation of a comprehensive and interrelated system of identification and control of the elements participating in the working environment through sensory and monitoring systems, can establish and develop what we have called the SATSE or System to Aid Decision Making in Occupational Medicine. This system consists of multiple modules that collect data and information about the company, the job (Identification of potential risks, protocols to be applied, pathobiological profiles to be determined, potential limitations, referrals ...), the worker (background, demographics and psychosocial factors, medical history at work,...), extra-clinical data (organizational profile, demographic profile, psychological profile...) and additional data (diagnostic algorithms, performance algorithms, medical knowledge data bases...) The use of physiological sensors that help us to determine both in the workplace and outside it, the physiological status or health of a worker at a given time at work or outside work (whenever necessary), complements the sensor system module, that is one of the basic elements of the project.

5.2. Personal Health Records (PHR)

Health data has been traditionally produced and owned by the Healthcare Systems and stored in Electronic Healthcare Records (EHR), focusing mainly in describing clinical procedures, tests and values. However, with the purpose of improving the characterization of the person and his environment, EHR data needs to be extended. This new information is stored, together with the EHR information, in other repositories. These repositories are known as PHR (Personal Health Record) [19] and collect data such as habits, preferences, information about the family, work, moods or nutritional profile. These repositories are, in opposition to the EHR, owned by the person, who has the option to share it with whoever he chooses.

The usage of PHRs in occupational health could enable that when an employee goes to work in a company for the first time, the enterprise’s health professionals can ask him to share his relevant PHR data, in order to have his personal file more complete and enable a more complete and accurate health risk management. Of course, this situation would require enhanced methods for privacy and data protection, ensuring no unauthorized and adequate usage of the health information is made. In general, current PHRs contain a summarized version of the EHR adapted to the patient’s knowledge and needs and, in some cases, home monitoring data. Future PHRs covering the area of occupational health should be based in the following aspects:

- Includes relevant health hearths and risks linked to the work conditions.
- It allows patient to introduce data (automatically or manually).
- It allows an exchange of information with the healthcare system (HER).
- It includes an option to generate summaries to share information with other PHRs.

One of the advantages, for example, would be when a worker goes to work in other factory. If the new factory is enabled, his PHR could be downloaded in the system of the new factory in order to have a more complete file and ensure continuity in the management of the risks for that particular worker. Stored Data can also be extracted for consultations and referrals in case health professionals need to, of course under the corresponding access control that prevents unauthorized sharing of the data. Data can be easily anonymized to be used for statistical and epidemiological studies in order to detect population based health problems.

5.3. Care plans, workflows and medical guidance

The normalization of processes is more and more present in current society. The formal definitions of procedures that are usually deployed in enterprises and factories are considered the best practices in order to support the management. This normalization allows not only predefine which is expected in the organization procedures but also allows a continuous monitoring of the processes that is crucial for their correct management and continuous improving. Moving these ideas to Health Care, Care Plans or Clinical Pathways [20] there are protocols for standardization of health processes. Nevertheless, the standardization of those processes is more complicated than usual enterprise processes. First of all, the Health processes are very complex. The high number of variables that are taken into account in a care process, the pluripatological patients and the wide quantity of different treatments that can be applied to them requires a special expressivity in the process definition. In addition, the processes standardized should be designed by health professionals. This requires that the specification language should be legible by those experts in order to ensure that the process can be understood and repeated. Moreover, the specification of the process should be non-ambiguous. This is a key problem in the specification of protocols. It is said that a protocol is ambiguous when more than one interpretation possible for the same specification exists. The presence of ambiguity in Care plans is a great problem that prevents that the process will be deterministic.

Usually, care plans are described as great manuals, free text written, that explain the care processes in a whole. Those manuals are available through big medical libraries like the Cochrane [21] or PubMed [22]. Although those manuals have the expressivity of natural language, they are often ambiguous and the high quantity of data is tedious to read. Other approaches are based on specific formal languages like GLIF [23] that provide tools to avoid ambiguity and to ensure the completeness of the protocols defined. Nevertheless, the use of rule based systems can difficult the creation of legible and controllable frameworks because the high number of rules to be taken into account. In enterprise environments, processes are usually defined as workflows. Workflows [24][12] are formal specifications of processes designed to be automatized. The main advantage of using workflows is that they usually have a graphical interface that makes easier their design and understandability to non-programming experts like doctors. In Care Plans environment there are works, available in the literature, which faces this problem using workflows technology [25][26]. The main problem of workflows against other approach like GLIF or traditional techniques is that workflows have less expressivity than them. Nevertheless, there are available workflows approaches in literature [26] that ensures a high expressivity for defining very complex workflows and even for the design of clinical pathways [27].

In addition to graphical design, Workflow has more advantages that can be useful for the design and deployment of care plans. Current workflow systems usually have associated an engine able to automatically execute the processes defined in the graphical way. That means that the formally defined processes can be automatically used for deploying the process by using automatic deploying systems. In addition, Process mining [27] technologies allows the application of pattern recognition technologies to support the iterative design of Care Plans. Furthermore, thanks to the low grammatical complexity of some workflow approaches [27] is possible to apply a great quantity of algorithms and tools for ensuring the completeness, the non- ambiguity and the simulation of processes in order to detect problems in their design before their deployment. In the Figure below is presented a basic specification of a Care Plan using a workflow based approach.

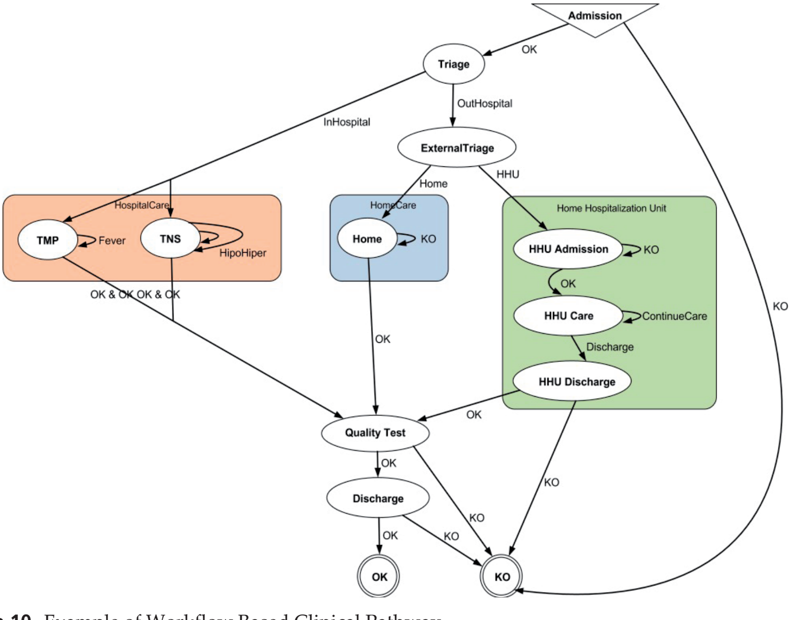


Figure 10. Example of Workflow Based Clinical Pathway

The continuous assessment of processes is critical for ensuring an efficient execution of the enterprise procedures. In this way, there are more and more business intelligence systems available to empower and support the management providing important information about the processes filtered by using Data Mining technologies. Usually those systems, present information about static or evolution of numeric data according to parameters that can, indirectly, help managers to detect inefficiencies or bottlenecks in the processes. In this scenario, an emerging technology is growing in order to enrich those business intelligence systems providing a more directly view about the process execution. This technology is called Process Mining (A.K.A Workflow Mining) [28]. Process Mining technology is research field, based on pattern recognition paradigm, that uses the events or activities of the process logs in order to automatically infer a graphical workflow that explains the actual execution of the process. There are many process mining algorithms in the literature, based on Events like Alpha [16] or Genetic Process Miner [29], or based on activities like PALIA algorithm [27]. Those algorithms are able to create workflows from samples and present graphically in order to know exactly how processes behave in real implementations. Comparing the results of those algorithms with the designed processes, it is possible to directly know what are the most usual paths followed by the processes within the designed workflows, what are the differences and exceptions to the designed processes that are occurring in the implantation of them, etc. In this way, there are available algorithms that allows the comparison between the designed workflows and their real implantation [27].

6. Real-time risk management solutions

The implementation of proactive risk management in a sensing enterprise demands that distributed reasoning capabilities are provided as a means for intelligence and analytics. FASyS should therefore not stop simply at the data collection and distribution level, but on the contrary, it should be able to analyse the vast amount of available information in the presence of data unavailability, ambiguity, imprecision and error. Therefore, for being able to detect, decide and act in real-time to dynamic risk levels various challenges need to ba addressed in the areas of (a) iindustrial safety reasoning engines (b) Real-time risk detection tools (c) Personalized decision support tools and (d) Semantic solutions for services coordination

6.1. Industrial safety ontologies and reasoning engines

To deal with safety in a quite heterogeneous environment in terms of information sources, it is required to create a formal data specification that will be used within ontologies structures that support risk management and context. The ontologies are used as a knowledge base for reasoning engines that are responsible for the detection of risk and responsible of the implementation of relevant actions committed in any situation [30].

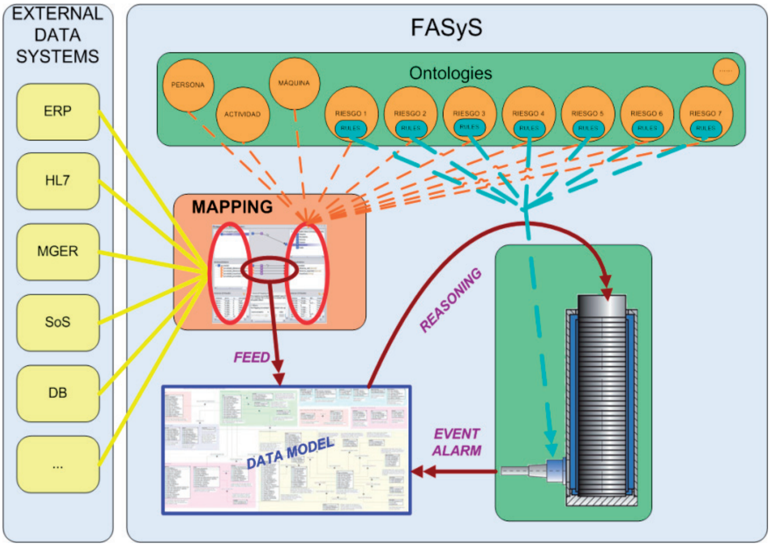


Figure 11. Ontology-based reasoning

As depicted by the Figure above, the selected architecture is based on a distributed multi-ontology structure. This way, there is an ontology for each risk and a set of context ontologies related to different existing and influent entities in each cycle of risk management in the factory. The context ontology properties might be source of different risk ontologies properties. Thus, risk can be defined through different entities or actors of the company.

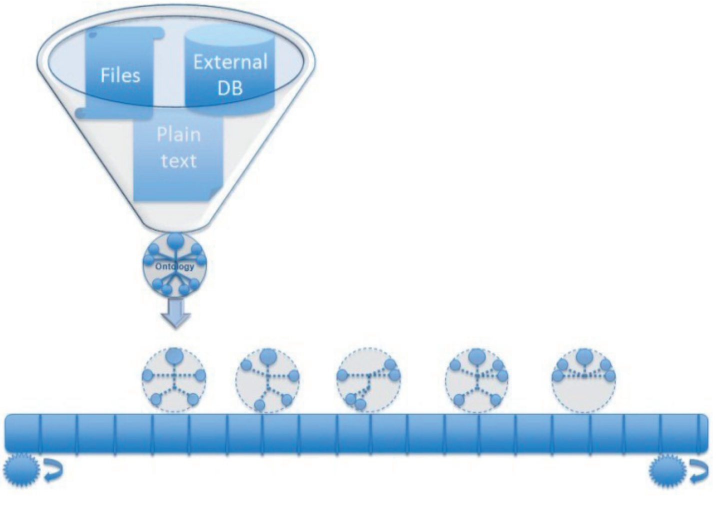


Figure 12. Mapping process results

The use of ontologies is related to the introduction of intelligence and reasoning on the information processing and the relation established between concepts that increase the relevance of such intelligence. Furthermore, this structures will be used for mapping data with information systems of the own company or potential external information systems [31]. The ontologies are considered therefore as the bridge between the heterogeneous information ecosystem and the actual services implementing the FASyS system logic.

6.2. Real-time risk detection tools

Inherent to the concept of proactive prevention is the aim of the system to predict a particular risk by evaluating all the variables of the worker, the working place and the environment. For being able to perform such evaluation the first requirement is to have tools that are able to process all the information in a complete, efficient and, at the same time, very light way. Currently, risk management is focused on monitoring the proposed actions after an evaluation process is periodically; e.g. yearly, performed in the companies. With nowadays facilities, there is no way to propose a 'real time' evaluation of the situation, due to the fact that the safety manager cannot control all the time what is happening in the shop-floor, the data of all the machines involved in the process and the information referred to the state of workers that are moving around the factory. This is inconceivable.

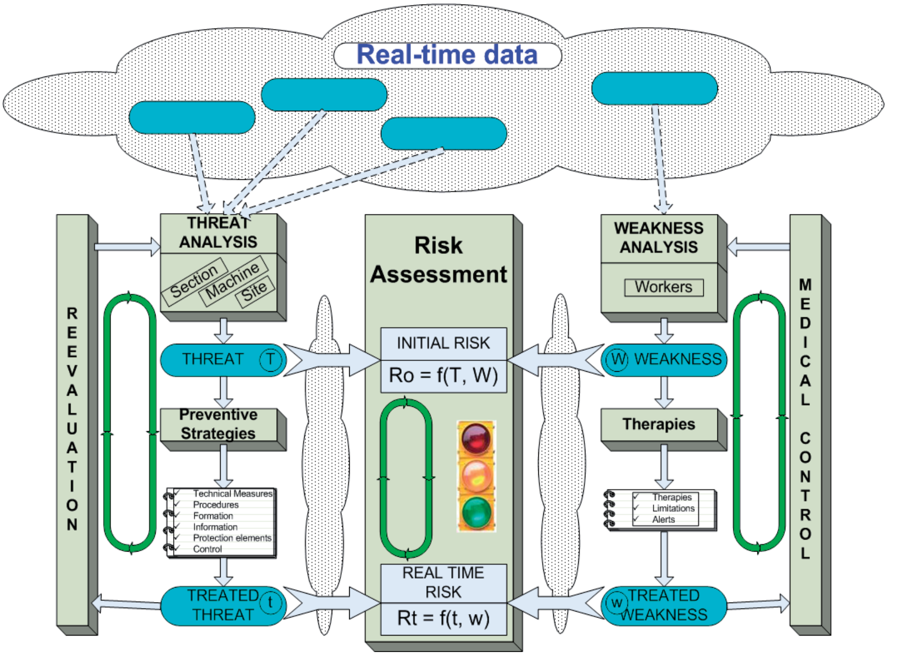


Figure 13. Continuous real-time risk management based on Complex Event Processing Technology

However, in the context of the sensing enterprise, ubiqutous smart object deployment and the availability of suitable universal virtual object abstractions make such control possible. Thus, a new risk management cycle can be designed based on the capability to master big data stream technologies in a scalable manner. FASyS has designed a consolidated framework for risk detection. FASyS does not only leverages effective data management but also consolidates an integrated approach to health (medical) and safety (security) risk management. The consolidated approach is shown below.

FASyS has proposed a complex event processing (CEP) [32] unit network that through a number of pre-set patterns, in the form of a complex formula compound by an undetermined number of factors, will feed and evaluate the patterns in a continuous way, creating alerts based on particular thresholds that have been previously defined and particular set of actions taken concurrently or sequentially place. These tools allow for big data volumes processing with really low computing infrastructure requirements.

However, FASyS data processing solutions go beyond big data volume. FASyS looks for flexible solutions that can create complex feedback and feedforward loops across CEP units to ensure that the time variable, the event frequency or event correlations-workflow can be processed at high speed.

6.3. Personalised decision support tools

As it has become apparent from the previous sections FASyS provides the tools and models for being able to process as much as information in less time as possible. Thus, the enterprise safety and healthy manager can work with relevant information to make informed decisions. The aim of FASyS personalised decision support tools is not just to warn and make apparent a particular risk level but also to ease the decision process based on strong knowledge support. FASyS Monitoring and Control Human Macine Interface (HMI) has therefore being designed to provide highly visual interfaces about risk levels. Moreover, the system also makes suggestions of the most suitable procedures to be aplied when a risk situation is detected, so that the reaction time can be hugely reduced and the user can send a highly effective execution action plan immediately.

The decision support system works with risk patterns that require a human interaction, either to provide additional information or to select preferred option in front of a multiple selection.

6.4. Semantic solutions for services coordination

In the context of the sensing enterprise, FASyS has to deal not only with the detection of risks but also has to support the actuation and deployment of the preventive actions selected by the safety and healthy manager through the personalised decision suppport tools. This implies that FASyS has envisaged a service oriented scenario, where the factory is populated by a large amount of services that exchange messages and perform are coregraphed or orchestrated to perform the designed actions by means of smart objects.

Therefore, in the FASyS platform, there is a huge amount of available services involved in risk management life cycle. In addition, those services have heterogeneous sources; they can become available, temporarely unavailable or even disappear suddenly; the availability of them can change anytime. In order to solve these situations, FASyS has proposed a highly effective service messaging and service management and coordination semantic solution that would use choreography techniques focused on browsing FASyS service topology [33], which is made using an ontology definition; e.g. through WSDL or USDL descriptions. With this solution, FASyS is able to adapt its reactions to available services at any time and ensure the best possible service performance based on the precedence of the risk to be addressed and the service load in the enterprise bus. Clustering techniques allow for optimum selection of services to be orchestrated otr coreographed to servce a particular aplication in the prevention workflow. Those services could be previously known or even configured at run-time.

All of FASyS technological developments and systems have a semantic service library that would ensure the availability of its own functionalities to the rest of systems. The access to those functionalities will be assured continuously [34].

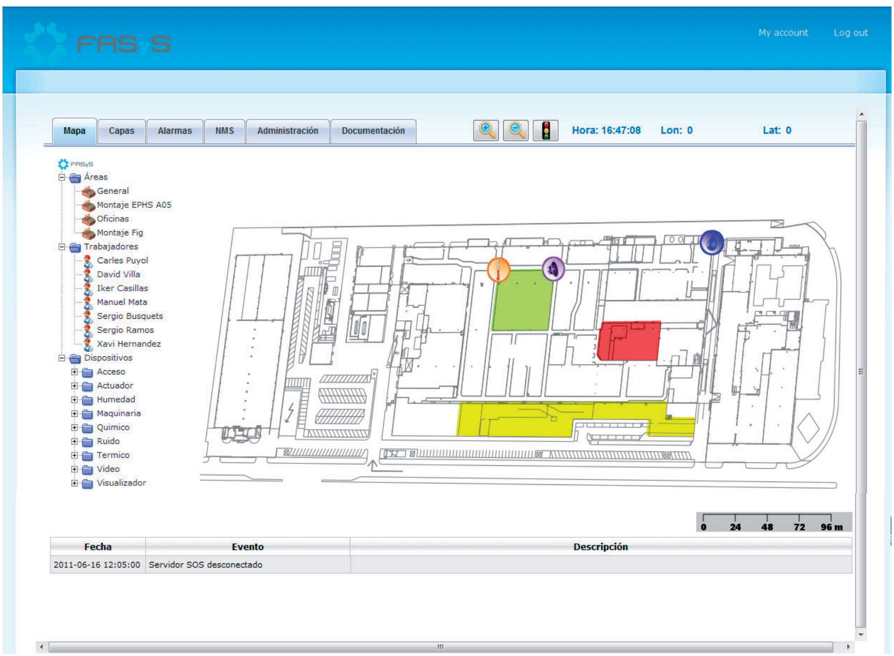


Figure 14. FASyS Monitoring and Control Human-Machine Interface (HMI)

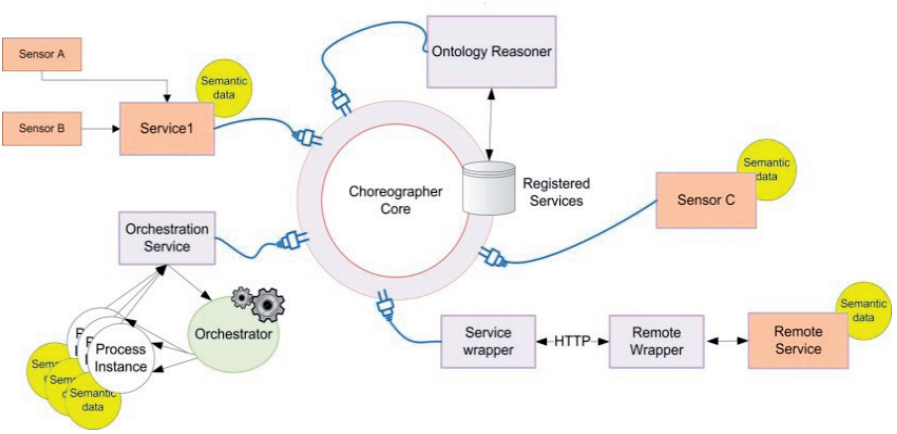


Figure 15. FASyS processes choreographer

7. Conclusions

This paper has presented the rationale behind the development of a risk management framework for personalised risk management in the context of the sensing enterprise. The paper has presented the main dimensions proposed for the model and it has presented the main technical components.

The paper has introduced the reference architecture and it has argued how this reference architecture is in complete alignment with European IoT movement currently under development. Moreover, the paper has provided evidence in terms of how the FASyS system is capable of providing a personalized and intelligent management of all the factors related, directly or indirectly, to the worker and its environment, in order to identify and detect warning situations, alerts and propose immediate actions required upon a worker, a machine or an area.

Author details

Óscar Lázaro and Alicia González
Innovalia Association, Spain

Agustín Moyano, Mikel Uriarte, Óscar López Etxahun Sánchez and Saioa Ros
Nextel S.A., Spain

Teresa Meneu, Juan Carlos Fernández-Llatas and Vicente Traver
Soluciones Tecnológicas para la Salud y el Bienestar (TSB), Spain

Benjamín Molina and Carlos Palau
Universitat Politècnica de Valencia (UPV), Spain

Antonio Moreno and María González
Sociedad de Prevención de FREMAP S.L., Spain

José Antonio Palazón, Miguel Sepulcre and Javier Gozálvez
Uwicore, Ubiquitous Wireless Communications Research Laboratory,
Universidad Miguel Hernández de Elche (UMH), Spain

Luis Collantes and Gonzalo Prieto
INDRA Sistemas S.A., Spain

Acknowledgement

This work has been partly funded by the CDTI through CENIT Project FASyS: Absolutely Safe and Healthy Factory (No. CEN-20091034). The authors wish to acknowledge the CDTI and the Ministry of Innovation for their support.

8. References

[1] Li M., Mazura M., Martínez C., Missikoff M. Future Internet Enterprise Systems (FInES) Position Paper; 2011. Accessible from: http://cordis.europa.eu/fp7/ict/enet/documents/fines-position-paper-fp8-orientations-final.pdf (accessed 16 May 2012).

[2] A. Willig, "Recent and Emerging Topics in Wireless Industrial Communications: A Selection", *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102-124, May 2008.

[3] M. Sepulcre, J. A. Palazón, J. Gozálvez and J. Orozco, "Wireless Connectivity for Mobile Sensing Applications in Industrial Environments", *Proceedings of the 6th IEEE International Symposium on Industrial Embedded Systems (SIES'11)*, 15-17 June 2011, Västeras (Sweden).

[4] J. A. Palazón, M. Sepulcre, J. Gozálvez, J. Orozco and O. López, "Heterogeneous Wireless Connectivity for Fixed and Mobile Sensing Applications in Industrial Environments", *Proceedings of the 16th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'11)*, 5-9 September 2011, Toulouse (France).

[5] A Sheth, C Henson, S S Sahoo, "Semantic Sensor Web" IEEE Internet computing, vol. 12, no. 4, pp. 78-83, 2008.

[6] Open Geospatial Consortium. Sensor Web Enablement Working Group Project. http://www.opengeospatial.org/projects/groups/sensorwebdwg (accessed 11 May 2012).

[7] S. F. Pileggi, C. E. Palau, M. Esteve, "Building Semantic Sensor Web: Knowledge and Interoperability", *Proceeding of the First Semantic Sensor Web Workshop (SSW 2010)*, 15-16 October 2010, Valencia (Spain).

[8] J.L. Martínez-Lastra y I. Delamer, "Semantic Web Services in Factory Automation: Fundamental Insights and Research Roadmap", *IEEE Transactions on Industrial Informatics*, Vol. 2, No. 1, February 2006.

[9] W. Ruh, "A Model for a Semantic Enabled Network", Proceedings of 22nd International Conference on Data Engineering, April 2006.
[10] L.M. Ni, et al., "Semantic Sensor Net: An Extensible Framework", *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 4, No 3/4, April 2009
[11] A. Kansal, S. Nath, J. Liu y F. Zhao, "SenseWeb: An Infrastructure for Shared Sensing", *IEEE Multimedia*, Vol. 14, No. 4, pp. 8-13, July 2007.
[12] J. Park and S. Ram, "Information systems interoperability: What lies beneath?", ACM Transactions on Information Systems (TOIS), Vol. 22, no. 4, October 2004
[13] V. Mojtahed, M. Eklof and J. Zdravkovic, "Towards Semantic Interoperability between C2 Systems Following the Principles of Distributed Simulation", Proceedings of 16th ICCRTS Conference on Collective C2 in Multinational Civil-Military Operations, Quebec City ¿Canada?, June 2011.
[14] International Labour Organization. Occupational Safety and Health. http://www.ilo.org/global/standards/subjects-covered-by-international-labour-standards/occupational-safety-and-health/lang--en/index.htm (accessed November 2011)
[15] Balogh I, Orbaek P, Winkel J, Nordander C, Ohlsson K, Ektor-Andersen J, et al. (2001). "Questionnaire based mechanical exposure indices for large population studies-reliability, internal consistency and predictive validity". Scand J Work Environ Health; 27(1):41–48.
[16] Leijon O, Wiktorin C, Harenstam A, Karlqvist L, MOA Research Group (2002). "Validity of a self-administered questionnaire for assessing physical workloads in a general population". J Occup Environ Med; 44(8):724– 735.
[17] J. Stranks, 2006. The manager's guide to health & safety at work. London: Kogan Page Limited
[18] Benny Lo. "E-AR (ear-worn Activity Recognition) Sensor". http://vip.doc.ic.ac.uk/benlo/m775.html (accessed 11 May 2012)
[19] PHR Reviews. Available from: http://www.phrreviews.com/
[20] S. Audimoolan, M. Nair, R. Gaikward, and C. Qing, "The Role of Clinical Pathways," Improving Patient Outcomes, vol. February, p. –, 2005.
[21] T. C. Collaboration, COCHRANE Library: http://www.cochrane.org/index.htm.
[22] P. Library, National Library of Medicine and The National Institutes of Health PubMed Library: http://www.pubmed.gov.
[23] M. Peleg, A. A. Boxwala, E. Bernstam, S. W. Tu, R. A. Greenes, and E. H. Shortliffe, "Sharable Representation of Clinical Guidelines in GLIF: Relationship to the Arden Syntax.," Journal of Biomedical Informatics, vol. 34, no. 3, pp. 170–181, 2001.
[24] WfMC, Workflow Management Coalition Terminology Glossary. WFMC-TC-1011, Document Status Issue 3.0, 1999.
[25] M. Sedlmayr, T. Rose, R. Rohrig, and M. Meister, "A workflow approach towards GLIF execution," in ECAI 2006 workshop: AI Techniques in Healtcare: Evidence-Based guidelines and protocols, 2006, pp. 74–76.

[26] C. Fernandez-Llatas, S. F. Pileggi, V. Traver, and J. M. Benedi, "Timed Parallel Automaton: A Mathematical Tool for Defining Highly Expressive Formal Workflows," in Modelling Symposium (AMS), 2011 Fifth Asia, 2011, pp. 56-61.

[27] C. Fernández-Llatas, T. Meneu, J. M. Benedi, and V. Traver, "Activity-Based Process Mining for Clinical Pathways Computer Aided Design," in 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2010.

[28] W. van der Aalst, A. Weijters, and L. Maruster, "Workflow Mining: Discovering Process Models from Event Logs," IEEE Transactions on Knowledge and Data Engineering, vol. 16, pp. 1128–1142, 2004.

[29] A. K. A. de Medeiros, A. J. M. M. Weijters, and W. M. P. van der Aalst, "Genetic Process Mining: A Basic Approach and Its Challenges.," in Business Process Management Workshops, 2005, pp. 203–215.

[30] Gruber T. R. "A translation approach to portable ontology specifications". Appeared in Knowledge Acquisition, 5(2): 199-220, 1993. Available from: http://tomgruber.org/writing/ontolingua-kaj-1993.pdf (accessed 11 May 2012)

[31] Information Mapping. "A research note by Namahn". Available from: http://www.namahn.com/resources/documents/note-IM.pdf (accessed 11 May 2012)

[32] Complex Event Processing. "What's behind CEP's Leaps and Bounds? Automation" http://www.complexevents.com/2012/05/08/whats-behind-cep%E2%80%99s-leaps-and-bounds-automation/ (accessed 11 May 2012)

[33] Howard R., Kerschberg L. "A Framework for Dynamic Semantic Web Services Management". Special Issue on Service Oriented Modeling, International Journal in Cooperative Information Systems. Available from: http://eceb.gmu.edu/pubs/IJCIS_Howard_Kerschberg.pdf (accessed 11 May 2012)

[34] Carlos Fernández-Llatas, Juan B. Mocholi, Agustín Moyano and Teresa Meneu. "Semantic Process Choreography for Distributed Sensor Management". Proceeding of International Workshop on Semantic Sensor Web, SSW 2010, in conjunction with the 2nd International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2010, 25-28 October 2010, Valencia, Spain.

Section 4

Risk Management Practice Across Different Projects and Industries

Chapter 14

Strengthening Risk Management in the Midst of Downturn Times

Amparo Marin de la Barcena

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51008

1. Introduction

Increasing interest towards Risk Management improvement and timely identification of firms' failure are driving demand for the set-up of risk management principles, particularly within the financial and banking industries. Although there is no-one accepted definition for appropriate and infallible risk management principles, research in this field has led to identify best practices that have proven successful results at a variety of industries. Zooming into the Financial Sector, this case study is aimed to present six tips which have strengthened resilience to bankruptcy and economic losses at a leading European Bank which has depicted positive growth in the midst of the Downturn and world-wide economic recession.

2. The Banking Business and the Importance of Risk Management Policies

Financial Institutions' turnover is influenced by their risk management policies and the target risk profile they want to achieve in the market. For instance aiming towards a low-medium predictable risk profile, jointly with geographical diversification can be key elements of differentiation for a firm to achieve a leading position ahead of other industry players. [1]

The effectiveness of Risk management policies enacts as a firm's identity card, specifically in downturn times. This largely explains why seeking continuous improvement of risk management policies has stayed as a top priority for key industry players which have shown positive growth in the last years of financial turmoil.

"Risk Management means looking into the future" [2]

Fuelled by increasing competition, more stringent regulations and higher levels of interconnection amongst different players, the marketplace is becoming more and more complex. Concerned with such trend, financial institutions have realized that if they want to survive, they need to rely on a robust risk management framework, given by corporate organizational policies and supported by sophisticated IT systems and technology.

Since 2008, financial markets have experienced sharp declines. [3-5]. Analyzing several consumer surveys allows us to conclude that consumers, particularly those who have been severely affected by the crisis, would rather have opted for higher levels of prevention and anticipation, supported by management schemes with enough level of risk management considerations. The aforementioned risk management considerations are playing a greater role in companies' overall strategy and strategic choices. Indeed, companies are increasingly realizing the need of relying on strong risk governance and a sophisticated hedging program. This trend is largely explained due to the fact that these are seen as essential elements of a robust financial risk management program that will help them secure resilience to the economic problems affecting the marketplace.

2.1. To what extent is ERM a "New" concept?

In the last years ERM has been gaining momentum as a new concept, a new trend, a MUST HAVE. But perhaps in the midst of everyone rushing-up to implement it, it is worth stopping for a second in order to understand its meaning and ask ourselves … is it really something new?

ERM stands for Enterprise Risk Management but its real origin is unclear. Some associate the birth of this concept to the definition of the Integrated Risk Framework by the Committee of Sponsoring Organizations (COSO) back in September 2004. Others believe that the starting point was in the 1970s when most of the Management theories arised. Being factual it would even be valid the mindset that it is a concept that has been in place since the beginning of times, when the Romans traded goods with the Carthagians. So if we agree that companies have managed risk forever the question here is what can Enterprises do better to manage Risk? Or what have been the best practices in risk management that companies have pursued to achieve recognized success?

Anyone involved in line management makes risk based decisions on a daily basis. The desire of such decision makers would be to count on a reliable crystal ball that would allow them to opt for the choices that would derive profit maximization.

Far from being the aforementioned magic tool, going down the road of ERM involves setting a structured approach or a reference framework to manage risk at companies. As of today, many companies have already understood the value proposition and experienced its benefits – which mainly tackle improved controls, better communications and decision-making, and a common language for risks. [7]. Nevertheless, for other companies the concern is still there: is there significant justification for an ERM program? Would the company really make different decisions if it did have an ERM program? According to some literature, a good ERM program has proven to enhance the company value through reduced costs, decreased variability in financial results, enhanced market reputation, and improved business decision-making. Such dimensions can even be measured and what is more, the Credit Rating Agencies have started considering these factors as part of their company assessment process. [8]

3. Elements of a robust risk management framework

Generally speaking, a robust risk management framework tackles five dimensions [9]:

1. Organizational culture
2. Processes
3. Technology
4. Risk measurement
5. Monitoring

3.1. Corporate governance = Positive risk culture

Relying on strong corporate governance that diffuses a positive risk culture from the top to the bottom of the organization is an imperative to set up a robust risk management framework. Unless the employees of a firm are aligned and somehow willing to contribute towards improving risk management at their workplace, the organization will not be able to manage risk appropriately. The whole organization, from the last employee to the most powerful member of the Board of Directors should see some benefits in managing risk, and as long as this happens, a positive risk culture can be promoted across the overall organization.

From another angle, a positive risk culture can also be understood as that which is characterized by individual accountability, creativity, transparency and honesty. As long as employees understand the importance of accomplishing the organization's risk management approach, have a clear view of what such policies mean and believe they are transparent and honest, they will feel more creative and a healthy attitude of constant challenging of decisions and ideas will arise from them towards improvement the management of risk at their workplace.

3.2. Policies = Procedures

The second key element of a sound risk management framework is to have in place a set of policies and procedures, which are consistent with the organizational culture and clearly defined, with enough level of detail.

In a simplified view, the output of a firm's risk management strategy is determined by the balance between the firm's commercial objectives (appetite for risk) vs. its risk contention goals (risk control). In other words, organizations look for profit maximization at the lowest risk possible. (Figure 1)

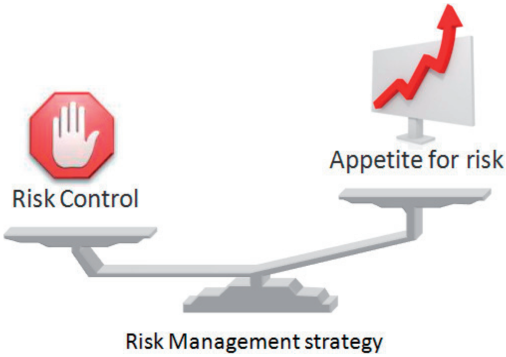


Figure 1. A company's Risk Management strategy should balance its appetite for risk over its measures for risk control.

Organizational processes are the means that a firm has to achieve its strategic goals. Therefore, firms' corporate policies should be well assembled, known by all and periodically updated not only with industry best practices, but also with the new regulation.

3.3. Data utilization = Efficient technology

Technology is changing the way people access and exchange information [10]. Information related to a firm's performance is a key success factor [11] and thus technological capability to extract such data can drive competitive advantage. [12].

Data collection and its enrichment are often tedious tasks but critical in any risk analysis phase. Such tasks have often a high manual processing component, which can be an inherent source of errors and productivity constraint. However, through technology and process automation, efficiency levels can significantly improve.

Timely and accurate information ease decision making, and therefore, counting on a well structured information management flow that incorporates advanced technology developments contributes extensively towards minimizing lead times whilst increasing a firm's level of pragmatism and effectiveness to meet its strategic objectives.

Additionally, it is worth referring, not only to the information gathering and processing dimensions, but also to how and when the information is communicated. Only when the information reaches the target recipient it can be utilized and in this context it is relatively obvious that automation can play a key role – let's assume that a piece of information is identified, checked, enriched and worked through; the outcome will only be useful if it reaches the right people at the right time which is achievable if there is for instance an automated process that can instantly distribute the information.

3.4. Risk measurement = Value-at-risk

The environment in which a firm operates is always subject to an uncertain level of risk, and it has to deal with it. A firm's know-how to measure risk levels is one of its main assets to understand to what extent it is worth to incur a certain risk and can even pre-empt a corporation from bankruptcy.

There are several risk measurement techniques which have been used for long. Some of the most common ones in the banking industry are:

- Size of open positions
- Degree of maturity mismatch in the net position
- Assets' exposures

The calculation of such measures often needs of many complex mathematical models. Furthermore, in some cases, their implementation at organizations is not easy. Therefore, since the late 1990s, companies have alternatively opted for another method: "Value-at-Risk" (VaR).

VaR calculates the risk of loss on a specific portfolio of financial assets. For a given portfolio, probability and time horizon, VaR is defined as a threshold value such that the probability that the mark-to-market loss on the portfolio over the given time horizon exceeds this value (assuming normal markets and no trading in the portfolio) is the given probability level [13]. See example in Figure 2.

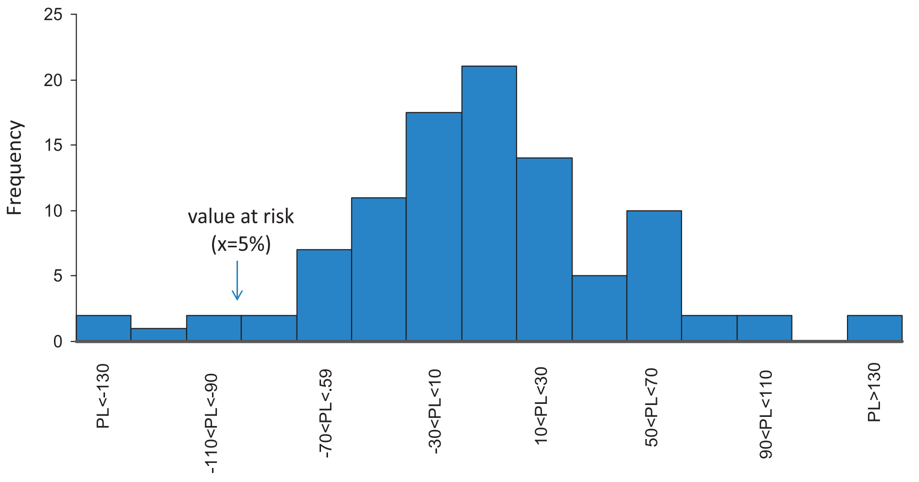


Figure 2. Histogram of daily hypothetic profits and losses over a long-term contract [14]

At this point it is worth highlighting that whilst VaR provides the quantitative perspective, it needs to be complemented by a more qualitative dimension resembling the market changes ("what-if scenarios" or "stress tests").

It is also worth pointing out that despite VaR is used in front of other methods which are perceived less powerful and more complex and difficult to understand, it is also constrained and only valid within a set of assumptions. This is one of the reasons why it is adviced that every VaR model is backtested, validating that the actual results are aligned to the forecasts.

3.5. Supervision = Ongoing monitoring

The level of risk undertaken by a firm derives from its ability to monitor it. High capability to monitor risk helps to optimize the risk taking process. Risk monitoring should not only rely on the automated output of the VaR calculation, but should also be supported by a risk management function that is able to do a correct interpretation of the results. Risk monitoring aids in risk assessment and can result invaluable to senior management when they need to take a decision and they look for advice.

At this point it is worth pointing out that although risk monitoring has mainly focused on risk avoidance, there is another dimension: "the upside" or opportunity aspects for risk. This dimension is scarcely looked at, but on the other hand it can, if appropriately managed, derive significant level of profits.

Most frequent means to monitor risk consist on processes, reports, and discussion venues. However to ensure their reliability, these must be backed by a solid function of data management and a usable model.

All in all, monitoring implies tracking and thus, the more effort and time consuming this task requires, the later the relevant information will pop up and the lower the level of responsiveness to potential problems.

4. Approaching the economic crisis: 2008 – today…

Although Risk Management has been for over a decade a central part of many organizations' strategic management [15], the trend towards adoption of more sophisticated techniques and technology, together with the bet for new models that can improve firms' Risk Management has been gaining pace. In general, the last years have seen rapid growth as concerns to the development and improvement of Risk Management tools which facilitate the evaluation of customers' credit worthiness, prediction of customer behavior and propensity of default [16-21]. Until 2008, a significant amount of investigation efforts within the financial environment had been towards providing assistance to banking agents on their daily decisions [22], but with the emergence of the economic crisis, researchers and practitioners began to strive even further to excel at their developments and come up with innovative and more diversified portfolios of Risk Management tools.

The economic downturn is seen as a worldwide phenomenon started in the US in 2008 and still ongoing. It is widely thought to be the result of a complex conglomerate of factors [23-26]. (Figure 3)

Such factors have been triggered by mistakes or defaults in Regulation and monitoring, as well as in the market behavior. Low interest rates jointly with financial innovation and globalization trends have evolved towards leverage and underestimation of liquidity and credit risk. The outcome of this mix has been excess indebtedness and excess asset valuation which has derived into a context characterized by: bankruptcy, recession, inflation, collapse in the Stock Markets, currency exchange volatility, dramatic decreases of monetary policy rates, slowdown of demand and consumption, boost of unemployment rates and shortages in profits.

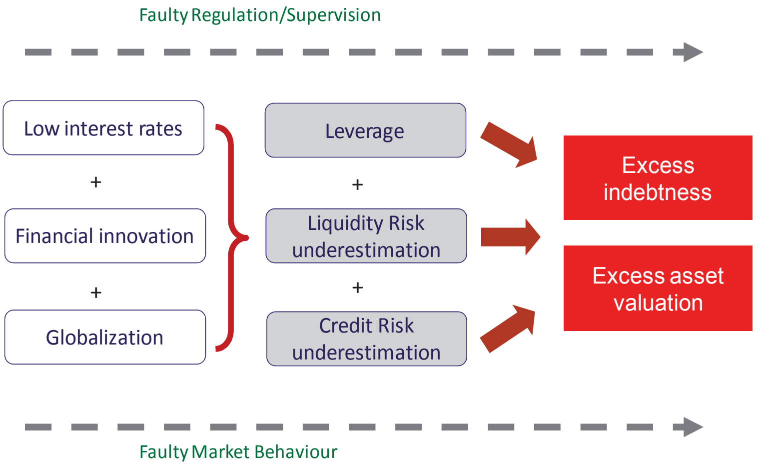


Figure 3. Possible root causes of the crisis

In order to recover from the crisis, regulators and country heads worldwide have convened on the need of regulatory changes to reform the market risk – Basel III. But restructuring the financial sector jointly with strengthening international regulation policies are necessary but not sufficient conditions to reverse the economic situation. These need to be supported by financial institutions' risk management models.

Consequently, awareness of the fact that a risk culture can constitute a very important safeguard against financial losses and collapse has risen amongst industry players. In particular, Banks and most financial entities have demonstrated particular high interest in the last years on the implementation of appropriate operational and control models aimed at an adequate management of risk.

5. Seven dreadful mistakes in risk management

Beyond looking at the economic crisis from a pessimistic angle, it is undeniable the existence of a brink of light considering the whole context – indeed the financial turmoil of the last years can be seen as an opportunity for change and learning from own or peer mistakes [27].

- So, what are the lessons learned?
- What have been the main mistakes that have led even former leading industry players to failure? (Figure 4)

5.1. Excess = Aggressive policies in risk lending

> *"… lend, lend, lend … we'll first get hold of the customers and afterwards make a fortune out of the interest rates …"*

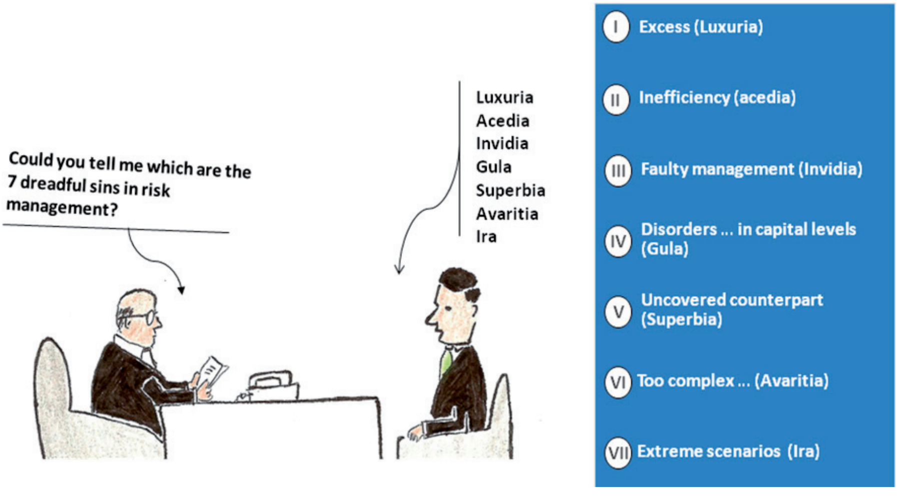


Figure 4. Seven Dreadful Mistakes in Risk Management

Excessive risk lending to others results in a sudden drought in available credit with catastrophic effects on business. Individuals are often more driven by their own interests, short-term impressive increases in market share and revenue, but under the assumption that this will continue for an indefinite period of time. Unfortunately, this attitude has a cost – forgetting the collective good of the whole group.

5.2. Inefficiency = Inefficient risk scoring techniques

> *" … - This risk scoring model does not seem very accurate - … - Don't worry, just apply the thumb rule - …"*

There is a high number of companies that use outdated, static or inflexible assumptions to elaborate their risk scoring models. It is key that such models are funded on the bases of a balanced combination of quantitative and qualitative data. On the one hand, quantitative data can be considered "*the facts*", but complementarily, the use of qualitative data can contribute to a more flexible and dynamic risk assessment processes.

Apart from that it is also a poor practice over-reliance on credit agency risk scoring. The right approach should be for companies to look for consistent and independent risk measurement methods, validating and challenging constantly their obtained results against those provided by others, for instance credit agencies.

5.3. Faulty management = Illiquidity derived from faulty management of assets and passive

> *" … – ups, it seems the firm is losing liquidity – … well but competitors are going for it and as long as ASSETS = PASSIVE … it's fine …"*

Wheras risk should be aligned to capital or liquidity positions, one of the mistakes incurred by companies is the failure to access cross-information. Such lack of information flowing fluently preempts that risk is managed by incentivizing business lines and consequently, it cannot be ensured that internal pricing reflects the current bear risk.

5.4. Disorders … in capital levels = Inadequate capital levels

" … if the capital level gets too low we'll see later on how to inject more from another source …"

Within an increasingly regulated market, Banks and financial institutions could be forced to build-up capital and liquidity buffers in the good times to secure their survival in a Financial Crisis. Many companies which have gone bankrupt in the last years, incurred this mistake and failed to have such buffers filled at the required levels. Therefore this is a clear area where former industry players should have learned the lesson.

5.5. Uncovered counterpart = uncovered counterpart risk

" … have you considered your counterpart risk? … - Do I really need it?"

Regardless the risk management function should be independent of day-to-day operations, risk components are not something that can be left out of the daily operations – risk limits, stop-loss limits, exposure limits and counterparty limits should be key elements covered by any business line and considered when defining targets and objectives.

5.6. Too complex … = Complex valuation of structured products

$$\text{“ }\ldots\textit{Valuation Prod. A } = \left(\left[\sqrt[18]{x} \cdot \oiint y \right] \oplus \sum_{-\infty}^{\infty} \nabla \lambda z_i \right) \ldots\text{”}$$

The role of structured products is mainly aimed at the design of investment products and customers' risk coverage. These products are highly subtle to market variations and therefore, too complex and time consuming methods for valuating these products allow more room for mistakes and less time to react to the volatility of the markets.

5.7. Extreme scenarios = Lack of extreme scenarios

" … - what if …? - forget it; it will never happen .."

In a high competitive and changing environment, such as the financial sector, it's an imperative for Banks and financial institutions to predict and prevent all potential challenges and threatening situations they may face at some point in time. Companies that have not defined severe, plausible and extreme scenarios in their management models have been more vulnerable and negatively affected by the crisis. Those organizations that have been better off, was because in the past the spectrum and range of possible risk scenarios they considered was more complete and consequently in difficult times they have had more margin to react and launch mitigation actions.

Having said this, the key to overcome such mistakes is based upon prudent management of risk, which entitles:

1. An organizational culture funded on prudent risk management
2. Prudent and anti-cyclic policies as concerns to provisions
3. Proactive risk management initiatives covering the whole credit life cycle; putting special emphasis on risk admission and monitoring processes

6. Six quick wins in risk management

After analyzing the root causes of the economic downturn and the main mistakes, it is time to provide some tips that have proven to be successful for a leading European Bank that, despite the challenging environment, has managed to excel in performance and obtain the 2011 Best Bank Award.

For this player, the quality in risk management is one if its core identity signs and thus, a top 1 line of action. For over 150 years, The Bank has combined prudency in risk management and advanced risk management techniques which have fostered recurrent and healthy turnover and value creation for its stakeholders.

The balance between The Bank's appetite for risk vs. risk control is driven by its business model and core principles:

- The Retail Business is the bulk of The Bank's activity. The Bank is present worldwide and has significant market share (>10%) in all the main markets in which it operates. The Global Business practice is mainly developed in The Bank's core markets.
- Based upon strong capital and liquidity, The Bank's business model enables recurrent profits and financial results.
- The Bank formalizes its presence through autonomous subsidiaries, both in terms of capital and liquidity, compatible with a common corporate control model. The shareholders' structure looks to be simple, minimizing the number of non-operative or instrumental subsidiaries.
- The Bank relies on a powerful and advanced technology infrastructure and corporate tools – which enable an agile compilation and processing of the information – timely and in a standard format
- The Bank's activity is framed within its social and reputational commitment, fully aligned with its strategic objectives.

All these principles are the main levers of The Bank's risk management guidelines. These can be summarized in six quick wins that for other industry players have turned into best practices (Figure 5):

1. Independent Risk Management Function

2. Heavily involved Executive Board
3. Vanguard tools and systems
4. Integrated risk control and management
5. Sustainable risk quality
6. Objective decisions

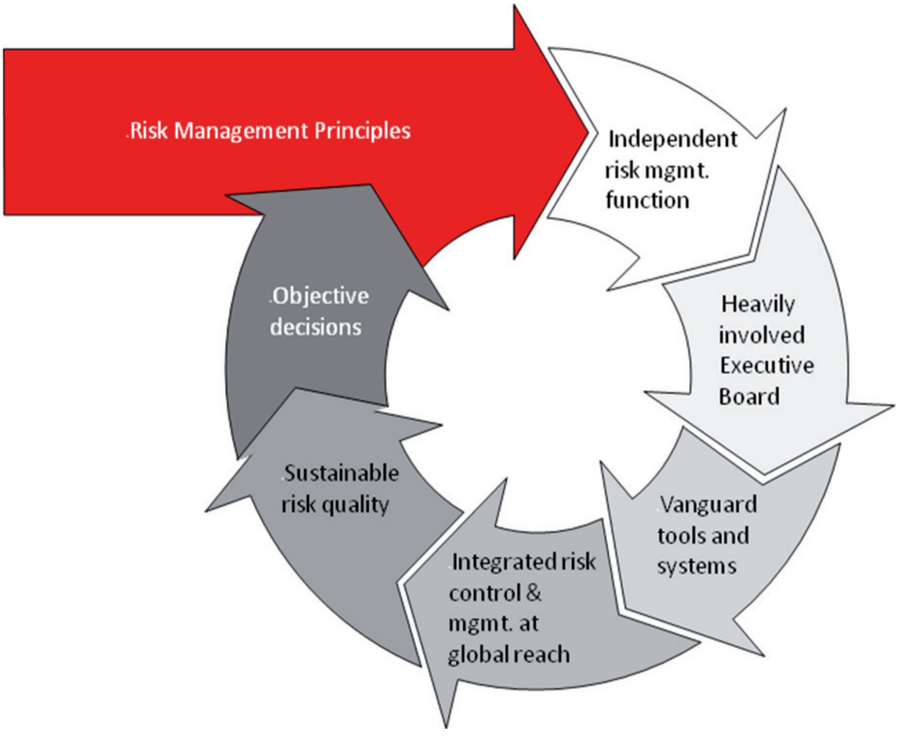


Figure 5. Quickwins in Risk Management

6.1. Independent risk management function

The Risk Management function has to be separated from daily operations. In this sense, The Bank counts on a Risks Management Unit, with its own Head of Department reporting directly to the Executive Board. Such Unit is responsible for risk admission and monitoring processes and acts on an autonomous way and thinking Global.

6.2. Heavily involved executive board

The Senior Management of The Bank is aware of the importance of Risk Management and thus, inspires such culture and principles to the rest of the organization.

6.3. Vanguard tools and systems

One of The Bank's core principles is:

> *"Technology drives Operations"*

Within the Risk Management arena, this translates into The Bank's keenness on the use of vanguard tools and systems. Having the right technology allows The Bank to analyze and measure risks at an acceptable confidence level.

6.4. Integrated risk control and management

Risk must be identified, quantified and homogeneously managed according to a common magnitude (economic capital). The Bank tackles such risk control and management activities in an integrated way across the whole corporate structure:

- every risk
- in each Business line
- at all geographic markets

6.5. Sustainable risk quality

In the last 20 years, The Bank has applied proprietary risk scoring methods to evaluate the risk of its customers and operations. Ensuring that The Bank does not incur risk over a pre-defined threshold, guarantees the sustainable development of its business.

6.6. Objective decisions

Decisions taken must be objective, taking into account contrast of opinions and avoiding decision making being exclusively restricted to an individual. The Bank pursues man-community decisions over credit operations both in the sales and risk areas.

As the objective of this chapter is not to describe in detail each of the aforementioned risk management guidelines, the main takeaway is that it is indeed critical for a firm to have in place some sort of core principles aligning its strategic objectives with its risk management model. Which specific principles are best suited at each particular organization is an open item that requires the organization an internal analysis.

For instance, although a risk management model could prove to be extremely successful for a leading player, it does not imply that the implementation of such guidelines by another competitor would prove the same degree of effectiveness.

In general, following industry best practices, lead to superior results than other techniques; nevertheless, the achievement of enhanced performance will only come by hand of ad-hoc and particularized risk management strategy.

7. Conclusions

This case study has outlined some reasons why Risk Management is important. Though the concept arised over a decade ago, the emergence of the economic crisis in 2008 has given thrust to organizations' adoption of risk management policies, specifically in the financial sector.

To date and during the downturn years we have evidenced common mistakes (the *"seven dreadful mistakes"*), that with the right risk management policies in place could have been avoided or at least ameliorated.

Nevertheless, in the midst of pessimism and underlying financial turmoil there is still room for hype and there are not only lessons to learn, but also quick wins that firms can put into practice to ease their recovery from the current situation.

Author details

Amparo Marin de la Barcena
Group for Automation in Signals and Communications, Technical University of Madrid, Madrid, Spain

Acknowledgement

To my family and peer leaders - for their support and advice at all times.

8. References

[1] Grupo Santander (2010). Gestión del riesgo. pp. 121-157.
[2] El blog del riesgo y la estrategia empresarial (2011) ¿Qué es la Gestión Integral de Riesgo Empresarial?. Available: http://riesgoyestrategia.wordpress.com. Accessed 2011 Dec 11.
[3] LIMRA (2009) The Role of Advisors in Retirement Income Planning. 16 p.
[4] Morgan Stanley (2010) Annuity Survey: Brighter Days Ahead … for Some. 5 p.
[5] Financial Advisor Magazine (2009) Clients Demanding Insured Investments (referred to survey by Financial Research Corp.)
[6] Prudential Financial (2009) Impact of the Market Crisis on Retirement preparedness. 4 p.
[7] Moody MJ (2010) Enterprise Risk Management. The value proposition. The RoughNotes Company Inc. j. 4. Available: http://www.roughnotes.com/rnmagazine/2010/may2010/2010_05p052.htm. Accessed 2012 Apr 15.
[8] KPMG (2011) Placing a value on Enterprise Risk Management. Available: http://www.kpmg.com/CN/en/IssuesAndInsights/ArticlesPublications/documents/Placing-Value-ERM-201003.pdf. Accessed: 2012 Apr 15

[9] JPMorgan Investment Analytics and Consulting (2011) Setting Up A Sound Risk Management Framework

[10] Montero JM, San Segundo R, De Cordoba R, Marin de la Barcena A, Zlotnik A(2009) Strengthening Web Based Learning through Software Quality Analysis. 8th International Conference on Web-Based Learning. Aachen

[11] Mathi K (2004) Key Success Factors for Knowledge Management. Internationales Hochschulinstitut Lindau. Germany.

[12] Marchand D (2004) Information Strategy as a Key Success Factor for Enterprises and Universities. 7th International Bielefeld Conference. Germany

[13] Philippe J Value at Risk: The New Benchmark for Managing Financial Risk. 3rd ed. McGraw-Hill. ISBN 978-0071464956

[14] Marcareñas J (2008) Introducción al VaR. Complutense University. Madrid

[15] AIRMIC (2002) IRM. A Risk Management Standard

[16] Sun J, Li H (2008) Data mining method for listed companies' financial distress prediction, Knowledge-Based Systems, 21(1) (2008) 1-5.

[17] Sun J, Li H (2009) Majority voting combination of multiple case-based reasoning for financial distress prediction. Expert Systems with Applications. j. 36(3): 4363-4373.

[18] Sun J, Li H (2009) Financial distress early warning based on group decision making. Computers & Operations Research. j. 36(3): 885-906.

[19] Sun J, Li H (2009) Predicting business failure using multiple case-based reasoning combined with support vector machine. Expert Systems with Applications. j. 36(6): 10085-10096.

[20] Sun J, Li H (2009) Gaussian case-based reasoning for business failure prediction with empirical data in China. Information Sciences. j. 179(1-2): 89-108.

[21] Sun J, Li H (2010) Business failure prediction using hybrid2 case-based reasoning. Computers & Operations Research. j. 37(1): 137-151.

[22] Aiken M (2001) Competitive Intelligence through Neural Networks. Wiley Interscience. j. 10(1): 49–53.

[23] Estepa JC, Lamothe P (2011) La gestión del riesgo financieroy Basilea III. Universidad Autónoma de Madrid. Madrid

[24] Las prácticas que nos hundieron en la crisis. Available: http://www.quequieres.es. Accessed 2011 December 11.

[25] Survey to Jesús Huerta de Soto (2010). Professor of Economy at Rey Juan Carlos University

[26] BBC news (2008) Recession in the US has arrived

[27] PriceWaterhhouse Coopers (2011)14th Annual Global CEO Survey

Risk Management on the Romanian Capital Market

Marin Andreica,
Stere Farmache,
Madalina Ecaterina Andreica
and Octavian Stroie

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50493

1. Introduction

The Stock Exchange is a barometer of economic and financial life in any country, an indicator of world business and economic prospects. Its movement is reflected in the stock market indices that express the overall performance of the stock market or just a certain sector, and offers the investors the base of their investment decision on the markets. Therefore, the authors have proposed in this paper to summarize the results of research in order to highlight the most significant issues of the risk management on the Romanian capital market and to provide solutions to them.

The first part highlights the role of stock market indexes in describing the state of the capital market, as well as the difficulties arising on their representation in the current financial and economic crisis. In order to reduce the risk of erroneous information to investors, the authors propose a composite index that summarizes the information from other indexes on the capital market in Romania. Finally, using the Risk Grades method, the authors make a comparative analysis of the risks on the Romanian capital market with those in the Czech Republic, Poland and USA capital markets.

The second part outlines the main classes of risks that arise in the settlement of exchange transactions that dominate in the risk hierarchy of the stock operational management. After highlighting some international regulations on reducing the risk of stock trading settlement, the authors develop a broad range of solutions for reducing these risks on the Bucharest Stock Exchange.

2. Risk assessment of stock market indexes

2.1. The role of stock market indexes

In the economic literature, the stock exchange is often characterized as being a true "engine" of the economic life, which stimulates business activity and allows the development of large-scale investment projects. Through its existence and operation, the stock markets provide liquidity for financial assets, enabling the sale of securities previously purchased, thus turning them to money. Economic and financial crisis that has monopolized economies worldwide in recent years has had a negative impact on the development of stock markets. Thus, stock prices of large companies listed on stock exchanges fell dramatically, the liquidity of stock markets fell heavily, and on this basis, the transaction volume decreased due to the reduction of interest shown by investors and speculators. Indexes saw an overall decrease and *a distortion of information provided to potential investors because of "interference" occurred in the structure of the companies participating in their training*. By their developments, stock market indices are able to provide clues on the economic outlook whereas in countries with developed financial markets, leading companies have almost without exception status of listed companies. Thus, the indices themselves may be true barometers of economic development or they may be included in the calculation of aggregate economic indicators, with a high degree of complexity.

Expressing the overall performance of stock markets or simply of a sector of the market, indices allow investors to perform analysis on yields of already performed transactions or on investment opportunities that may arise in the future. The existence of indices make the analysis of various stock markets much easier, even if the different methodologies used for calculating the indices affect to some extent the comparisons accuracy [1-3].

On any stock market, the indices are a reference for each investor when assessing the quality of the management portfolio. Due to the fact that in many cases individual performance fail to exceed indices' long-term growth rate, investors and fund managers especially were tempted to build their own portfolios using the precise structure and composition of the indices.

2.2. Critical elements in building and managing indexes

A first issue to be solved when considering the desirability of launching a stock index is if you choose to use a formula in the category of arithmetic or geometric average. Since the use of geometric replication creates difficulties when the stock index would be eligible as underlying for derivatives or as a benchmark for investment fund, managers will prefer to use a formula based on the arithmetic average.

Also, the option to use weights in the calculation of indices involves an analysis of market characteristics to be described by the respective stock market index. The decision to weight the prices which is part of an index is itself a controversial:

- the supporters of the unweighted variant believe that such an index only ensures equal treatment of each stock symbol which is part of an index;

- those who appreciate more weighted index say that the only way to avoid unwanted situation in which the relative change in price of shares issued by a company small in size will influence the stock index as much value as a change in the same type of the share price "Blue Chips".

Another issue that arises is the choice of the weighting factor. In theory there may be various options for the weighting factors, but in practice have won two:

- weighting by market capitalization (or the amount of the outstanding shares);
- weighting the liquidity of the shares that make up a stock index.

In the first case there is a risk that a relatively small number of companies - which have a market capitalization way above the market average - will decide the meaning and scope of developments in practice the whole index. It is a common situation in emerging markets. Therefore, it is often set a maximum level of the share capitalization of a single symbol that can play in the total capitalization of the symbols in the composition of a stock index.

In practice, the liquidity is the most common alternative criterion to using market capitalization as the weighting factor. Attractiveness of the weighted indices to the liquidity lies in the idea that virtually any listed price of a share is a fiction as long as it materializes through a stock transaction. Therefore, prices become relevant only insofar as they are the result of a large number of transactions, which require (at least at first) the use of liquidity as a weighting factor in calculating the indices. The disadvantage of using liquidity as a weighting factor consists, however, in the presence of many differences from one day to another in the number of shares made with a specific symbol, which significantly alter its influence in the evolution of the overall index [4-5].

In principle, a stock index is a tool for measuring aggregate price developments for the symbols which compose it, the sole result of supply-demand ratio of the market and the continuing process of reassessment of the price of each symbol individually [6-8]. Therefore is needed to incorporate the adjustment coefficients in the formula for calculating an index, allowing correction of artificial effects induced by certain events on the trading price of a symbol:

- events which directly affect the trading price of a symbol: the provision of dividends, granting of rights to subscribe.
- events that influence the number of shares taken into account: the change in share capital by increasing / reducing the number of the outstanding shares, split face value, strengthening the nominal value, granting free shares (bonus issues).
- changes in the composition of a stock index: the inclusion of a new symbol, a symbol of exclusion.

Payment of the dividends or issuing of preference rights involves reducing the value of traded shares of a company, unless the subscription rights have value “zero” or negative. When providing payment of dividends and share trading, price decrease is due to reducing the company's value by the amount allocated for dividends, while the rights issue of preference allocation occurs due to the decreased value of the company on a number of shares. These decreases in market value of a share are fully balanced by increasing the same amount of portfolio investors as a result of amounts received as dividends or the sale of preference rights indices as a result of the above events has the advantage of simplicity, though in practice most indices are adjusted when there are such price changes.

- Assuming that the granting of a certain amount as dividend per share produces a decrease in the share price with the exact amount of the dividend, the adjustment procedure will require adding value to the price of the stock dividend is taken when calculating the indices. Otherwise, you can use an adjustment factor, multiplied by the share price is taken when calculating the index.

In the construction and management of indices there are other critical elements, which primarily aim to adapt to the concrete conditions of the stock market that they are going to describe. One such factor is the choice of a *data base*, which is preferably selected so that it does not belong to a period marked by turbulence or excessive volatility. As a general rule, the base date of an index must be chosen *"in order not to overlap with a minimum or maximum market history"*.

The base value of an index is also an important element in building a stock index. It should be chosen so that the reporting of the current nominal value of an index value based may allow easy calculation of percentage change in that timeframe. Therefore, the most commonly used figures based on an index value are 100 and 1,000 points.

In the case of adjusted indices, selecting a large number of shares to be part of an index basket involves assuming the possibility of relatively frequent cases of adjustment, thus reducing the transparency index. Also, a too extended composition raises serious problems for managers who use index funds as a "benchmark" or which suggest as investment strategy to make investments in securities index.

If an index is used as underlying assets for the development of financial derivatives, its composition becomes more important. It is desirable that the selected symbols, especially those with high weight in the index, have a sufficiently high level of liquidity in order not to influence the price with the idea of obtaining a more favourable settlement value of futures contracts or options. The issue of settlement value of derivatives that are underlying indexes has been addressed and in terms of prices that are taken in calculating its value at the close of trading session. Intuitively, the last price recorded in a day by a certain symbol should generate and last values of stock market index. To avoid forcing the closure of certain levels of price indexes, administrators have considered closing the option that its value is determined either as the average of past values of the index calculated based on average prices in a given period time for each symbol in its composition. This concern for the accuracy of the closing price of the shares and hence the value of stock index led some stocks to decide to introduce at the end of the trading session of the tender period in which prices are determined on a multilateral basis, through a fixing algorithm.

2.3. Improvement of the management of Bucharest Stock Exchange indexes

Through the changes that were made in time upon the methodology of calculation and management of indices, the Bucharest Stock Exchange (BSE) showed that it understands their importance for the investors and the general public. However, according to the authors, there are still some aspects in the functioning of the indices in the future, as changes might occur, so that the indices calculated by the BSE to increase their representativeness and to become more attractive tools for use as benchmarks in management activity portfolio.

The first change is imposed by the profile described by the historical values of the indices. As it may be seen on Figure 1, joining the five indices calculated by BSE in the same graphic deprives us of any opportunity to make a comparison. This is obviously due to objective factors, as as BET, BET-C, BET-FI, BETXT and BETNG were launched from different basic data. Moreover, the BET and BET-C had worked in a period when inflation was very high and the national currency exchange rate at the time of their release was lower than the current one.

Therefore, it is required the denomination of all the BSE indices, so that on a specific reference date to have the same value for all indices. Bringing all indices at the same amount for a certain date, would open the way for a much easier tracking trends and comparison of their indices and market sectors they represent.

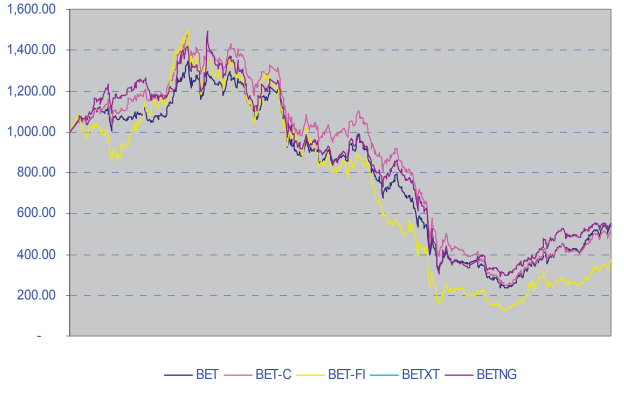


Figure 1. BSE indices (denominated) in January 2007-8 September 2009

To achieve this denomination, an important element consists in determining *the reference date*, which should be the first day of 2007, the day that marked the accession of Romania to the European Union, and thus integration of the regulated market operated by the Bucharest Stock Exchange in the unique European financial market.

Once the indexes are denominated, it appears the following problem consisting in identifying their new *base value*. Since 2 January 2007 had already established that the reference date for two of the indices BSE - BETNG and BETXT – which were then calculated for 1000 points, a value which seems to be the most appropriate and which will greatly ease the adjustment values of Historical BET index, BET-C and BET-FI. Adjusted historical series for these three indices of BSE will be as one that can be quite easily associated with the original series.

Another problem to be solved is to develop a synthetic stock index describing the overall market share in Romania, as the Bucharest Stock Exchange also calculates a number of indices for Rasdaq: RAQ-I, RAQ-II and RASDAQ-C (which is a composite index).

One of the main concerns of a stock market operator is to identify the most effective means by which information about the activities of the regulated market that he manages, reaches the participants, from potential investors or the general public. Indexes are highly effective tools in promoting or communication, and that is why special attention should always be given to their management. Regardless of the methodology used, the scope or the "brand" adopted, the indexes must reflect as closely as the stock market developments in a given period of time, and the investors and the general public should recognize this status. Therefore, we believe that the stock market in Romania would be even discussed whether to outsource the work associated with the major indexes by a specialized international company to calculate and disseminate indexes. Market indexes license is highly specialized, with a relatively small number of participants. Therefore, "rebranding" of BSE indices under the name of one of the leading players in this market (Dow Jones, FTSE or MSCI) it would significantly increase the international exposure of local indexes.

2.4. Construction of a synthetic index for the stock market in Romania

Currently, the shares of the Romanian companies are traded in two distinct market sectors (BSE - regulated market and Rasdaq), with the chance to add another market (alternative trading system). For historical, but also technical reasons, the two market sectors are currently covered by two families of indices:

- **“BET index family”,** for the regulated market;
- **“Rasdaq index family”**, for the RASDAQ section.

Therefore, as stated, there is not a single synthetic indicator to describe the evolution of prices in the two market segments. Furthermore, in the "BET index family" there is a composite index, BET-C that does not include in its membership the five SIF symbols. So, even the most comprehensive index of the regulated market does not include the five most liquid stock market symbols, which normally accounts for almost half of the total transactions on the BSE [9].

That is why, it is necessary to create a more comprehensive stock index, which describes the overall evolution of the quotations for all shares traded on markets and market sections administered by the BSE. A possible name for this could be ALL-X, and that structure will be an "index of indices".

Once the objective is defined and the name is established, the next step is to clarify the composition of such an index, so there is no overlap, but at the same time to remain as few "areas" uncovered as well. From the analysis of the five indices of the “BET family” and all three of the "RASDAQ family” it is very clear that the optimal variant is that the composition of ALL-X to enter the next three BSE indices:

BET-C: reflects the price development of all companies listed on regulated market, Category I and Category II, except for the SIFs. BET-C is a price index weighted by market capitalization of companies in its composition. The maximum weight of a symbol in the BET-C composition is 20%

BET-FI: reflects the overall trend in the shares prices issued by the five financial investment companies (SIF) on the regulated market. The maximum weight of a symbol in the composition of the BET-FI is 25%

RASDAQ-C: includes all the shares traded on the "RASDAQ market" and follows the synthesized global trend in prices.

Regarding the topic, the fact that the Rasdaq-C index is calculated at the end of the trading session (and not in real time) also requires that the ALL-X index to be calculated with the same frequency.

Also, the type of formula chosen is relatively simple, since all three indices (I_i) use the arithmetic average. Therefore, the ALL-X index will use the same formula - weighted arithmetic average, as the relative importance of the three indices in the general stock market in Romania is very different, both in terms of market capitalization and liquidity of the shares in their composition. Therefore, it is necessary to use a weighted arithmetic average.

For the price index, the most frequent weighting factors (Pi,t) are either capitalization or liquidity. Since all of the three indexes found in the composition of ALL-X are capitalization-weighted price index, at least apparently the natural alternative is to use the same weighting criteria: capitalization. Moreover, the choice of the capitalization weighting factor ensures that the advantage of not requiring frequent rebalancing for an investor who decides to adopt a passive investment strategy in relation to the ALL-X index.

Through its coverage, the ALL-X index will create replication difficulties. Therefore, in its construction, the focus should be on its descriptive issues and less on the investment. Consequently, it becomes important to find a variant that takes into account the weighting of liquidity, which has the disadvantage that the index can cause daily changes without evidence of membership to undergo changes, or changes in scope than the other induced increases or decreases in the BET-C index, BET-FI and Rasdaq-C. Such a compromise version could be given to the establishment of fixed predetermined weighting factors, which remain unchanged for a defined period of two consecutive data adjustment index ALL-X (one quarter).

In calculating these predetermined weighting factors, both the capitalization and liquidity will be taken into account. The formula for calculating the determining factors will be:

$$Pi,t = \frac{Ui + Vi}{2} \tag{1}$$

where:

Ui is the weight that the capitalization of the symbols within I_i index holds in the total capitalization of the market, at the time of reference;

Vi is the weight that the value of transactions with the symbols of the I_i index holds in the total market turnover in the reference period (quarterly).

To ensure a unified form to display all BSE indices, will opt for the next version of the equivalent representation:

$$(\text{ALL-X})_T = (\text{ALL-X})_{T\text{-}1} x \frac{\sum_{i=1,N} p_{i,T} \times I_{i,T}}{\sum_{i=1,N} p_{i,T} \times I_{i,T-1}} \quad (2)$$

where:

Pt,i: weighting factor at time T, associated with each index in the composition of ALL-X;
Ii,t: ALL-X indices, at time T (BET-C, BET-FI, Rasdaq-C).

To ensure the possibility to compare the new ALL-X index with the main stock market index, BET, the start value will be 2901.10 points and the start date will be fixed for 31.12.2008. The calculation of the weighting factor, Pi,t, will be done quarterly, on the last trading day of the final month of the quarter, and will enter into force on the first trading session of the next quarter [10].

The main advantage of building the index ALL-X is that it can be obtained a simple and effective tool to provide an overview of the entire stock market development in Romania. The index is limited to the evolution of shares issued by the Romanian companies, while any direct influence induced by the prices movement of foreign securities admitted to trading on domestic capital market was being removed from the beginning.

In the presented form, ALL-X index is defined primarily as a descriptive index and less like an investment. This deficiency can be corrected later, to the extent that the BSE will decide to use the same calculation methodology for all its indices, regardless of the section or market segment that is addressed: regulated market, Rasdaq market or alternative trading system. Moreover, in order to emphasize the character of investment, it will be necessary to modify the weighting factor (which may become the capitalization of each index), as well as the dates when its new value is calculated and effective.

To calculate the ALL-X index values retrospectively, after the release date was determined (December 31, 2008) and the reference value at that time (2901.10 points), the first step is to determine the Ui and Vi factors for the first quarterly interval (January-March 2009). To this end it will determine:

- capitalization associated to the symbols in the composition of the BET-C index, BET-FI and Rasdaq-C on the last trading sessions of December 2008, March 2009, June 2009, September 2009 December 2009, March 2010, June 2010, September 2010 and December 2010 (RON);

Date	BET-C	BET-FI	Rasdaq-C
23-Dec-08	25,441,147,900	1,715,098,084	12,099,865,129
31-Mar-09	20,494,699,824	1,473,851,929	11,799,297,770
30-Jun-09	28,577,557,323	2,276,366,573	12,359,818,473
30-Sep-09	35,329,737,297	3,454,691,266	12,868,187,477
24-Dec-09	35,313,895,995	3,264,425,651	12,346,442,408
31-Mar-10	46,106,028,659	4,391,516,633	14,104,548,987
30-Jun-10	36,894,347,800	2,715,614,092	11,065,581,044
30-Sep-10	41,488,093,266	3,282,932,587	11,556,865,887
30-Dec-10	41,901,691,868	3,004,035,475	10,832,645,290

Table 1. Capitalization associated to the symbols in the composition of the BSE indices

- liquidity (total amount in RON) associated to the symbols in the composition of the BET-C index, BET-FI and Rasdaq-C in Quarter IV 2008, Quarter I 2009, Quarter II 2009, Quarter III 2009, Quarter IV 2009, Quarter I 2010, Quarter II 2010, Quarter III 2010 and Quarter IV 2010 (RON).

Period	BET-C	BET-FI	Rasdaq-C
Oct - Dec 2009	435,647,850.34	468,296,885.00	186,402,641.20
Ian - Mar 2009	423,175,985.74	203,780,032.70	86,885,606.44
Apr - Jun 2009	668,823,417.53	658,946,311.40	102,202,564.63
Jul - Sep 2009	600,539,594.05	728,518,117.60	94,069,608.81
Oct - Dec 2009	485,044,466.85	691,750,042.50	113,518,426.21
Jan - Mar 2010	691,952,816.70	735,425,007.50	148,602,226.21
Apr - Jun 2010	695,515,698.95	1,039,348,556.50	130,801,236.36
Jul - Sep 2010	331,670,219.85	407,212,194.50	57,407,872.87
Oct - Dec 2010	352,178,129.60	304,259,023.50	87,486,365.49

Table 2. Liquidity associated to the symbols in the composition of the BSE indices

Based on the data presented in the previous tables, we will calculate the following elements:

- weights associated with each index in the total capitalization for each final quarter;

Period	BET-C	BET-FI	Rasdaq-C
Jan - Mar 2009	64.81	hicag30.82	30.82
Apr - Jun 2009	60.69	4.36	34.94
Jul - Sep 2009	66.13	5.27	28.60
Oct - Dec 2009	68.40	6.69	24.91
Jan - Mar 2010	69.35	6.41	24.24
Apr - Jun 2010	71.37	6.80	21.83
Jul - Sep 2010	72.81	5.36	21.84
Oct - Dec 2010	73.65	5.83	20.52

Table 3. Weights associated to the BSE indices in total capitalization (%)

- weights associated with each index in the total liquidity for each quarter;

Period	BET-C	BET-FI	Rasdaq-C
Ian - Mar 2009	39.95	42.95	17.10
Apr - Jun 2009	59.28	28.55	12.17
Jul - Sep 2009	46.77	46.08	7.15
Oct - Dec 2009	42.20	51.19	6.61
Jan - Mar 2010	37.59	53.61	8.80
Apr - Jun 2010	37.28	55.71	7.01
Jul - Sep 2010	41.65	51.14	7.21
Oct - Dec 2010	47.34	40.90	11.76

Table 4. Weights associated to the BSE indices in total liquidity (%)

Once the weights are determined, it will be necessary to calculate the weighting factors in the composition of each ALL-X index for each quarter:

Period	BET-C			BET-FI			Rasdaq-C		
	Ui	Vi	**Pi**	Ui	Vi	**Pi**	Ui	Vi	**Pi**
Ian - Mar 2009	0.6481	0.3995	**0.5238**	0.0437	0.4295	**0.2366**	0.3082	0.1710	**0.2396**
Apr - Jun 2009	0.6069	0.5928	**0.5999**	0.0436	0.2855	**0.1646**	0.3494	0.1217	**0.2356**
Jul - Sep 2009	0.6613	0.4677	**0.5645**	0.0527	0.4608	**0.2567**	0.2860	0.0715	**0.1787**
Oct - Dec 2009	0.6840	0.4220	**0.5530**	0.0669	0.5119	**0.2894**	0.2491	0.0661	**0.1576**
Ian - Mar 2010	0.6935	0.3759	**0.5347**	0.0641	0.5361	**0.3001**	0.2424	0.0880	**0.1652**
Apr - Jun 2010	0.7137	0.4391	**0.5764**	0.0680	0.4666	**0.2673**	0.2183	0.0943	**0.1563**
Jul - Sep 2010	0.7281	0.3728	**0.5504**	0.0536	0.5571	**0.3053**	0.2184	0.0701	**0.1442**
Oct - Dec 2010	0.7365	0.4165	**0.5765**	0.0583	0.5114	**0.2848**	0.2052	0.0721	**0.1386**

Table 5. Weighting factors Pi,t

Figure 2. Evolution of the ALL-X index vs. Evolution of BET index (Jan. 2009 – Dec. 2010)

The biggest weight in the ALL-X index is given by the BET-C composite index. The values of Pi,t associated with the BET-C for the full period 2009-2010 ranged between 0.5 and 0.6 percent.

2.5. Risk assessment indices

Stock markets investors constantly face several risks, which is why researchers have focused their concerns on issues related to developing risk management improved models of assessment and management. Among the latest methods used in economic practice in assessing risk in financial markets the method *Value at Risc (VaR)* is required [11].

Risk Grades indicator provides a synthetic image of the risk associated with a certain financial security, a portfolio or an index. This new way of measuring volatility is based on exactly the same data and analysis as VaR, while making it possible to translate the Risk Grades estimates in terms of VaR. Furthermore, Risk Grades methodology is intended to be more intuitive and easier to use than VaR. The characteristics of the Risk Grades are:

- It is a dynamic method of measurement that adjusts to the market conditions
- Allows comparison between different investment classes, regions and different areas, is a standardized method of measuring volatility
- Allows comparison of different market indices.

The Risk Grades formula for the financial asset *i* is:

$$RiskGrades_{(i)} = \frac{\sigma_i}{\sigma_{base}} x100 \qquad (3)$$

where: σ_i the volatility of the financial asset *"i"*
σ_{base} the base volatility

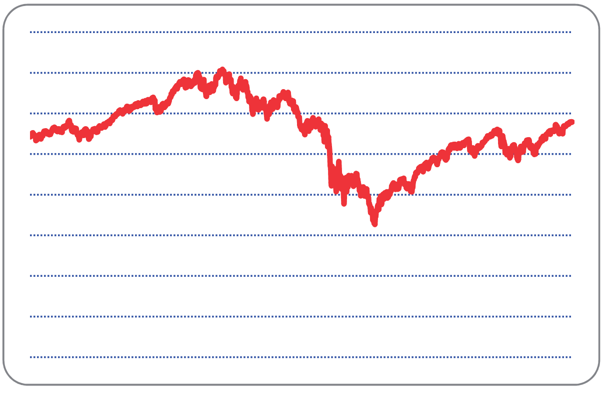


Figure 3. Evolution of DJIA between 1 Jan 2006 – 30 Dec. 2010

To determine the Risk Grades values associated to the main indices calculated by the BSE (BET, BET-FI, BET-C) it will be considered a decay factor λ = 0.98, which corresponds to a minimum number of records n = ln (0.1)/ln (λ) of 113. The series of values for indexes is 1 January 2008 to 30 December 2010 and the index used as reference for determining the Dow Jones Industrial Average (DJIA) is σ_{base}.

We can notice in the above chart the negative effects of the financial crisis that occurred on the DJIA mainly by early 2009, when the North American stock market indexes registered multi-minimum (on 9 March 2009, the DJIA was calculated for 6547 points). Choosing the interval 1 January 2009 to 30 Dec.2010 for the series of values Risk Grades indices used in the calculation took into account to cover both a pronounced downward trend of the DJIA index and an increasing trend after March 9 2009.

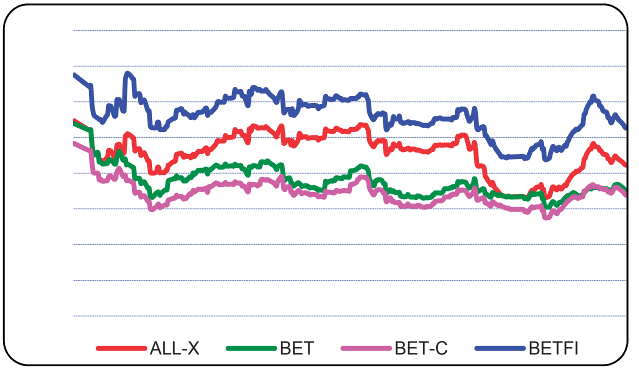


Figure 4. Risk Grades values for BSE indexes between 21 May 2009 - 30 Dec.2010

The Risk Grades values determined for the BSE indexes confirms that the Romanian capital market is riskier than the North American capital market. Besides that, as one can notice from the above chart, the Risk Grades values of the ALL-X index are consistently positioned below the BET-FI sector index and above the associated BET and BET-C.

If decay factor λ = 0,96 was chosen, the minimum number of records that we need is 228, sufficient for our purpose, consisting in comparing the risk associated with the capital market in Romania with those determined for other international stock markets or within the region.

In order to make a comparative analysis of the risks on the Romanian capital market and the main indexes of the Central and Eastern Europe we took into consideration the main indexes of Poland and the Czech Republic capital markets, meaning the WIG20 and the PX.

The **WIG20** index is the underlying instrument for futures transactions listed on the Warsaw Stock Exchange and represents a modified capitalization-weighted index of 20 Polish stocks which are listed on the main market. The base value was set to 1000 points in April 16th 1994. The WIG20 index may not include more than 5 companies from a single exchange sector and limits the maximum weight of one symbol included in the WIG20 to 15%. Currently, the most important companies included in the WIG20 index are PKO BP, KGHM, DAM and Bank Pekao.

The **PX Index** is an index of major stocks that trade on the Prague Stock Exchange. The 5th of April 1994 was selected as the starting exchange day (a benchmark date) for the Index PX 50 and its opening value was fixed at 1,000 points. At this time the index included 50 companies traded on the Prague Stock Exchange, accordingly named PX 50. In 20 March 2006 the PX 50 index was merged with PX-D into the PX index, having now 14 symbols of the Prague stock market in its composition. The maximum weight for a stock symbol included in PX was set to 20% and from the companies included in the PX index the most important are CEZ, Comercni Banka, Erste Group Bank and Telefonika.

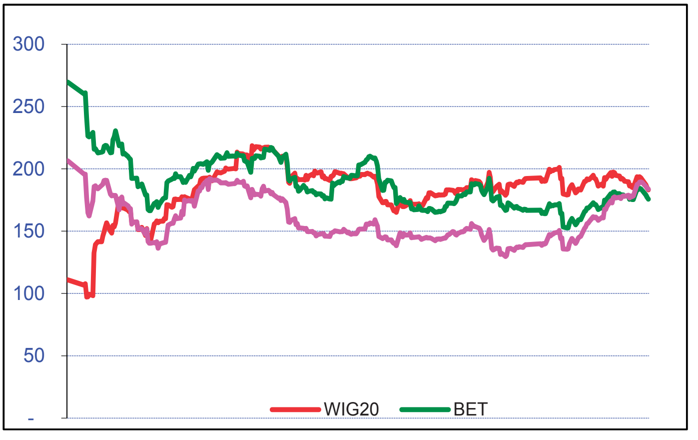


Figure 5. Risk Grades values for the main indexes in the region between 21 May 2009 - 30 Dec.2010

When comparing the Risk Grades values of WIG20 index of the Warsaw Stock Exchange and of PX index of the Prague Stock Exchange with those of the Romanian Stock Exchange (see figure 5), we reach a surprisingly conclusion, that the Risk Grades values calculated for the main indexes in the Central and Eastern Europe do not show major differences. Although constantly the Risk Grades values are above those associated with the WIG20 and PX index, the difference is very small [12].

3. Risk management in settlement of stock transactions

3.1. Clearing and settlement of stock market transactions

The risks managed by Institutions such as market operators, clearing houses and central depositories are not less important, but are maybe less visible to the public.

It is known that stock market transactions are followed by a mutual transfer of ownership from seller to buyer, and the amount of money equivalent, from buyer to seller. Clearing and settlement take into account that transfers of financial instruments and financial funds between parties are to be performed safely and with little cost.

To prevent and to reduce the risks specific to the settlement of financial instruments transactions, direct access to clearing and settlement system is limited to those participants who meet a set of capitalization criteria, operational capacity and professional experience, which means that the settlement of an exchange transaction involves a given number of financial intermediaries. The more financial intermediaries are involved in the settlement of financial instruments transactions, the more complex is the settlement process.

3.1.1. Clearing and settlement stages

Clearing and settlement process begins immediately after a transaction with financial instruments is closed and goes through the following main steps:

- **Confirming** the terms in which the deal was closed between the buyer and seller. Confirmation of transactions can be conducted directly between parties or using a system that generates such an acknowledgment. In the latter case, the terms of the transaction can be transmitted to the central system or automatically by the trading system, or separately by each party. Although apparently it is just a formality, confirmation of transactions is a significant milestone in the process of completing a stock transaction, since it can identify in due time differences that may affect the proper functioning of the system.
- **Compensation** is an algorithm of obligations and / or rights accrued to the participant, after transactions were inserted in the system and confirmed. Usually, the calculation of obligations and rights of a participant can be made on a gross basis for the settlement of transactions, based on a bilateral net or multilateral net basis. Therefore, the calculation algorithm can result in one or more instructions to participants in the system that have to be respected. Only in the case of net multilateral compensation, the calculation algorithm results for each participant in only one instruction, that expresses (cash settlement level) either the obligation/ write to pay/ receive a certain amount.
- The settlement itself, which is composed of:
  - **obligation to deliver securities:** falls to the seller that is part of a transaction or to the participant in the settlement that is in the position of a net debtor for a particular security;
  - **obligation to pay an equivalent amount:** falls to the buyer that is part of transaction or to the participant in the settlement with a debtor net quality in respect to financial funds.

The last stages of the settlement imply the highest financial risk throughout the process of finalizing exchange transactions, one of the main criteria that measures the performance of a clearing system and securities settlement transactions refers to the method the connection between the mechanism of transfer of securities (obligation to deliver) and transfer mechanism of funds (payment required) is made. The systems within which the permanent transfer of securities are to take place, subject to an exchange transaction, only if the final and irrevocable payment of amounts is made, fulfilling one of the main conditions to measure performance:

delivery versus payment (Delivery Versus Payment - DVP). By using mechanisms that ensure DVP, one eliminates the most important component of all financial risks that accompany the settlement of transactions with financial instruments**: principal risk.**

3.2. Risks in clearing and settlement systems

As presented before, between the moment a transaction with financial instruments is closed and the ownership transfer certain stages are to be completed in a given time frame. The following risks are present and occur:

CREDIT RISK, implies the possibility that one side of a transaction might not honor its obligations, or after maturity, in whole or in part. Types of credit risk:

- **Principal risk**: is a consequence of time the difference between payment and delivery of securities as part of an exchange transaction. In case bankruptcy occurs between the two moments or insolvency of one parties involved in the settlement of that transaction, it is possible for one party to incur a loss equal to the transaction left unsettled. If such a risk occurs, the buyer might find itself in the position of having the financial securities acquired through market transactions paid, while the counterparty is no longer able to deliver those securities. The risk of the principal, from the perspective of the seller is described by the situation in which financial instruments have been delivered, but the buyer no longer has the opportunity to order the payment to the seller.
- **Replacement Cost Risk:** refers to a situation where before a final settlement between the parties involved in securities transaction may not be able to complete the initial negotiated market transaction settlement, thus requiring a replacement of that transaction. Thus, parties' involved in a transaction are equally exposed to loss due to the price change of the security that was object of the original transaction, the time interval between the time the original transaction (but not settled) and the transaction time for the replacing one. The buyer involved in a transaction takes over the risk associated with the replacement cost if the initial price of the transaction performed with a particular security is below the market price of that financial instrument once the transaction shall take place. Looking at the situation from the seller's point of view, the same type of risk occurs when the transaction replacement loss is concluded at a price lower than that of the original transaction. The size of the losses due to the risk associated with the replacement of a stock transaction is directly proportional to the volatility of the securities that were part of the original transaction and with the time between the date of completion of the transaction and the time of the final settlement. The shorter this time frame is, the lower the probability for the risk to materialize.

LIQUIDITY RISK refers to the situation where the seller of certain financial securities does not receive on the settlement date the amount from the buyer and thus he finds himself in a position to borrow or sell other financial assets to honor at its turn various other obligations assumed in view of collecting amounts from the original sale of securities. Liquidity risk arises in the case of the buyer as well if upon settlement it does not receive the securities purchased that were possibly involved in a transaction following the initial sale. The size of the liquidity risk is inversely proportional to market liquidity: the more liquid the market is, the lower the liquidity risk costs are. Liquidity problems have a high potential to have negative influences on the entire settlement system, especially when occurring amid high volatility of shares market and any delay in completion of the settlement by a participant thus raising concerns about solvency.

SYSTEMIC RISK refers to cascade spreading of a situation when settlement is impossible, from one participant to the clearing - settlement system to other members in the system (or all participants in the system). Institutions that operate clearing and settlement of exchange transactions are required to provide mechanisms and procedures to prevent expansion of liquidity or solvency problems from one participant to the whole system. Investor confidence in post-transaction systems is essential for liquidity growth of the stock market and is a prerequisite to entering a virtuous circle: a more liquid market significantly reduces the impact of risks that may arise in clearing and settlement system, which is reflected in increased market liquidity.

OPERATIONAL RISK implies that defective processing of transactions by participants in a settlement system, failure of transactions, fraud, disruption of communications or any other operating problems turn into financial losses.

Depending on the structure of clearing - settlement systems and transaction type, there might also be other risk categories, such as bankruptcy risks of the settlement bank, legal risk or custody risk. The probability that one or other of the above risks could materialize depends substantially on the following two main factors:

- implementation of strict procedures to control risks in the clearing and settlement of securities transactions, procedures known collectively as "risk management";
- operation plan of the clearing and settlement system of transactions with financial instruments.

3.3. Recommendations on the clearing and settlement systems to reduce risks

Materialization of risks in clearing and settlement operations may have significant negative influences on a country's entire financial system [13].

Moreover, given the internationalization of financial flows, where some of the affected participants in a clearing and settlement system operate simultaneously in multiple markets or operational links exist between two or more national clearing and settlement systems, this may cause contamination and spread of local problems internationally. Thus a set of recommendations was elaborated by specialized international institutions.

3.3.1. G30 recommendations

“G30 Recommendations” consists of a set of nine recommendations that should be implemented by all settlement systems:

- Confirmation of all transactions must be achieved by all direct participants in the market (for example: brokers/ brokers/dealers and other members) in T+0. Details of all confirmed transactions must be transmitted to clearing – settlement system.
- Indirect market participants (e.g. institutional investors) should be informed on the implementation details of transactions in T +1.
- Every country should have an operational central securities depository, organized and managed so as to encourage activity of the direct and indirect market participants. The range of financial instruments eligible to be registered in the central depository must be broad. Preserving or dematerializing financial instruments should be as extensive as possible. If on the same market there are several depositors, they should operate on the basis of compatible rules and practices, in order to reduce settlement risk and to allow use of funds and make available any financial securities pledged as collateral in an efficient manner.
- Each market should be encouraged to reduce settlement risk by introducing "gross settlement in real time" mechanisms (Real Time Gross Settlement - RTGS) or by using compensation based on transactions that meet Lamfalussy recommendations.
- DVP system should be used as a method of settlement for all transactions. DVP is defined as "simultaneous dismissal, final, irrevocable and immediate securities and funds continuously throughout the day."
- Payments related to settlement of securities transactions and using a financial instrument portfolio should target all markets and financial instruments, the adoption of the "Same day fund" Convention.
- A single cycle of implementation of the settlement should be adopted for all markets. Final settlement for all transactions should be made no later than T +3.
- Loans of financial instruments should be encouraged as a method of developing financial instruments transactions settlement. Barriers and taxes that limit loan of financial instruments should be removed.

Each country should develop a standard for messages related to financial instruments based on ISO Standard 7775. In particular, each country should adopt the ISIN coding system for securities, as specified in ISO Standard 6166.

3.3.2. CPSS – IOSCO recommendations

CPSS (Committee on Payment and Settlement Systems) - IOSCO (International Organization of Securities Commissions) recommendations, completed in 2001, take and develop provisions of G30 recommendations, adapting them to the important changes occurring in the securities industry in recent years, but adding new ones.

1. The legal system: the settlement of transactions with financial instruments must have a solid legal framework, clear and transparent.
2. Confirming transactions: Confirming transactions between direct market participants should be completed as soon as possible after the time transactions are executed, but not later than the end of the trading day (T +0). When it is necessary to confirm the

transaction by indirect market participants (such as for institutional investors or global custodians), it must take place as soon as possible after execution of the transaction, preferably "T 0", but not later the "T +1".

3. The settlement cycle: settlement transactions mechanisms must adopt the system by which the value date is set after a specified number of days from trade date ("rolling settlement system"). Final settlement must be made no later than "T +3". Costs and benefits of a shorter settlement cycle than "T +3" should be carefully evaluated.
4. Central counterparty (CCP): costs and benefits of central counterparty mechanism should be evaluated. The introduction of this mechanism should always be accompanied by control of risk assumed by the institution which provides central counterparty services.
5. Loan instruments: lending of financial instruments (REPO contracts or other equivalent transactions) should be encouraged to streamline the process of settlement of transactions. Barriers that inhibit the use of loan activities for this purpose should be eliminated.
6. Central Depository (CSD): Preserving or dematerializing of securities and their transfer through electronic registration account must be used as widely as possible.
7. Delivery versus payment (DVP): CSDs should eliminate principal risk by linking the transfer of financial instruments and transfer of equivalent financial funds so as to achieve delivery against payment (DVP).
8. Complete Settlement: Completion settlement must take place no later than the end date of settlement. Successful settlement on the basis of a real-time settlement mechanism should be considered.
9. Risk control: CSD providing credit facilities to participants, including CSDs that operate net settlement systems, should establish control of risks, ensuring at least successful settlement if the participant with the largest payment obligation can settle the payment. The best way to control risks involves the adoption of a set of measures that combine the use of collateral to limit exposure.
10. Settlement money: Assets used for last payment obligation securities transactions should carry little risk or not be at all affected by credit risk or liquidity risk. If settlement scheme does not involve any use of central bank resources, measures should be taken to protect settlement system members of potential losses and liquidity constraints that may occur due to risk cash settlement agent used to complete transactions with financial instruments.
11. Operational Safety: Sources of operational risk arising in the clearing and settlement should be identified and minimized by adopting an appropriate architecture for the system, through control and adoption of specific procedures. Systems must be safe and have adequate capacity. Action plans and emergency back-up should allow recovery of data and timely completion of settlement.
12. Protect investors’ portfolio: Entities that have custody of financial instruments must use specific procedures to ensure full protection of investors' portfolio. It is essential to have investors' portfolios protected from creditors' claims on the entity that retains custody of financial instruments.
13. Access: CSDs and CCPs should have objective and transparent criteria established for participants, allowing fair and open access to the system.

14. Efficiency: Without affecting operations safety, securities settlement systems should ensure efficiency in terms of costs incurred by participants.
15. Communication standards and procedures: settlement systems securities transactions must use or adopt procedures and standards to facilitate communication to facilitate efficient cross-border transactions. Providing conditions for all participants to communicate in a fast, safe and clear manner is of the outmost importance for the functioning of clearing and settlement systems. It is therefore recommended for all participants in the system to know and to use standard procedures in relation to the messages content, securities identification or other participants. In cross-border transactions is advisable to use international standards: ISO 6166 and ISO15022.
16. Transparency: CSDs and CCPs should provide participants enough information for them to identify and evaluate risks and costs of using services provided by the central depository and central counterparty.
17. Supervision: settlement systems securities transactions should be subject to transparent and effective supervision. Central banks and capital market supervisory authorities must work together or with other authorities with competence in the field.
18. Risks associated with trans-border connections between central securities depositories. The CSD that establish inter-connections with another CSD to make cross-border transactions must design and operate such connections so as to reduce specific risks of cross-border settlements. Cross-border transactions carry most of and the most complex risks that you may encounter when conducting capital market investments. Besides, the costs of performing such transactions are not reduced. To reduce risks and costs involved in carrying out cross-border transactions of financial instruments that otherwise would be borne solely by investors; institutions responsible for the successful settlement of the different countries were involved in establishing relationships to facilitate these operations. Achieving inter-connections between different systems operated by Central Depository make life easier for cross-border investors, but involves close monitoring of such risks if transferred from investors to administrators of clearing and settlement systems.

3.4. Risk management of settlement transactions at the Bucharest Stock Exchange

The services traditionally offered by stock exchanges refer to admission to the stock exchange of the issuing companies, establishing criteria for maintaining a particular security traded, disseminating for the public such stock exchange market information, especially for ensuring mechanisms to concentrate liquidity for a particular security. But continuously exchanges were directly interested in the whole building technical and institutional infrastructure to ensure safe completion and cost of transactions traded on the stock market. Therefore, some scholarships have created their own departments through which to work out the full range of post trading: clearing, settlement and registration. This was the case with the Bucharest Stock Exchange, which since the resumption of operations (in 1995) has provided all operators on the stock market clearing and settlement services [14].

For settlement of stock exchange transactions, BSE chose from the beginning a variant of Model 2 settlement, which requires multilateral clearing and final settlement of funds applied to the net basis, followed by settlement of securities traded, settlement being applied to the gross. The modular structure of the electronic system of BSE that includes trading module, clearing and settlement module and registry module allowed real-time interaction between each of these components, which provided a minimum exposure of the system to operational risk. The settlement cycle used in the stock market in Romania is "T +3" and the final settlement of funds is performed in a separate section of the national payment system (ReGIS), BSE being authorized by the NBR as a inter-banking clearing house since June 9, 1997.

This brief presentation of the clearing and settlement of BSE is only an overview of its main features without revealing all relations established between participants on the stock market from the moment the transaction was closed on the exchange floor until settlement, meaning until the buyer settles all payment obligation according to the contract and the obligation to deliver the securities according to the contract stipulations, accrued to the buyer. To facilitate understanding of information flow that constitutes the transaction settlement system on the BSE it is useful to first review the following terms:

- **Trading Report:** is that statement issued by BSE with terms and conditions of sale contracts of financial instruments concluded on the Exchange floor at a certain date, called the trade date;
- **Compensation Report:** is that document issued by BSE reflecting financial obligations and rights of company's financial instruments or custodian agent resulting from the transaction;
- **Report of Settlement:** is that document issued by the Bucharest Stock Exchange outlining duties and the pay of brokers or custodian agent, resulting from clearing transactions;
- **Settlement Banking Report:** is that document issued by the Bucharest Stock Exchange based on the settlement report that outlines the duties and financial rights of each participating bank and brokers that have an open settlement bank account;
- **Balance for final settlement:** the document issued by the Bucharest Stock Exchange, upon which BSE is entitled to introduce in ReGIS the payment instructions for settlement to be applied to the net basis.

3.4.1. Participants in the clearing and settlement system

Bucharest Stock Exchange is at the heart of the whole system of securities transactions settlement concluded on the stock market, interacting with each of the participants in the system:

- **intermediaries**: are financial investment companies (SSIF) and credit institutions that have the right to trade on the BSE;
- **custodian agencies registered with BSE**;
- **clearing participants:** banks are authorized by the National Bank as credit institutions, that have a settlement contract with BSE, ReGIS participation contract and a set up and enforce securities agreement with NBR. Intermediaries have opened settlement accounts for exchange transactions at the level of the clearing participants.

- **National Bank of Romania**, which is the administrator of the Settlement Express Transfer (ReGIS) system within which the clearing participants hold settlement accounts and participate in the financial settlement transactions of securities concluded with BSE.

3.4.2. Settlement mechanism

Information flow and all the activities leading to successful settlement is achieved over three days (T +3). Presented sequentially, the process is done through the following steps:

1. After the trading session is closed, the Stock Exchange prepares the Compensation Report and the Settlement Report related to securities transactions executed by each intermediary. Intermediaries and custodians registered with BSE agents are required to confirm the next day, the contents of clearing and settlement reports.
2. BSE will issue and send by fax with delivery receipt or secured e-mail to the clearing participants (settlement banks) all bank settlement reports.
3. Participants in the clearing and settlement system managed by BSE are required to check the reports taken from BSE system. If participants in the clearing and settlement system managed by BSE object to the reports, they will communicate this in writing to BSE. After complaints are settled by BSE staff and reports are amended by BSE, documents and operations resume the circuit described above and the brokers modified reports, custodians and that participants will be forwarded by the BSE in the same day. If no objections are raised regarding the data presented in the reports provided by BSE, these are considered to be confirmed.
4. On transaction settlement date (T+3), debtor clearing participants will confirm BSE their consent to participate in the settlement, according to the amounts recorded in the Bank Settlement Report, by fax with delivery note or secured e-mail system. Also, participants will set up deposits equal to the net amount to be paid, in ReGIS in favour of BSE.
5. Based on the final Settlement Balance, BSE will initiate the payment instruction to settle all net positions.
6. National Bank will confirm to Bucharest Stock Exchange that final settlement was concluded by sending the document "Settlement report of net positions".
7. After payment confirmation is received, BSE will transfer financial instruments from the sellers' accounts to the buyers, according to the concluded transactions. This will be accompanied by sending a message confirming to all clearing participants that settlement transaction was closed.

Workflow described above is one in which all intermediaries and settlement system clearing participants honor their obligations that arise from the exchange transactions. Efficiency of clearing and settlement system is checked, but especially when during the settlement cycles a situation where one (or more) of the participants in the system cannot fulfil obligations. To prevent such situations and to solve them when they manifest, clearing and settlement system of the BSE has established a risk management mechanism.

3.4.3. Risk management

Risks that can disrupt the transactions settlement flow have been identified both at the level of the brokers and custodian (first layer) and the clearing participants (second floor). Consequently, risk management measures are structured on two levels:

- Risk management measures managed exclusively by BSE;
- Risk management measures managed together with NBR.

The main instruments to manage risk, available to BSE are the following:

1. In order to limit default exposure on brokers' obligations, BSE sets through procedures a trading limit for each broker.
2. To limit exposure in case of no payment, BSE ask from clearing participants and brokers to give priority to clearing and disbursement of amounts necessary to cover their debit positions.
3. If it is found that upon settlement date, the participant to the clearing and settlement system managed by BSE has limited solvability to cover payment obligations, the clearing participant with whom the broker holds an account may proceed to grant credit, under an agreement signed in advance.
4. If the clearing participant grants participants in the clearing and settlement system managed by BSE a loan needed to cover the debit position, the participant in the clearing and settlement system managed by BSE was in position debtor may proceed to establish a claim pledge of financial instruments existing own account, both those settled, and of the pending settlement in favor of the credit granted. Clearing participant is required to present BSE documents stating the guarantee institution.
5. If the participant does not grant credit or the loan does not fully cover the debit of participants in clearing and settlement system, than BSE shall be able to call the Guarantee Fund.

BSE applies risk management measures as presented above. However, it should be noted that enforcing trading limits to each participant and using the amounts from the Guarantee Fund are the measures with the greatest impact on the entire system thus helping BSE to control risks in the settlement system

3.4.4. Trading limit

The trading limit is the maximum amount limit within each of the participants on the market regulated by BSE can operate over a single trading session. This limit is calculated in real time by the electronic system of BSE, for each intermediate, being affected by each completed transaction and not by the orders entered into the system by the participant. The trading limit is determined by a calculation formula and is affected by exchange transactions concluded on their own by each participant, both on their own name and on behalf of clients.

The formula for calculating the trading limit is as follows:

$$3\ \mathbf{x}\ \left(\mathbf{CIFG}\ +\ \mathbf{CCFG}\ +\ \mathbf{VNC}\right)\ +\ \mathbf{VGB}\ -\ /\ +\ \mathbf{VMj} \tag{4}$$

where:

CIFG = participant's initial contribution to the Guarantee Fund;
CCFG = participant's current contribution to the Guarantee Fund;
NPV = participant's net equity amount, included in the "Minimum net capital and indebtedness indicators";
VGB = amount by which the trading limit may be increased as follows:

- *for purchases* - based on the amount deposited by the participant or his clients in the settlement account (or the accounts "clients") from the settlement bank or other banks, confirmed by presenting the statement of account or other bank documents (such as, for example, the payment order accepted by the bank);
- *for sales* - depending on the amount of sales declared by the participant in writing that he would perform during the trading session

VMj = value that has changed the total amount of loans to the participant and its customers and represents the more or less difference than the value previously reported by the participant. BSE will modify the participant's trading limit to this value.

During the trading session, each time a participant concludes a transaction, the Stock Exchange electronic system checks that the following formula is applied:

$$\mathbf{3x}\left(\mathbf{CIFG}\ +\ \mathbf{CCFG}\ +\ \mathbf{VNC}\right)+\mathbf{VGB}-/+\mathbf{VMj}>=\mathbf{3x}\left(\mathbf{CNCP}\ -\ \mathbf{VNCP}\right)+\mathbf{VNCl}+\mathbf{CNCl} \tag{5}$$

where:

CNCP = are financial instruments purchased by the participant in his own name in that trading session, expressed as values;
VNCP = are financial instruments sold by the participant in his own name in that trading session, expressed as values
CNCl = are financial instruments purchased by the participant in that trading session, in his own name and on behalf of his clients, expressed as values;
VNCl = are financial instruments sold by the participant in that trading session, in his own name and on behalf of his clients, expressed as values

To determine the trading limit, participants are required to submit monthly to BSE's specialized department "Minimum net capital and indebtedness indicators statement" This statement will be sent to BSE within 10 working days from the end of the reporting month. If, during the time between two successive reports, the net capital varies more than + / - 15% compared to the last reports, the participant is required to communicate to BSE the new value immediately.

If during a day, the transactions concluded by a participant affect the relationship of inequality, the participant exceeding his trading limits, the Exchange system will not allow the introduction or modification of the exchange orders on behalf of that participant, as their execution would lead to increase his exposure, and orders that are already in the market will be suspended automatically by electronic system of BSE. Basically, if a participant that has completed operations which no longer meet the inequality relationship described above, will not be allowed to the market access till increasing the trading limit. In order to increase the trading limit is however necessary that the participant expressly requests this by sending a form by fax to BSE in this respect, also assuming responsibility for carrying out the settlement. However, the participant shall submit an estimate of the value of transactions he intends to carry out during the day when the limit was exceeded.

In time, the usefulness of trading limits and especially the formula used to calculate it attracted much discussion at the level of the stock market participants. The main criticism concerns the fact that on the right side of the relationship of inequality the sales and purchases amounts made on behalf of clients are treated similarly and influence the limit in the same way. Even if by making a sale transaction the intermediary does not assume any risk regarding the cash settlement of transactions, however the value of sales transactions made on behalf of clients diminishes daily trading limit of the participant. In fact, you can even consider that sales transactions made by a participant does not induce any risk in the entire settlement system as long as before being accepted an order of sale by the Exchange's electronic system, in the account of the seller are performed verification of the securities, subject to trading order. Considered from the perspective of market participants and given the above considerations, the current formula for calculating the limit of transaction seems unduly restrictive. Intuitively, the limit of transaction reported only net amounts paid in settlement appears to be more reasonable and would create a much closer relationship of emerging risks into achieving financial settlement to the Guarantee fund.

3.4.5. Guarantee fund

As required by the regulations and procedures Bucharest Stock Exchange, "The Guarantee Fund", "the Fund ", has as main objective to ensure necessary resources for proper functioning of the mechanism of settlement of exchange transactions."

Sizing the Guarantee Fund and its constituent sources represent, for any settlement system, two central problems whose solutions can be found only by reference to at least the following elements:

- Settlement model used;
- Costs of setting up such a fund guarantee, costs supported mainly by direct participants in the settlement system;
- Financial strength of the participants in the system, which is closely related to the overall stock market development;
- The value of the transactions made on the stock market;
- The performance of the compensation system (the degree of compensation), which ultimately generates the size of the exposures to settlement;
- Other regulations of risk management;
- Accepting that, in fact, financial risks are minimized and in very few cases are complete eliminated.

As explained earlier, for settling the exchange transactions, Bucharest Stock Exchange is using a variant of Model 2 Settlement (BIS classification), which requires multilateral clearing and final settlement of the funds on a net basis, followed by gross settlement of securities. For such a system, in the literature were directed two ways of determining the optimal value of a Guarantee Fund:

1. For net settlement systems, there are six minimum requirements to be fulfilled. The fourth requirement of this report states that "*as a minimum requirement, compensation schemes and multilateral net settlement shall be able to provide for the daily settlement, when it is impossible to honour the payment obligations by the participant with the highest exposure*";
2. Another opinion is that the minimum requirement to be greater than or equal to the number of days of the settlement cycle (for BSE, three) plus one, multiplied by the greater of the net debtor position recorded in a recent period of time, e.g. last two weeks.

As these recommendations suggest, the total amount of the Guarantee Fund has no absolute meaning, but must be interpreted in light of the whole system of risk management, value of transactions made on the stock market and especially the amount of exposure to settlement . These considerations were taken into account when establishing criteria for setting up and functioning of guarantee fund for settlement of negotiated transactions to BSE.

Guarantee Fund of BSE was established based on initial contributions of all members of the Stock Exchange Association (later, with the transformation of BSE in a stock company, any intermediary agent, carrying out transactions at BSE became eligible to participate in the Guarantee Fund) as well as of all custodian agents authorized by the Stock Exchange.

For the Guarantee Fund to be effective and achieve the purpose for which it was created is necessary that the component assets to have a high degree of liquidity. Therefore, in managing resources in the Fund should be prioritized the safe and liquid investment criteria, the yield obtained in the administration of the Fund represents only a secondary criteria.

4. Conclusions

To increase the accuracy which describes the development of the stock exchange in our country, it is necessary to build a synthetic index. Regarding the risk on the capital market, Risk Grades values determined for BSE indices show that the domestic capital market is more risky than the North American market.

However, analyzing the Risk Grades values obtained for all indexes in the region we notice that they are very close and shows that the investors are not picking up the capital market in Romania as an excessively risky relative to other exchanges in the region. It confirms the hypothesis that the main problems of the capital market in Romania are not related to demand, but derive primarily from an inadequate supply.

The risks managed by Institutions such as market operators, clearing houses and central depositories are very important, even though there maybe be less visible to the public.

Clearing and settlement services ensure that transfers of financial instruments and financial funds between the parties of an exchange transaction to be performed safely and with low costs.

In order to prevent and to reduce the risks specific to the settlement of financial instruments transactions, direct access to clearing and settlement system is limited to those participants who meet a set of capitalization criteria, operational capacity and professional experience, which means that the settlement of an exchange transaction involves a given number of financial intermediaries.

Since the highest financial risk throughout the process of finalizing exchange transactions take place in the last stage of the settlement process, an essential role is associated to the way the connection is made between the securities transfer mechanism (obligation to deliver) and the funds transfer mechanism (payment required). By using mechanisms that ensure Delivery Versus Payment (DVP), one eliminates the most important component of all financial risks during the settlement of the financial instruments transactions that is the principal risk. In order to limit the exposure to the risk of default of the intermediaries' obligations, a trading limit is set for each intermediate part.

Furthermore, the BSE requires its participants to prioritize the amounts necessary to cover their redundancy payment positions. From the market participants' perspective, the current formula for calculating the transaction limit seems unduly restrictive. Intuitively, calculating the transaction limit based only on the net amounts paid in settlement appears to be a more reasonable assumption and could lead to a much closer relationship between the risks emerging in financial settlements and the guarantee fund. Meanwhile, for the Guarantee Fund to be effective and to achieve its purpose for which it was created it is necessary that the assets within it to have a high degree of liquidity. The transition process and the recent international financial crisis have significantly affected not only the evolution of stock markets in Central and Eastern Europe, but also significantly changed the outlook for their future development.

That is why the stock market management should not only consider short-term measures in order to mitigate the crisis impact, but also to try to ensure the prerequisites for further development. The improvement of stock management in general and of stock market risks in particular, requires a medium to long-term perspective, similar to the opinions and contributions formulated by the authors in this chapter.

Author details

Marin Andreica*
Department of Management, Academy of Economic Studies, Bucharest, Romania

Stere Farmache
BSE, Bucharest, Romania

* Corresponding Author

Madalina Ecaterina Andreica
Dept. of Economic Informatics and Cybernetics,
Academy of Economic Studies, Bucharest, Romania
The National Scientific Research Institute of Labour
and Social Protection, Bucharest, Romania

Octavian Stroie
Department of Audit, National Authority for Scientific Research, Bucharest, Romania

5. References

[1] O'Neill Wyss, B. (2001) - Fundamentals of the Stock Market, Mc Graw Hill Printing House
[2] Smith B. M. (2003) - Global Stock Market. From ancient Rome to Silicon Valley, Farrar, Straus and Giroux, LLC, Chicago
[3] Bucharest Stock Exchange (2007) – Manual for the Bucharest Stock Exchange Index, Available: www.bvb.ro
[4] Stoica, V., Ionescu, E. (2002)– Capital Markets and Stock Exchanges, Economic Printing House, Bucharest
[5] Anghelache, G. (2000) – Exchange and OTC, Economic Printing House, Bucharest
[6] Andreica, M., Farmache, S., Andreica, R. Andreica, C. (2007) – Controversies regarding the listing of shares issued by BSE, Vol. Pro-active Partnership in Creativity for Next Generation, Polytechnique International Presses, Quebec.
[7] Andreica,M., Nicolae, D., Andreica, M.E., Todor, L. (2009) – Working out forecast on the Basis of Statistical Extrapolation Principle, Economic Computation and Economic Cybernetics Studies and Research, No.2, Vol. 43
[8] Andreica, M.I., Deac,V., Tipa, S. (2009) – Towards Providing Low-Risk and Economically Feasible Network Data Transfer Services, Proceedings of the 9th WSEAS International Conference on Multimedia, Internet and Video Technologies (MIV), Budapest, Hungary
[9] Farmache, S., Andreica, M. (2010) – Improving Stock Management Assessment Instruments, Metalurgia International, No. 3.
[10] Farmache, S. (2011)– Ways of Stock Management Improvement, Cibernetica MC Printing House, Bucharest
[11] Kim J., Mina J. (2001) - RiskGradesTM Technical Document, The Risk Metrics Group LLC, New York, Available: www.riskmetrics.com
[12] Warsaw Stock Exchange (2009)– Warsaw Stock Exchange Indices, Available: www.wse.com.pl
[13] McAndrews J., Stefanadis C.(2002) -The Consolidation of European Stock Exchanges, Current Issues in Economics and Finance, 8(6) June, Federal Reserve Bank of New York, New York.

[14] Tapking J. (2007) - Pricing of Settlement Link Services and Mergers of Central Securities Depositories, European Central Bank, Working Paper Series No. 710, Frankfurt.

Chapter 16

Importance of Risk Analysis and Management – The Case of Australian Real Estate Market

Gurudeo Anand Tularam and Gowri Sameera Attili

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50669

1. Introduction

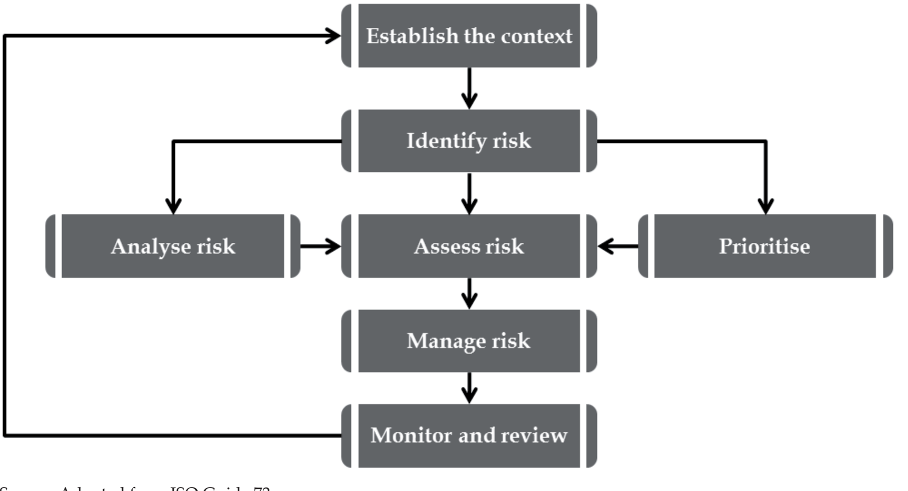
Life is full of risks for example risk is involved in simple things like turning on the gas at home or when dealing with life threatening medical emergency decisions. Risk plays an important role in the way we manage our economy, organization or our family. Risk can be rather complex when household money is involved; such as for individuals or families – for example, mums and dads stand to either gain or lose large sums of money. The types of risks involved influence decisions on how to manage or invest money in shares, bonds or property. When faced with risks, the challenge is how well prepared are we to overcome risks. Risk awareness may be limited in which case there is a high likelihood of risk turning into hazard -leading to disastrous outcomes. Successful businesses make constant efforts to change or update their in house administrative polices and frameworks to allow for possible risks in their business requirements. Some decisions that are likely to have been factored into the component of risk are: rigid corporate governance requirement, human resource planning, succession planning, training and development, merger and acquisitions, adapting to different cultures, foregoing or discontinuing some existing products, outsourcing, new market development etc. No matter how important a decision is made, strategic alignment is critical in business decision making. New ideas should be implemented according to the business needs a company. The introducing of novel ideas should involve all personnel particularly during the decision making processes of development and setting of targets. A well-managed business is also well prepared one and thus able to confront challenges of the modern dynamic business environments.

Yet managing risk is rather challenging for the world is mostly unpredictable. The processes are continuously changing and evolving in terms of resources that are available -

technology, innovation, human resources and time to name a few. In order to adequately address an impending risk, it is important to gather as much factual information as possible for analysis to help manage and thus minimize risk.

Risk can be classified into both voluntary and involuntary [1]. This classification depends on how an individual or an organization judges the situation. For example, a person with a habit of smoking or drinking fails to associate the habits as involving risks; yet often the habit becomes hazardous and they can significantly affect a person's quality life. Involuntary risk places a person or the organization in a state of ambiguity, where the people involved in the decision making process have not been exposed to a particular circumstance or they lack knowledge and awareness of the particular risk situation. The ability to deal with such risks is a crucial factor in determining successful outcomes irrespective of the stature of an individual or an organization.

For some individuals, the ability to deal with risk appears to be built in their character but for the rest of us it seems, it is knowledge that can be acquired through training. In order to gain the skill set required so that one to deal with risk, it is important to step out of one's comfort zone and be willing to change, learn, develop new skills, or be challenged to manage risk. Risk management is a methodical approach that could be taught and learnt by most. The general process and steps involved is presented in Figure 1.



Source: Adapted from ISO Guide 73

Figure 1. The process of risk management

This paper is organized in the following manner: In the next few sections risk is defined and risk management explored focusing on types of risks associated with real estate market. The Australian real estate market is then reviewed and possible risks involved are explored in some depth particularly in terms the global financial crisis. The paper compares the market with the rest of the world and summaries investor risks and rewards in Australian real estate market.

1.1. Definition of risk

In the international context, the ISO 31000/ISO Guide 73: 2009 [2] defines risk as the "effect of uncertainty on objectives" (p. 1). When there is a lack of knowledge or exposure to a certain event then such a situation can be termed uncertain. Taking decision on an uncertain event or situation may or may not be successful, which is what risk is about. Many definitions of risk exist in common usage [3-4]; however the ISO definition of risk was developed by an international committee representing over 30 countries and is based on the input of several thousand subject matter experts.

Risk is defined in Australia by the Australia/New Zealand standard for risk management [2] as "the possibility of something happening that impacts on your objectives. It is the chance to either make a gain or a loss. It is measured in terms of likelihood and consequence…" (p. 2). Risk can also be defined as the uncertainty of future events that might influence the achievement of one or more objectives such as an organization's strategic, operational and financial objectives [3]. Risk management may produce positive opportunities for developers although the negative aspects of risk are usually the once that are emphasized [4].

Likelihood of risk occurring varies from industry to industry and how complex a job maybe. Some areas where there is a high chance of risk are construction, transport, mining, health care, sports, finance and banking, insurance and superannuation.

Risk can be broadly understood and explained in three different scenarios [5]: risk versus probability; risk versus threat; and all outcomes versus negative outcomes. It is believed that any risk can be managed through the engagement of a proper risk management process.

1.2. Risk management

There seems to be an increasing demand of organizations to meet and exceed the financial expectations of shareholders. In the pursuit of growth, many organizations (for example: Toyota) have adapted and responded to expectations of the shareholders by becoming lean and efficient. It is always easy to think that risks and their potential consequences could have been predicted and managed. This is clearly not true when it comes to success in a business. Business success usually requires some acceptance of risk and, as such any risky strategy undertaken may lead to a failure.

In large organizations and corporations there are designated personnel; namely, risk managers. Hillson [6] argued that risk is mostly managed "continuously, both consciously an unconsciously, though rarely systematically" (p. 240). Risk manager's main role is to be aware of the market, collect data and predict forthcoming threats so that a company can manage the risks in a successful manner. Risk manager duties include developing and communicating risk polices and process, building risk models involving market, conducting credit and operational risk analysis, coordinating with concerned stakeholders involved in the process and creating a risk awareness culture in the organization.

Risk management not only prevents organizations from entering a dangerous and uncertain territory, which could lead to a catastrophic failures, but also ensure the development and growth of the business. The depth and clarity with which a risk is defined is critical for risk management. In an event where an organization has a low risk situation at hand and decides to postpone rather than resolve the issue involved for financial or other reasons, the risk may eventually become a threat of moderate to high level and this could prove to be disastrous for management. Ignoring the risks that apply to the business activities or the events that have been planned could impact on the following:

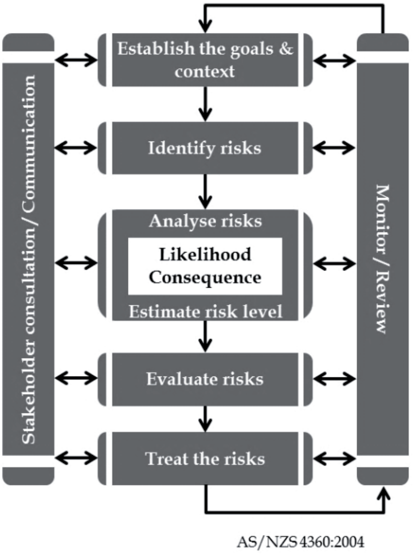
- customer and public confidence in the organization;
- credibility, reputation and status;
- equipment and the environment;
- financial position of the concerned; and
- health and safety of employees, customers, volunteers and participants.

A systematic approach to managing risk is now regarded as best management practice. The approach taken almost always benefits the organization irrespective of type of risk involved. Once the risk is identified it is documented in detail; subsequently the concerned stakeholders undertake possible risk management and mitigation processes. A comprehensive review of the situation and critical feedback are usually required that may ultimately lead to changes in the organizational polices and structures; particularly in case of a major events.

Organizations that thrive to be successful constantly monitor themselves and willfully undertake only calculated risks. In doing so, they enjoy a competitive advantage in addition to meeting their business objectives. In era of globalization, companies are often expanding their business opportunities and in the process, they may undertake challenging and ambitious projects. In most cases, they need to take a number of risks. In this regard, businesses such as Microsoft, Google, and Wal-Mart appear to have been successful global players mainly because they were able to manage risk in a timely manner.

Risk management decisions should be a part of business objectives. Every new project, policy or invention should include all the possible anticipated risks that one may possibly confront. Decision making process needs to consider threats identified, its impact and reaction on the business. By making a careful analysis, companies will have fewer surprises and thus may in the end spend less time recovering from the losses that may be inevitable at times. When companies do not have "a keen eye on the kind of risk", risk retention can become a legitimate way of managing the risk. Figure 2 shows the six steps involved in the risk management process: establish the context, identify the risk, analyze the risk, evaluate the risk, and manage and review the risk.

1.3. The steps involved in managing risk



Source: AS/NZS 4360:2004

Figure 2. The steps in risk management

1.3.1. Establish goals and context

To establish context and define goals is an important step. Once the context is established it is critical that the risk is defined and the objectives are set. Also important is to know the limitations of the risk strategies proposed. An effective risk management team understands the needs of the organization and the way it operates. Once the goal is defined there is a need to identify the scope of the context. In general, these factors can be classified into strategic and operational risks. Strategic risk management includes economic, social, environmental, political, legal and public issues; while operational risk management includes technological, human resource, financial, reputation and other relevant strategic issues. Clearly, management may not be able to totally control the many factors but the risks posed by them could indeed be minimized.

The process of risk management has to be simple, precise and effective. For it to be effective, organizations should consider strength, weakness, opportunities and threats (SWOT) type analysis of the situation. By conducting SWOT analysis, the management can identify and analyze different situations [7]. Once threats are identified, appropriate measures and decisions may then be taken to convert the threat into an opportunity. The organizational context provides an understanding of the organization, its capability and goals, objectives and strategies. In establishing the context the identification of stakeholders is critical; these are individuals who may affect, or be affected by decisions made by the risk management team. For example, stakeholders may be employees, volunteers, visitors, insurance organizations, government agencies or suppliers etc. Each stakeholder will have different needs, concerns and opinions; therefore it is important to communicate with the stakeholders involved in the process of addressing risks.

1.3.2. Identify risks

Identification of risk involves a systematic process of examining situations and finding solutions. The process includes stages such as group discussions and brainstorming sessions to generate a variety of ideas. While all the ideas or issues generated may or may not be relevant, it is important to document all problems, possible impacts and solutions identified. There are four primary areas in which risk can occur in a general business environment:

- financial: this could mean loss of funding, insurance costs, fraud, theft, fees etc.;
- physical: this involves physical assets of the organization, personal injuries and environmental;
- ethical or moral: involves a perpetuated, actual or potential harm to the reputation or beliefs of an individual or organization; and
- legal: this includes responsibilities and adherence to the law, rules and regulations of governing bodies such as the federal, state or local governments.

Risks can be identified by examining records of previous activities or events. Other ways in which risks could be identified are results from past experiences (personal, local or overseas) [8], through conduction interviews of stakeholders (example: Susilawati and Armitage [8]) or by analyzing specific real life or generated scenarios.

1.3.3. Analyse risks

This step determines and addresses the impact of threats that have been documented. Threats identified are rated according to the likelihood of occurrence. The potential of an identified risk can be estimated by the effect it has on financial and other resources. When analyzing a risk, one decides on the relationship between the likelihood of a risk occurring and the consequences of the risk identified. The level of risk is then defined and management of it is then explored. Managing risk can be done in several ways such as contingency planning, using existing assets or making an investment in new resources. The levels of the risks can be classified into

- extreme: an extreme risk requires immediate action as the potential could be devastating to the enterprise;
- high: a high level of risk requires action, as it has the potential to be damaging to the enterprise;
- moderate: allocate specific responsibility to a moderate risk and implement monitoring or response procedures; and
- low: can manage a low level of risk with routine procedures.

The tools most commonly employed to measure risks include qualitative techniques [10]. Melton [11] described the tools as probability and impact analysis tools and Webb [4] called these likelihood and consequences tools. A risk matrix presentation tool (qualitative technique) can provide better insights to the nature of a risk. Risk matrix is often used as a tool to display different risks once they have been analyzed. It allows an organization to mark a threshold above which risks will not be tolerated; or will receive additional treatment from the board or delegated staff. In Figure 3 the threshold is set at risks score of 5 or above. It is then important to ask the following questions in relation to each of the identified risks:

- What is the likelihood of the risk occurring?
- Are there any controls currently in place to manage the risk - if yes then, are there any remaining risks?
- What are the consequences if the risk should occur? and
- What is the level of the risk?

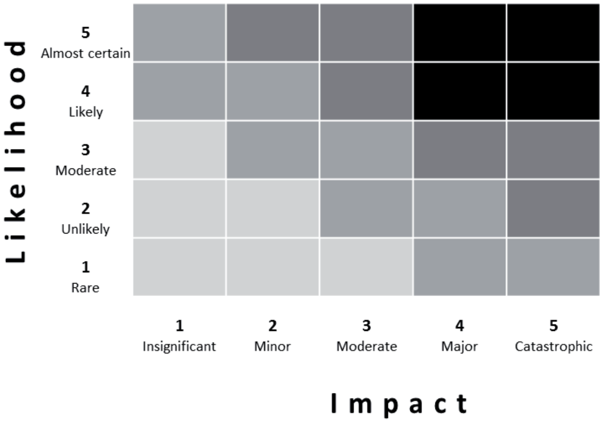


Figure 3. Risk matrix Source: adapted from Austrac

1.3.4. Evaluate risks

In this step the tolerance of the risk is determined; that is, whether the identified risk is acceptable or unacceptable. The evaluation takes into account the following:

- importance of risk management and possible outcomes of a risky activity;
- potential and actual losses that may arise from the risk;
- benefits and opportunities presented by the risk; and
- degree of control one has over the risk.

An acceptable risk is a type of risk that that a business can tolerate; a loss for example- the risk does not have major impact on business. An acceptable risk has to be constantly monitored, reviewed and documented so that it remains tolerable. A risk is deemed to be an acceptable risk because of following reasons:

- risk level is low and the benefits presented by the risk outweigh the cost of managing it;
- risk level is so low that it does not warrant spending time and money to manage it; and
- risk presents opportunities that are much greater than the threats posed by it.

A unacceptable risk is when a business is bound to experience significant losses and such losses cannot be tolerated. In such an event it is important to address and treat the risk in an appropriate manner.

1.3.5. Treatments of risks

Risks may be dealt with in several ways; it can be avoided, reduced, shared or retained. Risk is avoided when appropriate decisions are taken to eliminate all possible pitfalls thereby preventing the situation from occurrence. In most decision making processes, calculations are made and ideas are contemplated to strike a balance between the cost and effect. In such situations calculated risks are accepted and a high risk situation may be reduced by:

- identifying options to treat the risk;
- selecting the best treatment option;
- preparing a risk treatment plan; and
- implementing a risk treatment plan.

In other cases, risk is shared between the stake holders in terms of how profits and losses are shared. This is done mainly to share the impact of a risky event when it occurs. For example, in the era of globalization it is challenging for the companies to enter new markets and countries. In order to minimize uncertainty and exploit business situations that may exist, companies often decide to share risk; careful consideration and research undertaken by the companies often suggest risk sharing. Risk sharing develops opportunities while engaging all partners in achieving strategic goals and the gains and loss are then shared accordingly. The nature of strategies to mitigate risk often depends on the experience of the risk manager who may consider one or more of the following [3]:

- avoid the risk by deciding not to proceed with the activity or choosing another way to achieve the same outcome;
- control the risk by reducing the likelihood of the risk occurring, the consequences of the risk or both;
- transfer the risk by shifting all or part of the responsibility of the risk to another party who is best able to control it; and
- retain the risk after accepting that the risk cannot be avoided, controlled or transferred.

It seems the simplest of all methods of addressing a risk is by retaining an identified risk that may not potentially impact upon the operations of a business. It is important to continuously monitor such risks for in the absence of careful monitoring, the risks may become threats in due time.

A dedication towards risk management often projects a wiser professional image to the community. In doing so, the stake holders recognize the fact that the concerned organization has a keen interest in safeguarding its assets as well as that of its employees, visitors and volunteers among others. In the process of identifying, analyzing and evaluating risks an organization improves its management team's ability to make educated decisions.

1.3.6. Monitor and report effectiveness of risk treatments

Every organization irrespective of size clearly strives to reduce the risks involved. In order to reduce risk organizations have to align their policies and structures in a consistent manner and constantly monitor business activities. Also, there is a need to allocate resources (financial, human resource, technology etc.) efficiently to improve performance and to win the approval of all stake holders. It is also important to ensure personnel working at different levels in the organization report to the appropriate authorities when a risk is identified. Such a culture enables an organization to document and then undertake suitable and timely measures to avert risks. In the risk management process, data capture and reporting can provide valuable insights into the risk management process. A sample risk management planning template is shown in Table 1. As discussed, risk management team play a vital role in identifying and addressing risks.

Risk Assessment Plan
Business name:
ABN:

ACTIVITY STEPS **(Step 1)** List the steps required to perform the activity in the sequence they are carried out.	**POTENTIAL HAZARDS/RISKS** **(Step 2)** Against each activity step list the hazards that could cause emission of refrigerant and describe the risk these hazards pose.	**RISK RATING** **(Step 3)** • Rare; • Unlikely; • Likely; • Almost certain	**RISK CONTROL MEASURES** **(Step 4)** Describe the identified Risk Control measures.	**RISK RATING** **(Step 5)** • Rare; • Unlikely; • Likely • Almost certain	**PERSON RESPON-SIBLE** **(Step 7)** Document the name of the person responsible for implementing risk controls	**TIME-FRAME Monitor & review** **(Step 8)** Document when **step 3** was conducted and when **step 6** is planned
	• • •		• • •			
	• • •		• • •			
	• • •		• • •			
	• • •		• • •			
	• • •		• • •			

Table 1. Risk management planning template

It is necessary to constantly monitor and evaluate the strategies that are employed to manage risks. This is because risks do not remain the same - new risks are created, existing risks are increased or decreased, some risks may no longer exist and previous or existing risk management strategies may no longer be effective. In the end risks can originate from accidents, legal liabilities, natural causes and disasters, uncertainty in financial markets, credit risk, project failures (at any phase in design, development, production, or sustainment life-cycles), or events of unpredictable root-cause. Several risk management standards exist including those from the Project Management Institute, National Institute of Science and Technology, Actuarial Societies, and ISO standards. The risk management definitions, methods and goals vary widely according to the context of project management, security, engineering, industrial processes, financial portfolios, actuarial assessments, public health and safety and real estate.

An important aim of the paper is to study and review the real estate market in Australia to identify risk and rewards as well as compare the Australian market conditions and performance with the rest of the world. Therefore, the focus of the next section is on risks in the real estate market.

1.4. Types of risk associated real estate market

Figure 4. Types of risk in real estate market

As is the case with every other industry, there are several risks in the real estate market. For example, there exists a risk factor in land procurement; housing development; asset management; property management; tenancy management to name a few [13]. The risks may be classified as internal or external risks (Figure 4). In turn, the internal and external risks can be divided into various other risk categories shown in Figure 5 and Figure 6 [14]. Builders, project managers, owners and investors who plan to make an investment or hold an investment in the property market may need to consider one or more of the following risks and then implement appropriate strategies for their projects to be successful.

1.4.1. Internal risk

Internal risk can be divided into financial management, human resources, property management, legislative compliance, corporate governance and housing management as shown in Figure 5.

Financial management: A detailed analysis of any proposed or existing projects need to be conducted for project viability. It is also important to plan the cash flow and management of the same. A poor cost control may lead to a budget over shoot and the project may run into un-chartered territories. When it comes to servicing the debt due care needs to be given to income streams - to take into account either reduction or loss of future income streams. In this regard, banking organisations need to be diligent in testing the capacity to repay the loans that are being offered. Fraud often occurs in real estate market mainly involving the use of false documents regarding number of properties, outgoing fees or rates, income streams and so on.

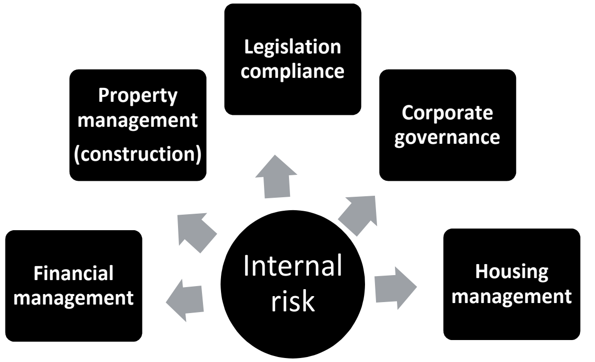


Figure 5. Internal risk Source: adapted from Sheryl and Adam [14]

Insurance also plays a vital role in financial management of a project or investment. Adequate insurance is needed to cover the various risks that may be involved such as the type of property, its location, exposure to natural calamities etc. to name a few. Insurance also need to be updated with the changes in conditions.

Property management of a construction project: During the construction of a new project the builders needs to plan their inventory and keep control of their stocks irrespective of the size of the project. Stock control starts from buying goods to using and maintaining them, and also reusing or reordering as required. Quality of the stock also plays a vital role in real estate business. To maintained quality several techniques are adapted. Just in time technique (where items are ordered when necessary and used immediately), minimum stock level technique and stock review technique.

Contractors play an important role in success of a construction project. They are responsible for recruitment and supervision of employees working on the project. Contractors are also responsible for material management coordinating with suppliers thus acquiring necessary goods in time for the construction phases. Poor response from the contractors or failure to perform their duties will delay the project and overshoot budgets.

Legislation compliance: Often a property holder has to disclose his personal and financial information to third party. Protecting information plays a key issue in this business. Once all the parties are ready to proceed it is necessary to have a privacy act is in place so that all information is secure. The corporation act provides the guidelines for conflicts or issues arising in construction or maintenance of a property. There are several agencies that provide comprehensive legal services to better understand the litigations involved. Anti-discrimination law and disability service act also play an important role in real-estate. Property owners are liable for any discriminatory acts.

Occupational health and safety (OH&S) also arises in real-estate and a number of OH&S compliance officers are usually assigned to monitor the safety and health; for example, conditions provided to the workers at construction sites. OH&S officer duties include inspecting construction sites and providing support to internal clients. It is important to report any hazard or incident and all incidents should be attended to and documented for future reference.

Corporate governance: Corporate governance plays an important role in risk management in the real estate industry. It is important to properly align the ideas, interests and decisions of managers to the interests of both internal and external shareholders. For example, failure to recruit appropriate personnel may lead to conflicts of interest. If the conflicts are not managed effectively they may have a substantial impact on the company bottom line. It is required and expected of the managements or boards of construction companies always carefully analyze performance in terms of the market so that they are able to keep track of their company's performance and progress in a dynamic environment. It is also expected that the managements re-inspect and update their policies and procedures to meet the market trends and demands of all concerned stakeholders.

Housing management: A holistic management of the investment made in real estate can be defined as housing management. Housing management includes keeping track of maintenance and financial arrangements. As a common and popular practice the management of an investment property is outsourced to property management companies who appoint property managers to manage and oversee duties as required. Property managers on a daily basis are responsible for taking maintenance requests, collecting rent, dues or other fees and are responsible for the overall upkeep of the property. They also perform routine property inspections and organize inspections for the owners. Poor performance of the property managers leads to more grievances for the tenants as well as the owners.

1.4.2. External risk

External risk depends on a number of factors such as economic risk, funding, regulation, environment, reputation, competition, partnerships and natural disasters (Figure 2.6). Each of the factors noted are discussed briefly in turn.

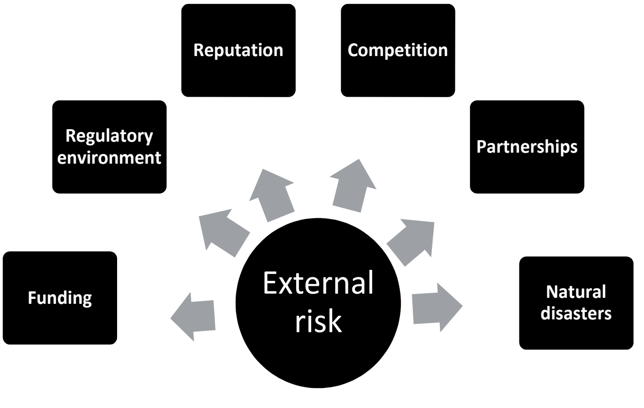


Figure 6. Eternal risks Source: adapted from Sheryl V and Adam W, 2008

Funding: The availability of funding depends on a number of aspects such as the economic situation in general, market performance, and credit based upon any future cash flow. Some factors that influence economic performance are: change in political regime, rise in the price of raw materials, emergence of a new competitor and disruptions in production process. Market performance usually depends on changes in interest rates, changes in laws, and political and financial market factors. The risk of loss of principal or loss of a financial reward stemming from a borrower's failure to repay a loan or otherwise meet a contractual obligation falls under the funding risk. It is important to take into consideration as many of the previously mentioned factors while undertaking an investment decision, even when one already has an investment portfolio. Investors often anticipate future cash flow situations while borrowing money to pay a current debt. The failure of the anticipated cash flow leads to credit risk. However credit risk can be considered less likely since most often the investors are compensated by way of interest payments made by the borrower in end.

Regulatory environment: Investors in real estate projects should be aware of the local, state and federal laws and regulations. These laws depend on economic, credit and market risk as explained above. Failure to comply with the rules and regulation often leads to delays or in the worst case - complete scrapping of the project; all of which may lead to a complete or partial loss of capital invested.

Reputation: The reputation of a project developer often attracts investor attention and also provides favorable environment for investments. Joint ventures and partnerships are possible if the reputations are well known and have been built over time - providing partners the opportunities to win potentially new clients and investors, as well greater opportunities for new investments. An investor has to study the "people" perception of the organization and the credit history and rating of the project developer. An investment made into a company with poor credit history may end up losses of the principle amount invested. It could also be wise for an investor to know the value of the tangible and intangible assets and the market value of the organization into which an investment is being planned.

Competition: Property market plays an important role in the economy. There are several players in the market who usually try to attract investors. While a healthy competition is good for growth in the industry, it is important for the investors to research exactly what they are being offered because the agents often utilize high pressure selling strategies to gain client's cash. It is possible that in the process the investors may receive inappropriate financial advice. For example, consumers may not be aware of non-disclosed information pertaining to advice they receive.

Partnership: Partnership plays an important role in investing, as it reduces the impact of potential risk on the individual or company investment. For an investor to be successful in a real-estate partnership it is important to know the partner well and therefore trust plays a vital role. The role of each partner does need to be well defined and documented. Having a clear legal document will protect the interest of all partners. It also important to plan and document an exit strategy for all involved, because personal situations may change over time. Clearly, before a partnership agreement is made it is necessary to conduct a detail research to become self-confident about the deal.

Natural disasters: In the real real-estate market, location plays an important factor in the investment decision. A property purchased at an appropriate location is expected to provide a good return on the investment. One of the main factors affecting location is the potential exposure to natural calamities such as bushfires, floods, sea level raise and erosion to name a few. If the location has a history or is likely to be exposed to a natural disaster it can be expected that the property prices will eventually be exposed to the risk. Therefore, it is wise to not be enticed into such toxic locations. Other factors that need to be accounted for are the costs of maintenance of properties and the nature and level of insurance required for risky locations, if chosen.

1.5. Risk and reward

The nature of risk definition and management process is such that it should be integrated into “the philosophies, practices and business plans” of any individual investor or large organization’s culture (Hillson [5], p.240). It is certain that there are many risks involved in real-estate market as mentioned. While real-estate provides variety of investment options every investor has to find their comfort level upon taking risks involved. It is not easy to decide if a selected property for investment is appropriate, but the decision should be made based on the consideration of all the factors discussed earlier. In the end however, the willingness to take risks largely depends upon individual preferences and circumstances.

The elements that usually determine the scale of risk or reward are the amount of money that is invested, length of time investment, rate of return or property appreciation, depreciation, fees, taxes, inflation etc. While it is natural for the individual and organizations to invest and expect returns it is important the investors make the informed choice to reduce the odds of losing the principle invested. The potential risks and rewards in investing in the Australian real estate market are investigated next.

2. Real estate scenario in Australia

The speculation about Australian housing market has been intense since 2003. First it was the international monitory fund (IMF) which warned of the housing bubble in Australia “would bust” [15]. In mid-2008, IMF stated that the Australian property market was overvalued by about 25% [16]. In more recent times (April 2010), “The Economist” house price indicators estimated Australian house prices were the most overpriced in the world (56.1% overpriced - against long-run average of price to rents ratio) [17]. The US based analysts Jeremy Grantham (Boston-based hedge fund GMO analysts co-founder) and Heather Hagerty (Fidelity Investments), were also speculating whether or not the Australian residential market is experienced a housing bubble, after the US housing crisis. According to Edward Chancellor [18], a US-based investment strategist and financial author, Australia was "in the midst of an unsustainable housing bubble that could burst at any time" and the "house prices are more than 50% above their fair value - a once in 40-year event." (p.1). In 2011 Morgan Stanley’s global strategist Gerard Minack said that "we've had 20 years where the Australian consumers have been willing to borrow more to buy an asset that they believe always goes up in value. The classic sign of an asset bubble." and that "home prices are 30 to 40% above fair value [p.1, 19].

The house price-to-income ratio has been the main focus in Australia. The house price-to-income ratio is comparatively high when compared to other countries. Also, the price-to-income ratio in Australia since has been more than 40% higher than the long term average. In the next sections a discussion of the fundamentals that govern the house prices in Australian residential housing market is examined. Also, the potential risks and rewards to the investors are explored in terms of the risk analysis framework presented earlier.

2.1. Introduction: How Australian real estate compares to the rest of the world.

Since the U.S. housing crisis, analysts have been speculating about the potential housing bubble in the Australian residential property market. A report by Real Estate Institute of Australia (REIA) argued that analysts primarily focused their attention on the higher house price-to-income ratio in Australia as compared to other countries (REIA 2010). Moreover, it is observed that the house price-to-income ratio levels are at levels that are similar to that in the US before the housing market there crashed in 2008. The raise in the price-to-income ratio in Australia since 2003 by over 40% higher than the long term average adds fuels the speculation. However, it is important to analyze the fundamentals that govern Australian residential market price growth against the rest of world.

2.1.1. Some aspects of the residential finance system in the U.S. and Australia

In the US, the residential finance system played a significant role in the housing bubble of 2008. The regulation, residential finance institutional arrangements, and mortgage characteristics aided the excessive demand for housing finance. Housing finance was available and offered to borrowers with poor borrowing capacities. Consequently, excessive borrowing led to the housing bubble and the collapse of the financial system in the U.S in 2008. There are some fundamental differences in the lending practice in Australia when compared to the US [21].

In Australia the lending process is highly regulated by the institutional arrangement. The lending practices enforce the regulatory provisions on financial institutions forcing them to avoid excessive risk taking behavior. Table 2 outlines the characteristics of housing loans both in the U.S and in Australia. The table highlights the systemic susceptibility to riskier mortgages in the US and that availability of such funds to finance the mortgages were more common than in Australia.

Australia	US
Regulation is high on mortgage loans	No full recourse of mortgages
No negative amortisation of loans	yes
Securitization is low in housing finance	Securitization is high in housing finance
Non-conforming loans	Subprime loans
Full recourse of mortgages	No full recourse of mortgages

Source: RBA [21]

Table 2. Mortgage characteristics of Australia as compared to US

In the US, the non-conforming housing loans represent 13% compared to 1% in Australia [21]. Negative amortization loans are common in the US but no such loans existed in Australia at the time of the crisis. In Australia the mortgages are "full recourse" lenders and hence the incentive that is offered to households to take out loans they cannot repay is reduced. This is also deters financial institutions from offering risky loans. These primary differences stand out to support and contribute to a relatively strong performance of the housing loans in Australia when compared to the US. It is important to note that the share of non-performing loans in Australia were less than 1.5% even during the financial crisis.

Another fundamental difference is that there is no government sponsored enterprise (GSE) in Australia while they exist in the US. The GSE in the US holds a guarantee of the loans that are offered. This potentially provides an impression that bad loans offered to borrowers with poor repayment capacity would be covered by the Federal Government [23]. This is not so in Australia where commercial banks provide 90% of all housing loans. The commercial banks are mainly funded by the bank deposits, short term and long-term wholesale debt [24]. The absence of the so called Federal guarantee restricts Australian banks from any excessive risk taking behavior. In 2007, at the beginning of the financial crisis, GSE's possessed 90% of these securities. The shadow banking system in which the financial institutions have a greater participation and the GSE's can be said to have led the excessive risk taking behavior and practices in the US [21]. In addition, according to the RBA [21], the regulation level of financial institutions in Australia is about 80% while in the US only 50% of all the financial institutions are regulated [21].

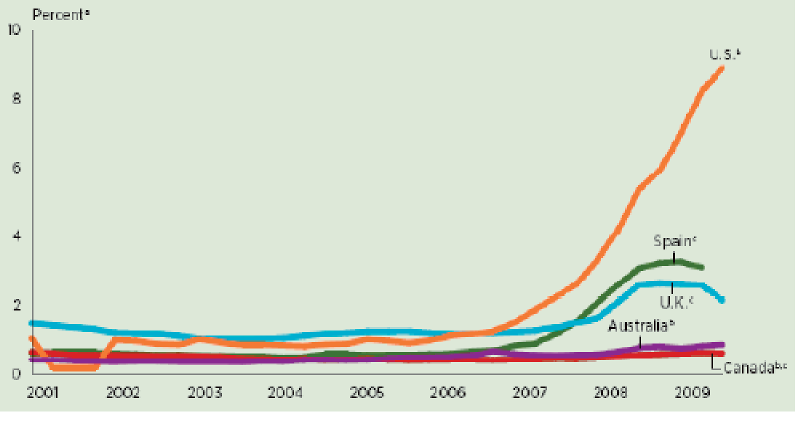


Figure 7. Non-performing housing loans Source: Real estate Institute of Housing America

The Loan to Value Ratio (LVR) refers to the amount of money borrowed against the total value of the property in a home equity loan. For example, a $50,000 loan against a home worth $200,000 has a Loan to Value Ratio of 25%. In Australia, loans with an LVR exceeding 80% require mortgage insurance - the risk of the borrower defaulting is far too great for the lender. The value of the property is determined by the lender and is often significantly less than the purchase price, which often surprise first-time borrowers. Typically, the amount that lenders have been prepared to lend for housing has been restricted by one or both of the following:

- scheduled repayments should not exceed some fixed share of the borrower's income – the repayment-to-income, or serviceability, constraint; and
- the loan should not exceed a certain proportion, most commonly 80% [21] of the property's purchase price – the LVR constraint.

2.2. Australian real estate market compared to the rest of the world

The analysis presented in the previous section shows that Australia is fundamentally different to the US when it comes to the residential housing market. But, how does Australia compare to the other countries in the world? New research conducted by Lloyds TSB [27] - International Global Housing Market Review, shows that Australia just made it into the top 10 list of countries with the highest house price increases over the past decade (Table 3). Four of the six top performing housing markets since 2001 were in the emerging economies of the world. India with a booming real estate market tops the list - house prices rise by 284% over the last decade; Russia coming second - house price increase of 209% over the same period. China faired only marginally when compared to other major economies - ranked 14th with a 47% growth rate since 2001.

Country	Real house price change % 1 year	Real house price change % 10 year	Real house price change % 10 year per annum
India	8.7	284	14.4
Russia	-24.3	209	12
South Africa	-1.1	161	10.1
Lithuania	-1.3	143	9.3
Hong Kong	13.6	125	8.4
Bulgaria	-10	106	7.5
France	4.3	82	6.2
New Zealand	-2.1	79	6
Australia	*-5.7*	*76*	*5.8*
Norway	6.9	72	5.5
Belgium	0.4	69	5.4

Source: Lloyds TSB International [27]

Table 3. Real house price changes – A global comparison.

According to the findings of the report Australian house prices increased by 76% and had the ninth fastest growing house prices during 2001-2011. During the same period house price declines were seen in the world's largest economies such as Germany, Japan and the

United States. Japan registered the largest house prices fall of 30%, while house prices in Germany and US were down 17% and 2% respectively during the same time. Other major findings of the research include:

- housing markets have typically risen fastest in countries with the fastest growing economies. On average, the countries with the biggest rises in house prices since 2001 have seen GDP increase by more than 100%. Countries that had large rises in pre-crisis times lost the most after the GFC affected their economies; and
- house prices within countries that form part of the Euro have climbed an average of 23 percent since 2001. France saw the largest increase with 82%, Belgium rose 69%, Spain 26% and Italy was up 31%. But Spain has seen a major decline in 2012.

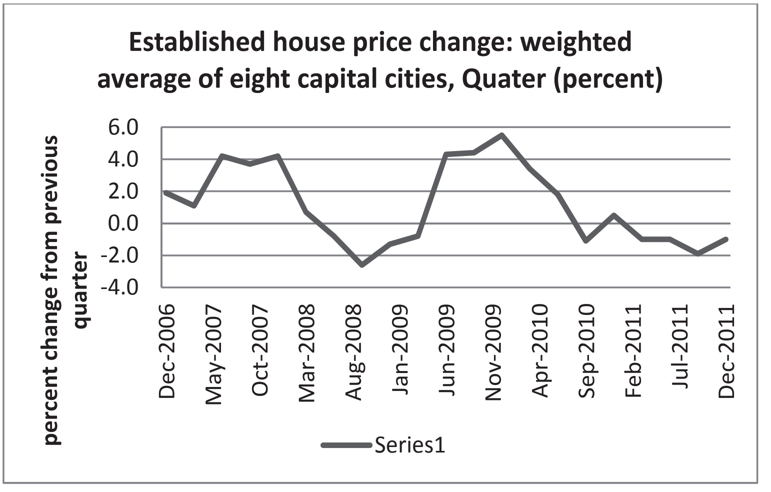
The performance of the established house prices in Australian housing market provided by the Australian Bureau of Statistics (ABS) is presented in Figure 8. The Australian housing over the past five years has seen some corrections. The period can be divided into pre-global financial crisis (GFC), during GFC and post GFC. Prior to GFC, there has been a considerable growth in the established housing prices. This growth pattern however changed course and reached the worst levels in August 2008 when the GFC was setting in. However, the prices of established homes climbed steeply during the peak of the GFC when markets around the world were playing havoc. This defiance could be mainly attributed to the management initiatives taken by the RBA [21] and government of Australia. The RBA drastically reduced the interest rates to a record low of 3.25% supported by the federal government incentives such as economic stimulus plan, which included substantial increase in first home grants among others.

This financial incentive was “too good to miss” for anyone considering their first home purchase. This led to flood of first home buyers entering the market that drove the prices up against all odds. Since the time the incentives have been wound back, and the market and investor sentiment took over. This led to a fall in the growth when compared to the preceding three years and has been mostly in the low sentiment in the past two years. Therefore, although Australian market prices are influenced by the global events, a collapse similar to that seen in markets elsewhere seems appears a distant possibility. This can be attributed to the underlying government incentives to manage the risks during the crisis. Other micro-economics aspects also helped manage the downturn.

2.3. Australian house prices and the fundamental influences

Australian housing demand has been strong and can be also attributed to the following:

- strong overseas migration from 2004 to 2007;
- housing shortages due to a rapidly growing population;
- Australian household sizes are shrinking;
- lending standards stricter than most advanced economies including the US; and
- interest rates at record lows.

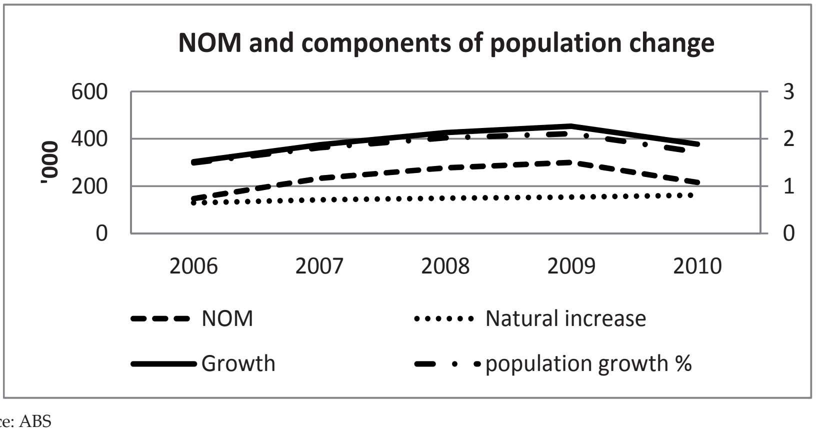


Source: ABS

Figure 8. Australian annual house price change in the last decade

2.3.1. Trend of net population increase and net overseas migration increase

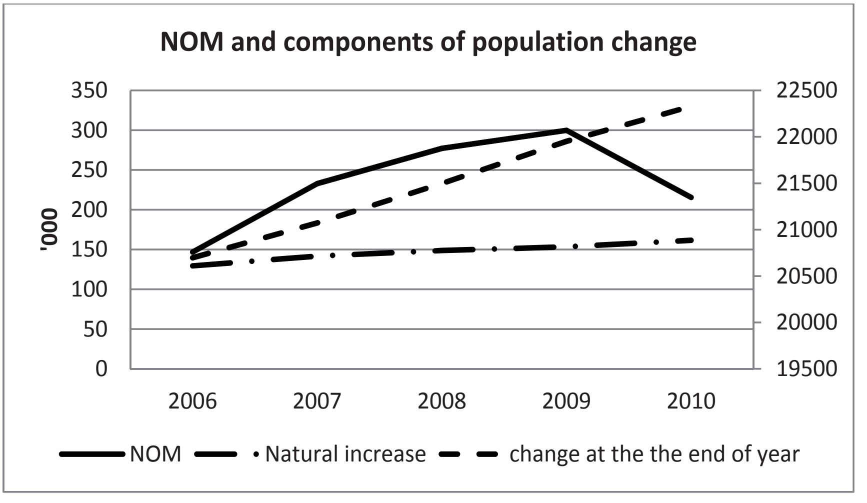
House prices have been underpinned by a chronic housing shortage in Australia. This was brought about by an ever increasing population and constraints placed on housing supply over time. Figure 9 shows the increase in population growth from both natural growth and migration since 2006. From 2006 to September 2010 natural population growth has only seen a marginal increase, but during the same period the net overseas migration growth has been substantial.



Source: ABS

Figure 9. Trend of natural population increase and net overseas migration

Figure 10 shows that there has been an increase in the total population by about 1.6 million people 2006–2010. During the same period, the Net Overseas Migration (NOM) accounted for 1.02 million people compared to only 600 000 increase in natural population. However, given that there has been a large influx of people into Australia, the question was whether there was enough housing infrastructure in place.



Source: ABS

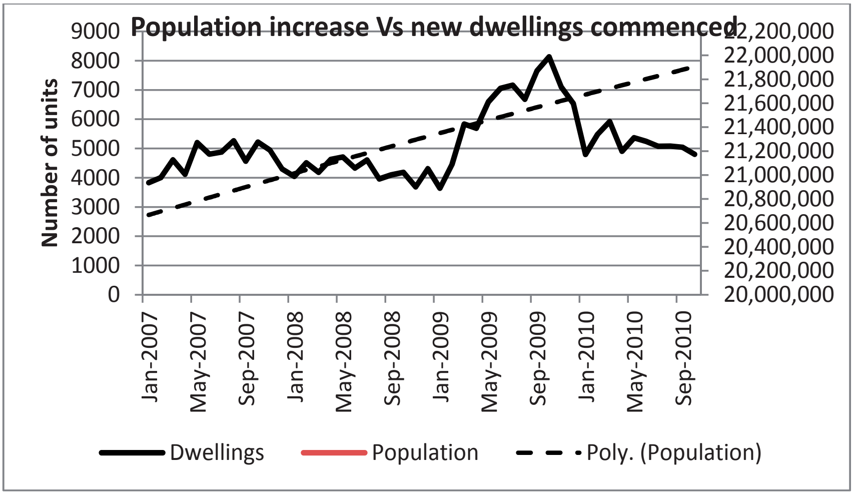
Figure 10. Net overseas migration and components of population change

2.4. Trend in the number of dwellings commenced and population

Figure 11 shows the trend in the population and dwellings commenced from January 2007 to October 2010. As shown earlier, the population growth showed an upward trend over the entire period. The number of dwellings commenced shows a rather distressing trend. Figure 11 shows the commencement of new dwellings significantly fell short and did not keep pace with the rapid growth in population. For an addition of 1.25 million people during this period only about 235 000 new homes were built demonstrating a significant shortage in the housing market. Interestingly, this situation presents a case for more property investment as people search for a place to live.

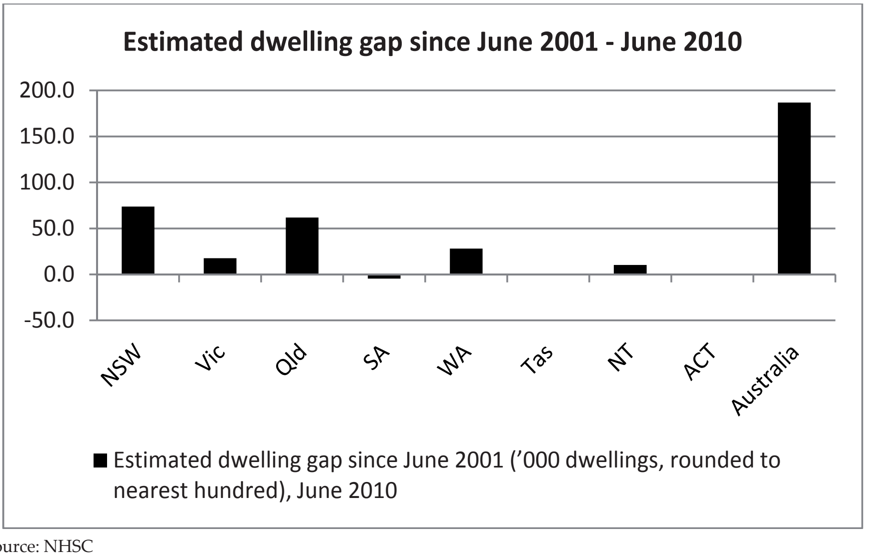
2.5. Demand and supply scenario

Historically, Australia has been behind in the demand versus supply of residential dwellings, but more so in the last decade than any time earlier. Figure 12 shows the dwelling gap in the previous decade. Australia continues to run large annual deficits in housing supply - the underlying demand for dwellings and the completion of dwellings has not matched. In view of this it can be expected that in the longer term Australia's housing market is underpinned by insufficient supply in addition to robust underlying demand.



Source: ABS

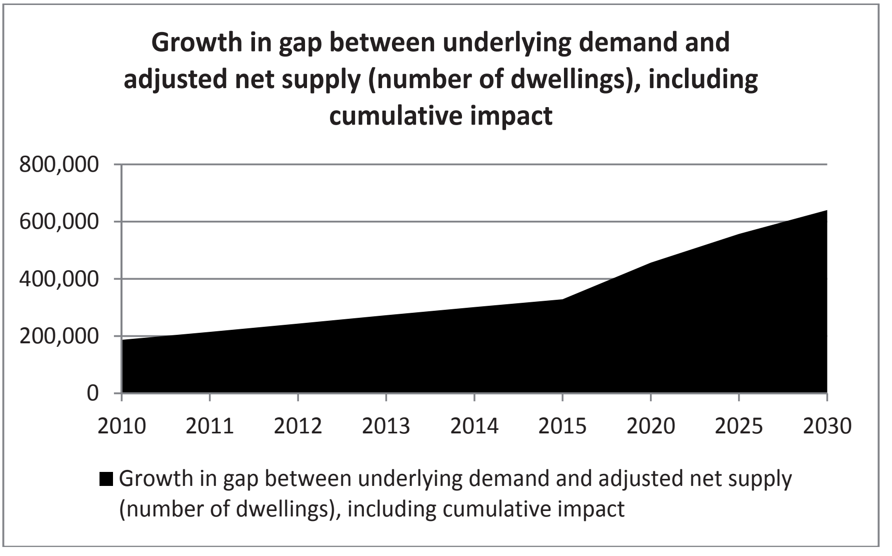
Figure 11. Trend in the number of dwellings commenced and population



Source: NHSC

Figure 12. Estimated dwelling gap in the last decade

National housing supply council (NHSC) estimates a demand versus supply gap of approximately 640 000 houses in 2030; and an increase in the gap from 250 000 in 2012. Figure 13 shows the projections in the supply gap to 2030. The figure shows an increase over time till 2015, and indeed a higher rate of increase predicted from 2015till 2030.

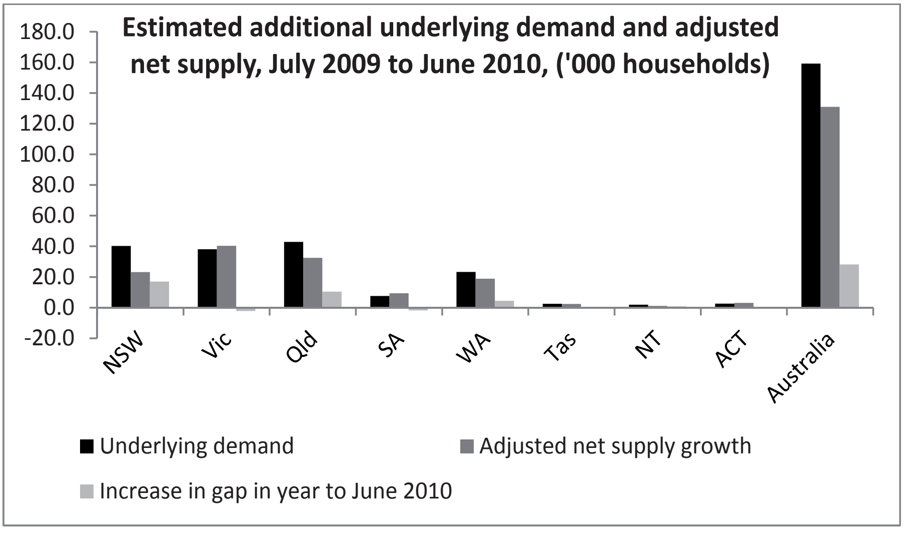


Source: NHSC

Figure 13. Supply and demand gap projections to 2030

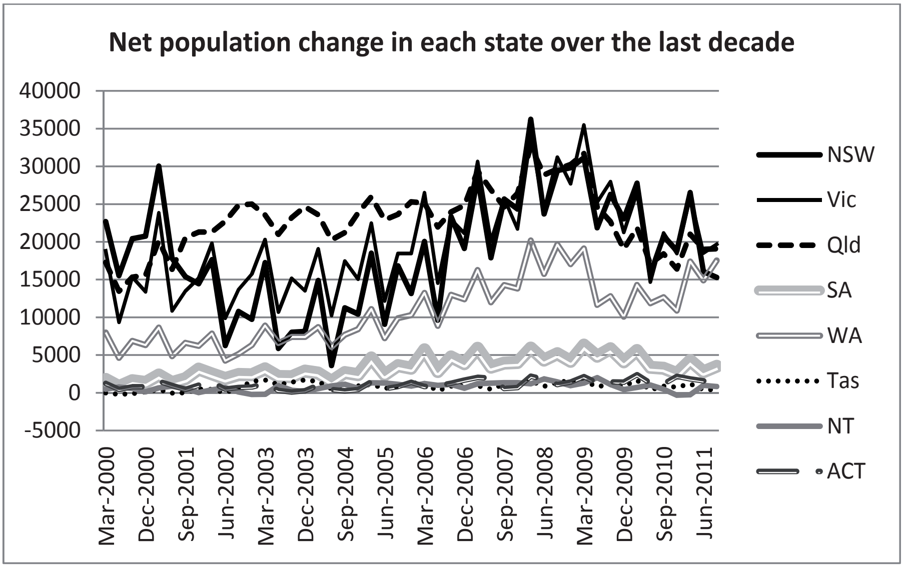
To examine whether the situation is the same throughout Australia or mainly confined to a few states, data from all the states are explored in more depth. Figure 12 and Figure 14 both show that not all states have an acute shortage of housing such as South Australia (SA), Tasmania (Tas) and Australian Commonwealth Territory (ACT). Their data runs against the trend for the last decade but more so during 2009-2010. The larger states of New South Wales (NSW), Victoria (Vic), Queensland (Qld) and Western Australia (WA) all continue to have high deficits year after year and the deficit is increasing – however, Victoria being an exception in 2009-2010 where it managed to go against the trend temporarily (Figure 14). To further understand the nature of the differences between states, the net population increase in the demand across states needs to be compared. Figure 15 shows the state by state net change in population as well as housing issues. The states with a high influx of population showed higher dwelling demand.

Not surprisingly, the high demand has led to a rather strong rental market particularly in the larger states and this has provided an impetus for higher rental returns and an ideal time for new investors to consider for the longer term. With recent housing approvals declining, this demand supply gap can only be expected to widen. Clearly, the population increase cannot only be driving the market. Therefore, other aspects need investigation such as house price to income ratio; and house hold debt to income ratio.



Source: NHSC

Figure 14. Housing demand and supply by states



Source: ABS

Figure 15. Net population change - state by state over 2000-2011

2.6. House price-to-income ratio

The house price-to-income ratio is generally calculated using average income of the whole population. This method of calculating house price may not be appropriate in that a set of buyers whose incomes are above the average income of the wider population, and have the ability to service the loans tend to bid in the auctions there by inflating house prices [28]. Such competition is visible across all capital cities but more so in Sydney, Melbourne, Perth and Canberra than other cities. Figure 16 shows the median change in the house prices across eight capital cities since 2007.

Figure 16 shows that the increase in house prices in the major capital cities have been greater than those of other cities. This suggests the increase in house prices in Australia over the past five years was driven mostly by house prices in the most expensive cities, where home buyers tend to be higher income earners. The house price-to-income ratio does not seem to pick up the distributional differences. The household debt to disposable income ratio can provide valuable insights while assessing the vulnerabilities. Therefore, disposable incomes of people need to be considered when assessing the vulnerability of an average mum and dad investor.

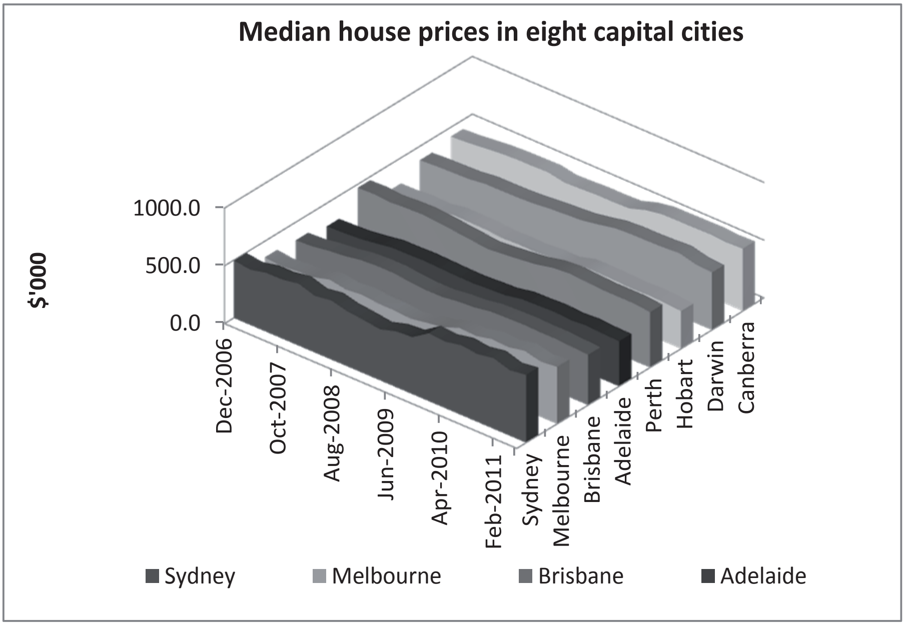


Figure 16. Dwelling prices in capital cities in Australia Source: ABS

2.7. Owner- Occupier debt

Figure 17 shows the distribution of debt to income since 2006. The data indicates that the debt to income ratios has been fairly high – but consistent around 160% for the total debt, of

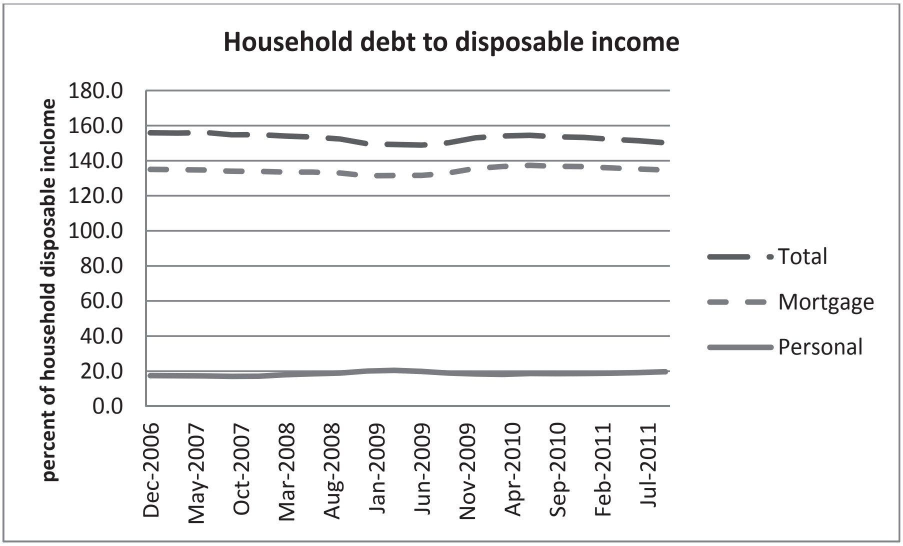


Figure 17. Owner occupier debt Source: RBA

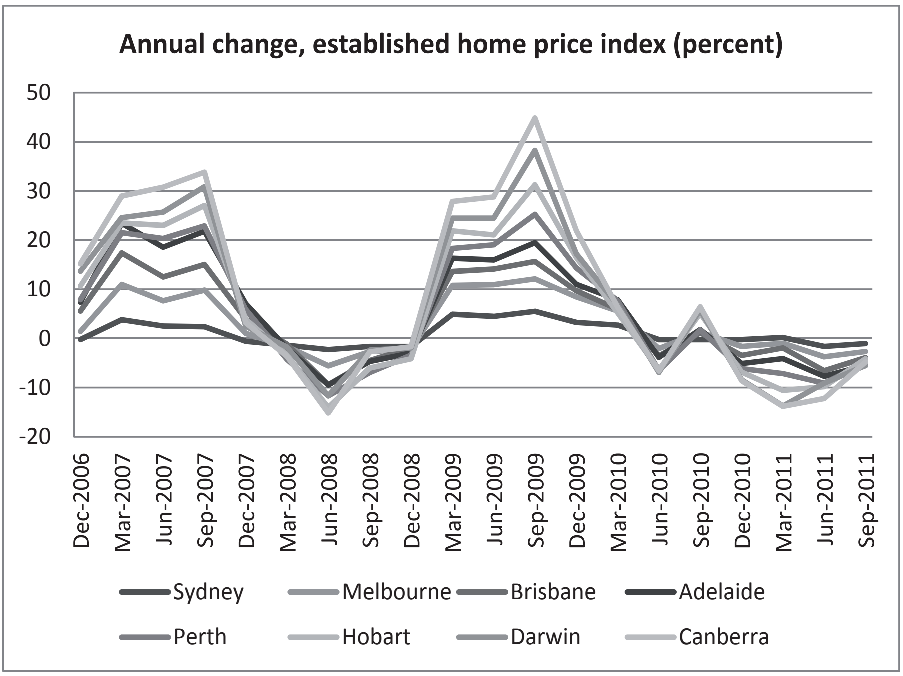


Figure 18. Annual change in established home prices Source: ABS

which close to 140% is towards the mortgage. An indication to the scale of vulnerability can become salient when the house hold income to debt and the annual change in established home price are compared. Figure 18 shows that there has been a somewhat volatile situation in the housing market in all capital cities during 2006-2011; yet, during the same period, the debt to income ratio seem to be approximately constant over time. The comparison shows the average households are not so vulnerable to at least a change in their income situation given there was volatility in house price changes over time.

3. Conclusion

The aim of this paper was to define risk and risk management in terms of real estate investment thus demonstrating the in depth nature and complexity of the process. Another aim was to conduct risk analysis of the Australian real estate market in particular, in terms of the global financial crisis – pre GFC, during GFC and post GFC. The review shows that risk analysis involves a number of steps with each step in turn involving another set of procedures. Risk analysis is a process that it is often ignored by investors particularly by the individual or smaller investors who tend to be more vulnerable. Similarly, risk management involves a number of processes and stages with steps and these have been outlined in the paper. A risk analysis is conducted here for investors in Australia real estate market. The results are rather interesting in that several conditional differences exist between Australia and the rest of the world. The factors identified that influence Australia's house price are different from the rest of the world; including for example the rather stricter and well regulated lending practices of Australia's financial institutions. A tight financial system regulation in Australia means a highly disciplined financial sector. The tougher regulation of the industry therefore prevents financial institutions from taking on excessive risks, contrary to the US counterparts. In fact, increasing house prices was identified in Australia after the crises of 2007-8; and this was associated with the changes in mortgage lending rates, rising family income, increasing overseas migration demand, government incentives to name a few. Together the market situation suggests that Australia is unlikely to face a US style housing bubble. The results of the risk analysis show that:

- rising incomes and population growth ensure the demand for housing outpaces current supply, thereby increasing the prices;
- high capital growth in larger cities where there has been large population migration such as Perth, Sydney and Melbourne;
- high demand still exists for residential and commercial real estate to accommodate growing expatriate working community;
- increased property prices has to many Australians increasingly seeking rental accommodation, making housing investment a healthy growth area for investors;
- higher growth rate in property investment in Australia - superior to most OECD countries, including the UK, Spain and the US; and

- foreign exchange rate changes have been favorable, making property purchase in Australia a valuable option; that in turn driving property prices higher. This has changed in 2011-12 when the higher Australian dollar has posed interesting challenges for the Australian investments.

The findings are in line and relate to that of the Australian housing and urban research institute's findings [29], which further suggest:

- investors are motivated to invest in the private rental market for a number of reasons such as financial factors, personal goals (retirement or future home for children at university), and household circumstances (proximity to their own dwelling);
- investors use their own measures of quality and personal preference when selecting a dwelling even though they will not be living in the property;
- investors perceive property as a long-term, safe and stable investment that is low risk and will produce guaranteed returns;
- investors largely expect capital gains from investing rather that rental yield only and this is how success is measured; and
- informality characterizes investor approaches to the housing market where property is considered familiar, relatively easy to invest in when compared to other investments.

In summary, Australian housing industry continues to experience significant housing shortages in major cities due to a rapidly growing population; in particular, the growth has been fueled by strong overseas migration during 2004-2007, but the Australian current government immigration laws suggest that the strong levels of immigration will continue for some time due to the lack of skills in the labor market. The housing demand is further supported by the fact that the size of the Australian household appears to be shrinking adding to the pressure on housing both in rental and investment. The demand of rental housing together with somewhat lower house prices in recent times (buyer marker) has lured many new investors in the market. This aspect, the negative gearing benefits, and the first home ownership schemes supported by significantly lower interest rates have all led to a favorable and stronger real estate market in Australia. All of this has occurred within a framework of a stronger, tightly regulated financial sector that has been more-stricter than most advanced economies including the US. Such a regulated real estate market appears to have kept the mortgage repayment failure and housing related bad debts at a minimum in Australia.

Author details

Gurudeo Anand Tularam and Gowri Sameera Attili
Griffith University, Australia
Science Environment Engineering and Technology [ENV], Environmental Futures Centre, Brisbane, Australia

4. References

[1] Sunstein CR. A note on “voluntary” versus “involuntary” risks. Duke Environmental Law & Policy Forum. 1997 Fall; 8(1):173-180.

[2] Standards Australia/Standards New Zealand. AS/NZS 4360:2004 : Risk management. 3rd ed. Sydney: Standards Australia International Ltd and Standards New Zealand. 2004.

[3] International Standard Organization. Guide 73: Risk management — Vocabulary. 2009; [cited 2012 March 4] [about 24 screens] Available from: http://pqm-online.com/assets/files/standards/iso_iec_guide_73-2009.pdf

[4] Comcover Insurance & Risk Management. The Introduction of AS/NZS/ISO 31000:2009 - Risk Management Principles and Guidelines. Issue 32; [cited 2012 March 17] [about 2 screens] Available from: http://finance.gov.au/comcover/docs/IB_Issue32_Oct09.pdf

[5] Hillson D, Murray-Webster R. Understanding and managing risk attitude. 2nd ed. Aldershot: Gower; 2007.

[6] Webb A. The project manager's guide to handling risk. Burlington: Gower; 2003.

[7] IRM, AIRMIC, ALARM. A risk management standard. London: IRM; 2002.

[8] Hillson D. Extending the risk process to manage opportunities. Int J of Project Management. 2002 April; 20(3):235-241.

[9] ACT Insurance Authority. Guide to Risk Management. 2004 February; [cited 2012 March 10] [about 10 screens] Available from: http://treasury.act.gov.au/actia/guide.doc

[10] Susilawati C. Can risk management boost the supply of affordable housing development and management?. Int J Housing Markets Analysis. 2009; 2(4): 392-402.

[11] Susilawati C, Armitage L. Affordable Housing: who supply it? Tenth Annual Conference, Pacific Rim Real Estate Society; 2004 January.

[12] Elenor C. Managing Risk in Community Housing Operations: Concepts and Tools. NCHF [serial on the internet]. 2006 April; [cited 2012 March 19]: [about 52 screens]. Available from: http://pandora.nla.gov.au/pan/105045/20090903-1455/www.nchf.org.au/downloads/ManageRisk2006.pdf

[13] Melton T. Project management toolkit: the basics for project success. 2nd ed. Oxford: Elsevier; 2007.

[14] Hillson D, Murray-Webster R. Understanding and managing risk attitude. 2nd ed. Aldershot: Gower; 2007.

[15] Bisset H, Milligan V. Risk Management in Community Housing: Managing the challenges posed by growth in the provision of affordable housing. NCHF [serial on the internet]. 2004 December; [cited 2012 March 18]: [about 76 screens]. Available from:

[16] http://pandora.nla.gov.au/pan/105045/20090903-1455/www.nchf.org.au/downloads/risk_management_summary.pdf

[17] Sheryl V and Adam W. Housing Association: Risk Management Plan (template). Version 5. 2008 December. [cited 2012 March 18]: [about 22 screens]. Available from: http://google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CFQQFjAA&url=http%3A%2F%2Fcommunityhousing.org.au%2Fresourcebank_docs%2Frisk_mgmt_docs_GovGPG3%2FHA%2520template%2520v5%2520Dec%252008.doc&ei=AiiNT5KVMumviQes6a3_DA&usg=AFQjCNEAsSuQamg9mQaask5Xrbe3mHReXg
[18] ABC [homepage on the internet]. ABC Materials; [updated 2003 Apr 13; cited 2012 March 8]. IMF predicts Australia housing bust; [about 3 screens]. Available from: http://abc.net.au/news/2003-04-13/imf-predicts-aust-housing-bust/1835644
[19] Smart Company [homepage on the internet]. Smart Company.com.au Pty Ltd; [updated 2008 Apr 4; cited 2012 February 28]. Australian property bubble could be about to burst: IMF; [about 3 screens]. Available from: http://smartcompany.com.au/
[20] The Economist [homepage on the internet]. London: The Economist Newspaper Ltd; [updated 2010 Apr 15; cited 2012 March 25].You can't keep 'em down; [about 3 screens]. Available from: http://economist.com/node/15911113?story_id=15911113
[21] The Australian [homepage on the internet]. News Pty Ltd [updated 2010 May 3; cited 2012 March 15]. Housing tipped for price implosion; [about 4 screens]. Available from: http://theaustralian.com.au/archive/business-old/housing-tipped-for-price-implosion/story-e6frg9gx-1225861304871?from=public_rss
[22] ABC [homepage on the internet]. ABC Materials [updated 2011 Mar 11; cited 2012 March 8]. Australian house prices 56 per cent over valued: The Economist; Available from: http://abc.net.au/pm/content/2011/s3155728.htm
[23] Real Estate Institute of Australia. Australian house prices: bursting the bubble myth. 2010; [cited 2012 April 5]. Available from: http://reia.com.au/search-result.php
[24] Reserve Bank of Australia [homepage on the Internet]. Reserve Bank of Australia C2001-2012; [cited 2012 March 7]. Available from: http://rba.gov.au/
[25] Research Institute of Housing America [homepage on the Internet]. Washington: mortgage bankers association; c2008-2012; [updated 2011 March 30; cited 2012 March 12]. Available from: http://housingamerica.org/default.htm
[26] Carmassi J, Gros D, Stefano M. The Global Financial Crisis: Causes and Cures. J Common Market Stud. 2009 November; 47(5): 977-996.
[27] Brown A, Fabbro D, Hanrick T, Davies M. Recent Developments in Banks' Funding Costs and Lending Rates. Bulletin. 2010 March: 35-44.
[28] Reserve Bank of Australia. Financial Stability Review. 2010 September 29.
[29] Reserve Bank of Australia. Recent Developments in Margin Lending in Australia. Bulletin. 2009 December: 11-17.
[30] Lloyds banking group [homepage on the internet]. Lloyds banking group plc [cited 2012 March 15]. Emerging markets top global house price league over the past decade. Available from: http://lloydsbankinggroup.com/media/pdfs/LTSB/2012/1703_global.pdf

[31] OECD. Recent house price developments: the role of fundamentals. OECD economic outlook. 2006 Jan 19; 2005(2): 123-154.

[32] Seeling T, Thompson A, Burke T, McNelis S, Pinnegar S, Morris A. What motivates households to invest in the private rental market?. 2009 Jun 22; [cited 2012 March 5]. Available from: http://google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDQQFjAA&url=http%3A%2F%2Fwww.ahuri.edu.au%2Fpublications%2Fdownload%2Frap_issue_115&ei=aDGNT8nfK--QiAeC6K3JDA&usg=AFQjCNGpcAlAjmJeWptc7zFKMCKw0iJ5TQ

Quality Risk Analysis:
Value for Money in the Pharmaceutical Industry

Jordi Botet

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50450

1. Introduction

Pharmaceuticals play an important role in keeping people fit, but they can also put live at risk if they don't have the required quality. Contamination and mix-up may have a great impact on them because of their tiny active doses and because of the often precarious state of health of the patients, not to mention the existence of routes of administration, which skip certain defense barriers of the body.

To cope with this problem "Good Manufacturing Practice" (GMP) was introduced in the 1960s with the intention of providing a kind of common quality baseline for all laboratories. GMP, however, consists of general rules, and as such, it can neither give an answer to every practical situation, nor replace the need to study and understand processes in depth, as some people wanted to believe.

This is why the American FDA initiative on GMP, launched in 2002, underlined the need of taking decisions based on knowledge and science "in [1]".

The ICH (International conference on harmonization of technical requirements for the registration of pharmaceuticals for human use) has given world-wide diffusion to this initiative and put it into practice by publishing several closely related Q (quality) guidelines "in [2-5]".

These guidelines have to be applied conjointly in order to ensure that the quality of a product is, first of all, developed and, then, monitored within a quality management system. This pharmaceutical quality system, as defined by ICH guideline Q10, has two "enablers": knowledge management and quality risk management.

Knowledge management is defined as *a systematic approach to acquiring, analyzing, storing, and disseminating information related to products, manufacturing processes and components*.

Whereas, quality risk management is described as *a systematic process for the assessment, control, communication and review of risks to the quality of the medicinal product across the product lifecycle.*

They are called enablers because they constitute *a tool or process which provides the means to achieve an objective.*

The importance of quality risk management (QRM) is such that a whole ICH guideline, Q9, has been devoted to it.

Thus, QRM combined with GMP and science is a kind of "magical potion", which we can use to become "wizards" ensuring quality. However, this is only true if we understand what QRM is and use it in the right way.

Unfortunately practice shows that the real role of QRM is not always understood and as it has already happened in the recent past (e.g., with validation) it can become something that is only done, because it is required by the Authorities, but that it does not yield what it might and is just written for the occasion, shown and filed. And this is not something unimportant because resources which are misused here become resources that lack there…

Let us then review some key points for making the most of quality risk management (QRM).

2. Knowledge is the base of everything

QRM is a powerful machine but to move has to be filled with a kind of fuel called "knowledge". If we don't know anything about our product or process, we cannot manage risk. Risk, by definition, is the combination of the probability and severity of a hazard. And if we don't know the hazards, not to say their likelihood and importance, there is no way for QRM.

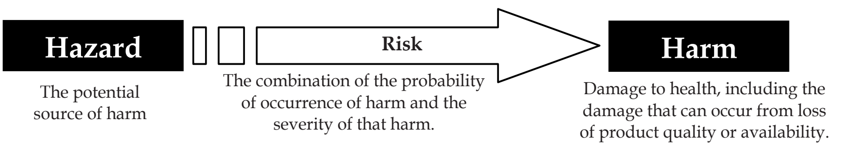


Figure 1. Hazard – Risk - Harm

This explains why knowledge management is put beside QRM as an enabler. Experience shows that more often than not information on the products and on the processes is "lost", either because it is not duly registered and disappears or because it is just kept away by a given person and never diffused within the company. The result is that in many cases there is or there was information but it is not available when needed by the persons of the company who have to take a decision.

And the fact is that if we cannot gather information on the product or process, either in house or outside it (other sites, publications, courses, etc.) it is very unlikely that we might follow a QRM approach. In other words, both enablers are linked and thus, the amount and characteristics of the knowledge at our disposal will be one of the main factors which will dictate how to perform our QRM.

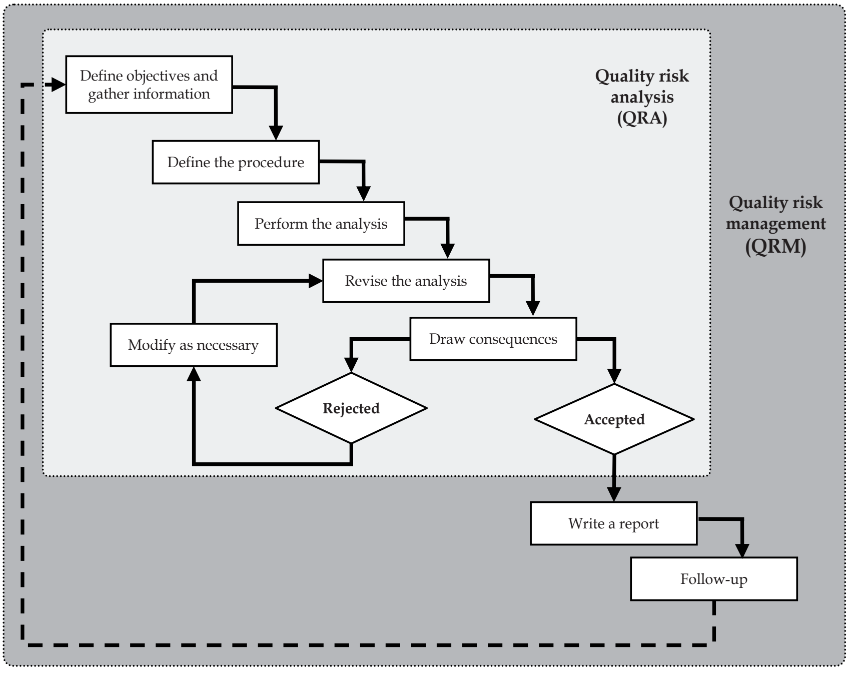


Figure 2. Quality risk management steps

QRM is the result of a certain number of operations or steps, which can be summarized in different ways differing only slightly one from the other. The first part of QRM is evidently devoted to the quality risk analysis (QRA), that is, to *the estimation of the risk associated with the identified hazards*, whereas the second one concerns, properly speaking, the administration of this risk. Any QRM process has to start by defining its goal (what is intended) and by gathering information. All other steps are shaped by this first one.

But, speaking in practical terms, what kind of knowledge we need? Let us try to respond to this question by considering three different cases.

2.1. Product

If we are dealing with a product, we might need to be familiar with:

- Its characteristics and specifications
- Its CQAs (critical quality attributes), that is *physical, chemical, biological or microbiological properties or characteristics that should be within an appropriate limit, range, or distribution to ensure the desired product quality* "in [4]".

- The factors that might affect its quality
- The likelihood of these factors and the effects of the loss of quality

And this can only be achieved by people possessing or gathering this knowledge and analyzing it subsequently by means of brainstorming. A table like the following one can be helpful in performing this.

Problems (what can affect its quality?)	**Causes** (what might provoke this problem?)	**Likeness** (is that likely to happen?)	**Preventive measures** (what could we do to control or mitigate this problem?)
Particle size	Inadequate particle size	Yes	Provide the supplier with detailed specifications
Contamination (bioburden / pesticides)	Product of natural origin	Maybe	Tight control of the sources of product
Degradation	Too high temperature	Probable	Determine acceptable temperature range

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 1. Table for product brainstorming

Several approaches have been proposed to facilitate this brainstorming, such as,

- Imagine different situations (e.g. change of supplier, lack of electricity, flood, error, mix-up, etc.) and then answering to "what would happen"?
- Ask people working or having worked in this situation.
- Inquire regulatory authorities.
- Review technical literature and attend congresses.

2.2. Process

If we are dealing with a process we might need to be aware of:

- Its flow-chart;
- Its equipment;
- Its CPPs (critical process parameters), that is *process parameters whose variability has an impact on a critical quality attribute and therefore should be monitored or controlled to ensure the process produces the desired quality* "in [4]".
- The factors which can affect them and the ways for doing this;
- The likeness and the consequences of these deviations.

Exactly as said above, this has to be done by people possessing or gathering this knowledge.

In this case, the table might appear like this one.

The approach to be followed is the same that was mentioned before.

Process stage	**Problems** (Deviations from the normal situation?)	**Causes** (what might provoke this problem?)	**Likeness** (is that likely to happen?)	**Preventive measures** (what could we do to control or mitigate this problem?)
#1	Wrong weight of starting material	Error	Maybe	Scales will be provided with a printer
#3	Irregular dosage of vials because dose is very low	Filler has not enough capacity	Yes	Qualify filler and organize a monitoring system for filled vials
#8	Lack of leaflet	Leaflet fell outside the box	No	A weighing machine controls each box

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 2. Table for process brainstorming

2.3. Comparison

And finally, if we are comparing two different situations, then we should identify:

- Their characteristics (their main attributes and sub-attributes);
- The elements likely to influence them;
- Our needs / requirements.

Attributes	**Sub-attributes**	**Influence** (which role they may play?)	**Preventive measures** (what could we do to control or mitigate this problem?)
Premises	HVAC configuration	Risk of cross-contamination	Dedicated HVAC system without recycling
	Pressure differentials	Protection of the environment	Containment of the product by negative differential pressure surrounded by positive differential pressure
Personnel	Training	High risk of error / confusion	Ensure adequate training
	Protection	Manipulation of active products	Introduce isolators where product is exposed

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 3. Table for comparing two different situations / elements

In this case brainstorming can be facilitated by:

- Carefully analyzing the attributes and sub-attributes (e.g., the factors which intervene in the constitution of this situation / element).
- Ask people about their experience on this.
- Inquiring regulatory authorities.
- Reviewing technical literature and attending congresses.

3. GMP and QRM

An interesting question that might arise as a consequence of the application of QRM is about precedence. What is more important a GMP statement or the result of a QRA? This is however a completely false question because if GMP can be considered the logic baseline in pharmaceutical production, then how can QRA be anything different?

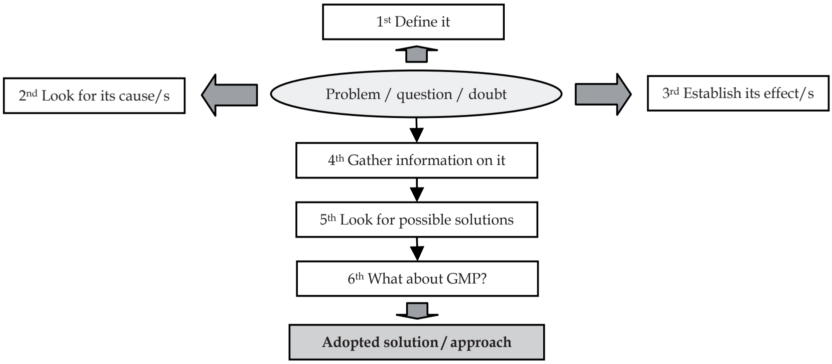


Figure 3. Problem management

In fact QRM is complementary to GMP because it provides the frame for taking a decision, whereas GMP proposes us some practical and well known solutions. It is important to understand that “good manufacturing practice” has been, and still is, accompanied by some amount of “bad practice”. And this is the result of seeing in GMP as a kind of oracle which will provide us with magical solutions.

No solution for a problem should be looked for, before having defined it perfectly. This assertion might seem surprising but experience shows that analysis is often left aside because all attention is eagerly focused on the quick search of a solution.

Once the problem is well understood and its causes and likely effects have been determined is when the search for a solution can be started. In any case this requires, as said before, possessing knowledge. It is evident that GMP has to be taken into account, but only when the problem has been analyzed and understood in depth. Then, an answer to the problem can be found and it is evident that it will be science and knowledge-based and GMP compliant.

4. Selection of tools

There are different tools which can be used in QRM, both unspecific and specific for this task “in [2, 6]”. In reality, and very unpleasantly for some people, tools just organize in a more or less sophisticated way the information that we have. They will neither provide us with the knowledge that we don’t have nor liberate us from the task of thinking about the question. Brainstorming will always be inevitable.

According to their function QRM tools can be classified in three main groups: risk analysis, risk comparison and statistical support.

Risk analysis tools are either inductive or deductive. The first, starting with the normal operation, try to detect possible problems. The second, starting with the problem, try to find the chain of events that led to it. It is also very common to talk about formal and informal tools. In fact the former have really been devised for this purpose, whereas the latter are just data given in a certain way and they can be only considered "tools" in a loose sense.

It might seem surprising to explain that there are QRM tools that do not consider risk, but hazard. Although this is discussed in more detail below, here we need only point out that "risk analysis or management" is a general concept which can consider just hazards or their associated risk too.

<table>
<tr><th colspan="5">Characteristics</th><th>Tools</th></tr>
<tr><td rowspan="8">Tools for risk analysis</td><td rowspan="7">Inductive / single factors</td><td colspan="2">Basic / Informal</td><td>Simple organization of data</td><td>Flowchart / Process map / Checklist/ Ishikawa diagram (fishbone), etc.</td></tr>
<tr><td rowspan="7">Formal</td><td rowspan="2">Identification of hazards and their potential effects</td><td>Risk is not estimated</td><td>PHA (Preliminary Hazard Analysis)</td></tr>
<tr><td>Risk is estimated</td><td>PRA (Preliminary Risk Analysis)</td></tr>
<tr><td rowspan="2">Evaluation of failure modes and of their potential effects</td><td>Risk is not estimated</td><td>FMEA (Failure Mode Effects Analysis)</td></tr>
<tr><td>Risk is estimated</td><td>FMECA (Failure Mode, Effects and Criticality Analysis)</td></tr>
<tr><td colspan="2">Evaluation and monitoring of hazard</td><td>HACCP (Hazard Analysis and Critical Control Points)</td></tr>
<tr><td colspan="2">Analysis of deviations by means of "key words"</td><td>HAZOP (Hazard Operability Analysis)</td></tr>
<tr><td>Deductive / multiple factors</td><td colspan="2">Identification of real or potential problems</td><td>FTA (Fault Tree Analysis)</td></tr>
<tr><td colspan="5">Tools for risk comparison</td><td>Risk Ranking and Filtering</td></tr>
<tr><td colspan="5">Supporting statistical tools</td><td>Control charts / Design of experiments / Histograms / Pareto charts / Process capability analysis</td></tr>
</table>

Table 4. Basic classification of tools used in QRM

Even if it is true that specific tools have been developed with an intended objective, and this somewhat restricts their scope of use, there is often overseen that they can have a wider utility. In fact tools are at our service and we should use them to organize information in order to get the most of it. Except in those cases when there is a need for coordination, e. g., different sites of the same group, or a requirement by the Authorities we should feel free in the way we use QRM tools.

In practice, it is possible to distinguish six basic cases in QRM:

1. As it has been indicated before the presentation of data is a basic need in QRM, hence informal tools, such as flow charts are necessary.
2. A first task in any QRM is hazard assessment and for doing this PHA is the right tool.
3. If risk assessment is desired, then FMECA is what we need.
4. HACCP is an appropriate tool for the monitoring of processes by means of their parameters.
5. When the search for the root cause of an event is required then FTA will do.
6. RRF allows for comparisons.

These six cases will cover practically all the needs regarding QRM. Understanding and using them can thus be considered a must.

QRM tool	Practical utilization
• Check lists, reports, graphs, etc. • Diagrams (flow, Pareto, Ishikawa), histograms, etc.	Basic data presentation
PHA (Primary hazard analysis)	Basic hazard assessment
FMECA (Failure Mode, Effects & Criticality Analysis)	Risk assessment
HACCP (Hazard Analysis and Critical Control Points)	Process monitoring
FTA (Fault Tree Analysis)	Root cause identification
RRF (Risk ranking and filtering)	Comparison

Table 5. QRM most common tools and their practical utilization

4.1. Basic data presentation

Any system may be used to gather and present data and further on we provide some examples. Ishikawa and Pareto diagrams not only show data but the first organizes them at a certain level and the second treats them statistically. This is why they are also mentioned for the root cause identification.

4.2. Basic hazard assessment

In every situation (product, process, etc.) there is a period of time, in the beginning, when knowledge is very limited and unsure. Suppositions count more than facts. Then, there is a very simple tool, perfectly adapted to this situation: PHA (Primary hazard analysis) "in [2, 6]".

It is developed by using a table, which might vary slightly according to specific needs or requirements, but which considers basically these items: hazards, causes, effects and preventive measures. Note that often it is not necessary to consider the "effects", as they are either evident or are already somewhat included in the "hazard".

The hazards which put the quality of the products at risk during their manufacturing may belong to five categories:

- Contamination (external): Any contamination of a material or of a product not related to other materials or products manufactured in the factory (e. g., pollen, sand, hair, scales, dandruff, fibers, microorganisms, etc.)
- Cross-contamination: Contamination of a material or of a product with another material or product.
- Environmental contamination: Contamination of the production rooms, the operators or the surroundings of the pharmaceutical unit because of the voluntary or accidental liberation of materials or products
- Mix up / error: Operation inadequately performed (error) or where one thing is taken by another one (mix up)
- Degradation: Loss of quality of the product because of inadequate conditions

The following table uses this approach:

Hazard	Cause	Preventive measures	Comments
Contamination (outer)	Inflow of dirty air	Separation of production areas and air control	Airlock doors are interlocked. Air is filtered.
Cross-contamination	Inadequate CIP cleaning	Vessel is dirty	Cleaning is validated.
	Dust liberation	Dust extraction system	Effectiveness of extraction is verified
Mix up / Error	Transfer of wrong material	Inadequate product	Double verification. Materials are controlled and registered while entering and while being used.
Environmental contamination	Isolator is not kept at depression	Room and people contaminated	Isolator is provided with a pressure gauge. Isolator has an alarm system for pressure.
Degradation	Inadequate temperature	Air conditioning	Temperature is controlled
	Inadequate humidity	Air conditioning with control of humidity	Relative humidity is kept ≤ 30%

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 6. PHA table

4.3. Risk assessment

In order to analyze quality risks in the operations, when there is a significant amount of operator participation, the choice tool is FMECA (failure mode, effects and criticality analysis) "in [2, 6, 7]". It is performed by using a table possessing these main headings:

- Failure mode: the way an element can have a potential failure (in relation to specifications) or do not develop its functions. They are detected as an answer to the question "what might go wrong? It has to be pointed out that a simple function may have several failure modes.

- Cause: the grounds which provoke a failure.
- Effect: the results which appear when the failure mode comes out.

#	**Operation / Process stage**	**Failure mode**	**S** (Severity)	**Cause**	**P** (Probability)	**Existing controls**	**D** (Difficulty of detection)	Risk priorisation **(PR = S x P x D)**	**Risk accepted? (comments)**
3	Agitation	Speed < 1.400 rpm	2	Failure	1	Monitoring by computer	1	2	Yes (a tachymeter is in place; speed can vary without sensible effect)
5	Dissolution	Degradation	3	Temperature > 80ºC	1	Monitoring by computer	1	3	Yes (T is controlled and materials are not affected)
9	Filling	Inadequate dosage	3	Failure / wrong adjustment	1	Every tube is weighed after filling	1	3	Yes (equipment is qualified and scales are calibrated)

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 7. FMECA table

This table may include risk reduction and be used as an instrument of risk management.

Risk evaluation										Risk reduction					
#	**Operation / Process stage**	**Failure mode**	**S**	**Cause**	**P**	**Control in place**	**D**	**RP**	**Risk accepted?**	**Actions**	**Reevaluation**				**Risk accepted?**
											S	**P**	**D**	**PR**	

Table 8. FMECA table for risk reduction

4.4. Process monitoring

It is evident that if a process is well understood it is possible to identify its CPPs and if we can keep them under control by a process monitoring system then the quality of the products will be ensured.

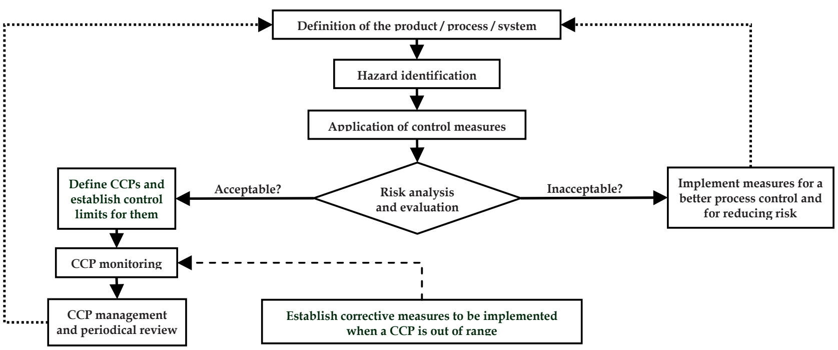


Figure 4. HACCP rationale

HACCP (hazard analysis and critical control points) is a method that detects the hazards for the quality of the products (or for the safety) and then their "critical control points" (CCP) "in [2, 6, 8]". The rational of HACCP is exposed in the annexed figure.

The flow chart is studied to identify potential hazards, which might affect the quality of the product. Then these hazards have to be assessed. Do they have to be controlled?

		Seriousness of harm				
		Insignificant	Minor	Severe	Critical	Catastrophic
Probability of harm	Always					High risk
	Frequently (probable)					
	Sometimes			Medium risk		
	Rare (improbable)					
	Non observable	Low risk				

Table 9. Example of risk evaluation

This leads to the determination of the CCPs.

Operation / Process stage	**Potential hazard**	**Is risk significant?**	**Why?**	**Preventive measures**	**Is it a CCP?**
Test for endotoxin in water	Presence of endotoxin in water	Yes	Endotoxin is not allowed	Monitoring	Yes
Measure of pH	pH outside range	Yes	Precipitation	Verify pH	Yes
Filter sterilization	Viable microbes in the filter	Yes	Product not sterile	Process validation and monitoring	Yes

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 10. HACCP – I: CCP establishment

The rationale for the establishment of CCP can be summarized as follows.

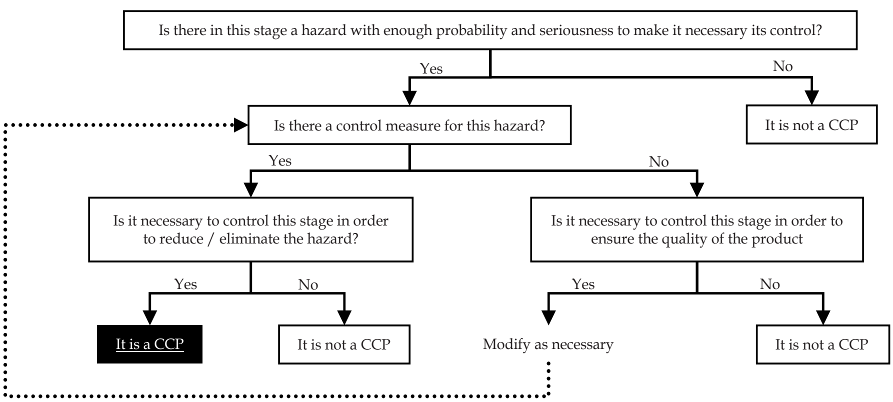


Figure 5. Rationale for the establishment of a CCP

Then, for each CCP are defined alert and acceptance limits. As a consequence, monitoring procedures are established and they are accompanied by the set up of corrective actions in case of deviation. Finally as in any monitoring system it is necessary to define how it will be managed and periodically reviewed to verify that it performs as expected.

CCP	Acceptable range	Monitoring			Corrective actions
		Who?	How?	When	
Test for endotoxin in water	< 0,25 U.	QC technician	LAL test	Before starting the production	Stop production and call the supervisor
Measure of pH	pH = 6-7	Production supervisor	pH-meter	In process	Call the supervisor. Add more sodium hydroxide
Filter sterilization	Sterile	Production technician	Control process parameters (P, T and t)	After sterilization	Stop production and call the supervisor

Note: The contents of this table are just given as an example; they don't intend to represent any real situation

Table 11. HACCP – II: CCP monitoring

4.5. Root cause identification

When a deviation is detected it is necessary to implement corrective and preventive actions (CAPA system), But this is only possible if the root cause has been identified. To do this we can use several tools, both unspecific (such as Ishikawa or Pareto diagrams) and specific (FTA).

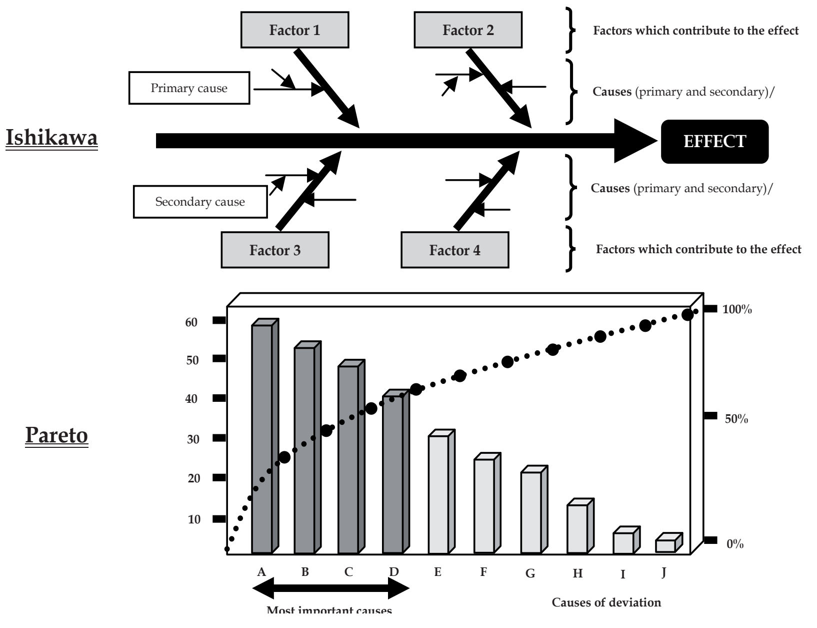


Figure 6. Diagrams of Ishikawa and Pareto

Two simple tools can be used for a primary analysis of causes.

The first one is the diagram of Ishikawa, also known as cause-effect or fishbone diagram, which shows in a graphical way cause relations and their interaction to provoke an effect.

The second one is the diagram of Pareto, which orders the data in relation to their importance and this allows for the distinction among frequent and infrequent causes of failures. It is prepared by listing all the elements and determining their frequencies. Then, the elements are classified in relation to their cumulative frequency.

FTA (fault tree analysis) is a deductive tool which uses a pictogram to represent in an organized way the factors (causes) which produce or contribute to the production of an undesirable event "in [2, 6, 9]".

The tree is started by placing the top event. Then the events which contributed to it are analyzed. Events are united by gates, which show the relation amongst them.

Gates can be very varied, but the most common ones are the "and" and "or" gates. Although there are specific gate symbols defined by an international standard, in practice for most cases it would suffice to represent them by a circle and write inside to which type they belong.

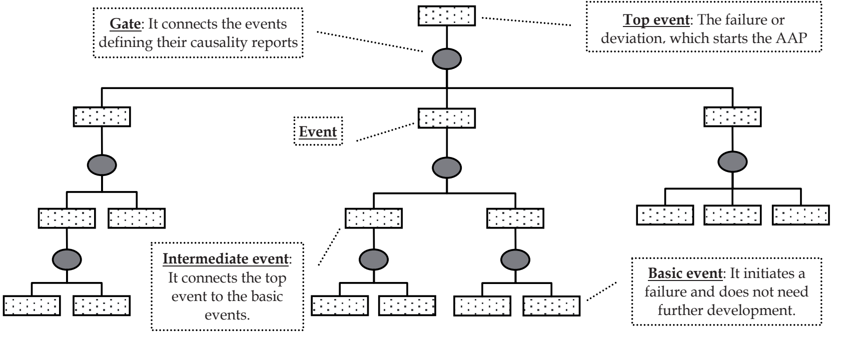


Figure 7. FTA diagram

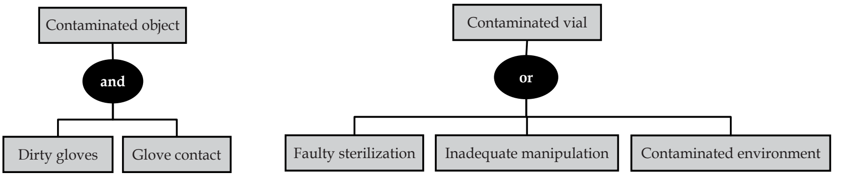


Figure 8. FTA: Examples of "and" and "or" gates

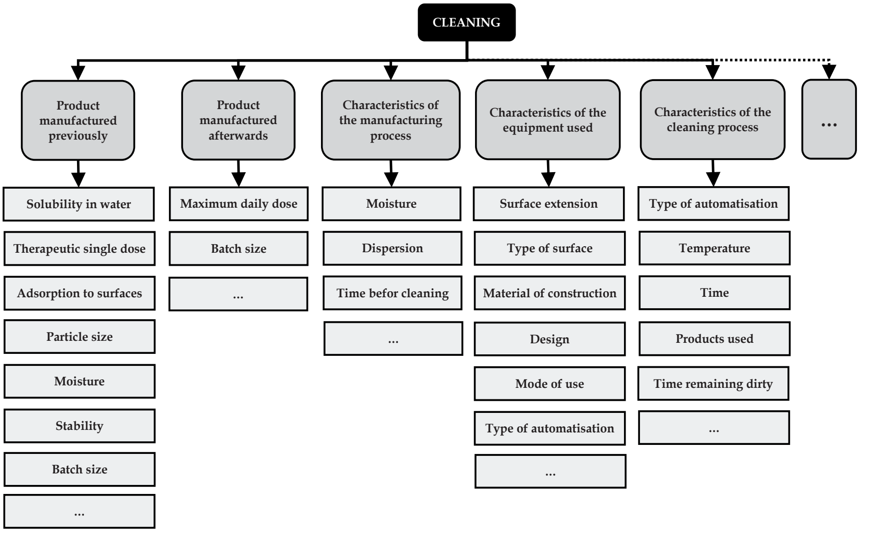


Figure 9. RRF: Example of hazard or problem decomposition (factors intervening in cleaning)

4.6. Comparison

RRF (Risk ranking and filtering) is a tool specifically devised for the comparison of different sets (units, processes, companies, etc.) possessing varied levels of risks “in [2, 6, 10]”. Once they are reduced to a common denominator they can be compared and this allows for the establishment of priorities.

As in any method, it is necessary to individualize first the hazards or problems and then the different attributes, components or elements, which contribute to them.

Then each attribute, component or element is evaluated in terms of risk.

Problem / Hazard				
	Classification			Evaluation
	Low	Middle	High	

Supplier compliance (example)				
Element	Classification			Evaluation
	Low (1)	Middle (2)	High (3)	
GMP	Deficient	Acceptable	Certified	2
Q system	Deficient	Acceptable	Certified	1
Audits	None	By other	By us	3
Documents	Deficient	Good	Excellent	2
History	>5 problems	<5 problems	No problems	2
				10

Figure 10. RRF: Risk estimation

It is possible to get a comprehensive evaluation taking into account all the intervening elements. This allows for the ranking of the problems or hazards, which can be filtered.

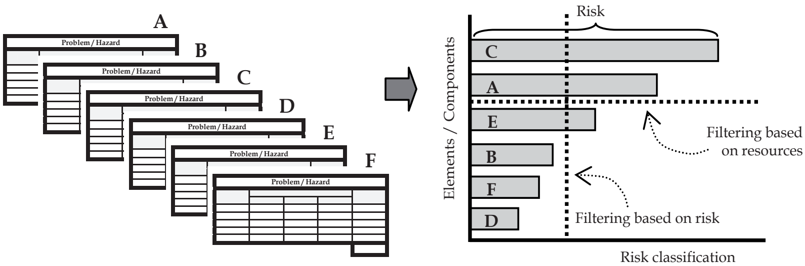


Figure 11. Hazard – Risk - Harm

It is also possible to rank and filter risks by using tables combining the three well-known factors (probability, severity and difficulty of detection).

Risk classification		Severity		
		High	Middle	Low
Probability	High			Class 1
	Middle		Class 2	
	Low	Class 3		

Risk filtration		Difficulty of detection		
		Low	Middle	High
Risk classification	Class 1			1st priority
	Class 2		2nd priority	
	Class 3	3rd priority		

Figure 12. RRF: Combination of the risk factors

5. Is it really necessary to determine risk in QRM?

This is not the kind of stupid question that, at first sight, might seem. And in fact this is not a question, but two. The first one might be related to the fact that in everyday's life we somehow tend to mix-up hazard and risk. As the latter is the consequence of the former, we tend to consider both practically as synonymous. The second one comes out because of the fact that identification of hazards is a prerequisite for the determination of their risk, and risk allows for an assessment of hazard. Thus, why to limit ourselves to something of "low level" like hazard when we can get something "better" like risk? Unfortunately this is not correct. By definition, to determine risk we have to start by knowing the probability of occurrence of the hazard. And this is often very difficult and in the end it turns to be just an inference. Then, we should know the seriousness of the harm and, although this is usually clearer, it is neither an easy task and often requires some degree of imagination. Consequently, in many cases, risk is not more than an estimation, that has to be improved along the time as more experience is collected. The case might be that hazard, usually a concrete thing, is substituted by risk, an estimated value and this can hardly be something better.

In practice, our first aim should be to determine hazard and then, only if there is an objective possibility of estimating risk, do it. Speaking in general terms the evolution should be hazard detection > qualitative risk estimation > quantitative risk estimation. In projects or new processes we would move towards the left (hazard), whereas as we gain process knowledge we might move towards the right (risk).

Summarizing, although we talk about QRM it is perfectly acceptable just to determine hazard and manage it in terms of making the appearance of harm an unluckily event by applying corrective measures. It is better to limit a QRA to the hazards that trying to estimate risk without having enough information for doing it.

6. How to determine risk?

Risk is determined by the combination of the two already mentioned factors, probability and severity, to which a third one, detection, can be added when a system of detection is in place. This gives us the classical formula:

Risk = Probability of occurrence of the harm x Severity of the harm x Detection of the harm.

Instead of probability it is often used the term frequency, pointing out that most of times what we really know is how often it happened in a well established process. It is evident that even if we don't know how to estimate probability, if we really do know that harm never happened in our process we can affirm that probability is very low.

Severity is easier to understand because we are only asked to assess the importance of harm.

The capacity of detection of harm is linked to the existence of a system for its detection. Thus, its assessment tends to be more objective, as it is related to the equipment. This factor, however, has a marked particularity: risk increases when the capacity of detection decreases; it is an inverse factor. This is not a problem if we bear in mind this fact, but it can be easily overcome by changing the way we express it, for instance, instead of talking about "detection of the harm" we could say "difficulty of detection of the harm", thus turning it a direct factor like the other two.

7. How to better evaluate risk?

All three risk factors can be evaluated either qualitatively or quantitatively. Again, a quantitative estimation of risk (e.g. of 45 over 50, say) might appear much more satisfactory than a qualitative one (e.g. middle), but this might be too, and often is, misleading, because it provides a false sensation of precision.

As it was discussed above, the main objection one can do regarding risk is that its determination is too subjective and this makes it unreliable. Although there is some amount of truth in this, we may, however, counterattack by explaining that the objective of assessing risk is not getting a faithful estimate of it, but obtaining a risk baseline to be used as an indicator for future improvement. This is why what really matters is providing a good deal of information on the rationale which led to the estimation of the factors of risk.

Independently from that, it is necessary to determine how many risk levels will be estimated for each hazard. The simplest case and maybe the commonest too is the utilization of three levels, a very intuitive approach, as we are talking about "a lot" (high), "medium" (middle) and "little" (low). More levels allow for a better classification, but they turn more complicate the assessment.

Another point that needs to be discussed, when using quantitative evaluation, is the relation between these levels. Beside the normal series of values (1, 2, 3, etc.), we might use other with wider gaps between the levels, either regular (e. g., 1, 3, 5, 7) or even irregular (e. g., 1, 3, 7, 12). These last cases would give more weight to the higher levels, thus increasing the sensation of higher risk. It is not necessary to insist on the fact that any type of characterization of the levels is acceptable, provided that it is clearly defined and indicated and that it is not used to take false conclusions regarding higher evaluations of risk.

7.1. Qualitative evaluation

For instance in three levels:

Evaluation	PROBABILITY	SEVERITY	DIFFICULTY OF DETECTION
High	The failure /accident occurs **frequently**	The consequences of the failure /accident are **important**	The failure /accident **will very likely not be detected**
Medium	The failure /accident occurs **periodically**	The consequences of the failure /accident are **moderate**	The failure /accident **might be detected**
Low	The failure /accident occurs **rarely**	The consequences of the failure /accident are **low**	The failure /accident **will very likely be detected**

Table 12. Example of qualitative evaluation in three levels

More than three levels:

Evaluation	PROBABILITY	SEVERITY	DIFFICULTY OF DETECTION
Very high	Always	Catastrophic	It cannot be detected
High	Often (probable)	Critical	It can only be detected when the process is already finished
Medium	Sometimes	Serious	It can be detected during one of the stages of the process
Low	Rare (improbable)	Minor	It can be detected during the stage in process
Very low	Non observable	Insignificant	It can be detected instantaneously

Table 13. Example of qualitative evaluation in five levels

An important inconvenient of the qualitative evaluation of the factors appears when determining risk. Everybody would agree that low x medium x high = medium, but the score of low x low x high is not so evident (medium?). And this becomes even more unclear if we consider more than three levels.

This is however less important that it might seem at first sight. As we have seen above, risk estimation, and particularly the qualitative one, is likely to be more a rough approximation than a very exact value. Thus, adding some more roughness should not be considered critical. In any case, this inconvenient can be overcome by using a complementary table which would provide a homogeneous estimation.

Risk = Low x Low x Low = Low	Risk = Medium x Medium x Medium = Medium
Risk = Low x Low x Medium = Low-Medium	Risk = Medium x Medium x High = Medium-High
Risk = Low x Medium x Medium = Medium-Low	Risk = Medium x High x High = High-Medium
Risk = Low x Medium x High = Medium	Risk = High x High x High = High
Risk = Low x High x High = Medium-High	----

Table 14. Example of table for the qualitative estimation of risk

It is evident that if the risk factors are estimated by using more than three levels, then the table becomes more complicated.

7.2. Quantitative evaluation

If it is done in three levels, then the approach might be exactly the same that shown above for the qualitative estimation, but instead of low, medium and high there will be used numbers (1, 2, 3) and this will facilitate calculation (e. g., 3 x 1 x 2 = 6).

The next table provides an example of quantitative evaluation in five levels, both following the natural series of numbers and an irregular (or "enhanced") one. Both are, of course, acceptable, but it is necessary to bear in mind that the final risk quantification will depend on the system which has been chosen and thus comparisons have to take this into account.

Evaluation (two options)		PROBABILITY	SEVERITY	DIFFICULTY OF DETECTION
1st	2nd			
1	1	Expected >80% of times.	Batch is OOS and it is rejected.	No detection.
2	3	Expected between >50% and ≤80% of times	There are deviations and batch is investigated and rejected.	Detection but only when the process is finished
3	5	Expected between >10% and ≤50% of times.	There are deviations and batch is investigated but accepted.	Detection during a stage of the process, before finishing it.
4)	7	Expected between >1% and ≤10% of times.	A trend is detected, but limits are not exceeded and the batch is not investigated.	Detection during the stage in process.
5	9	Expected ≤1% of times.	There is no trend and limits are not exceeded. Batch is not investigated.	Instantaneous detection.

Table 15. Example of quantitative evaluation of the factors of risk in five levels

8. How to introduce QRM in a company?

One might tend to think that what matters is to teach people how to use QRA tools and how to evaluate risk. Although this is something that has to be done, it is necessary to bear in mind that this is only second to the understanding of the purpose and practical utility of QRM.

QRM is not just a kind of new task in a company. It is, in fact, a new way of looking at things, a new approach in analyzing the problems and in proposing solutions to them. This is why the most important and basic task is to make up the mind of people.

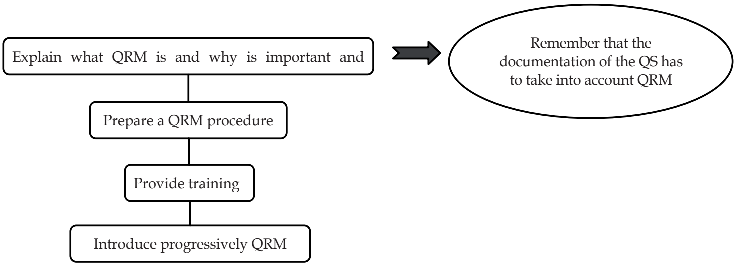


Figure 13. Introduction of QRM in a company

To simplify matters it is generally agreed that a good practice is to prepare a detailed protocol describing how the company intends to develop and to apply QRM.

9. Is there a way to facilitate QRA?

Everybody can have a personal approach towards QRA, but experience shows that there are no simple ways. As said above a good knowledge of the subject is essential, then some dose of experience is helpful and a big amount of patience is necessary.

It has already been described the paramount importance of getting information on the subject and how this can be done. It has also been explained how to choose and use the more common (and practical) tools. Thus, getting an adequate QRA is just a question of work. Get a first draft, review it and improve it little by little until getting something satisfactory. There are no rules, but the approach of working by progressive "retouches" is certainly very appropriate.

It is worth to mention that while working on QRA one of the most puzzling aspects is the frequent confusion that one tends to experience regarding "hazard", "cause" and "effect". This can be overcome in two complementary ways. Firstly by understanding well the meaning of these terms and applying them carefully to the elements being analyzed and secondly by reviewing the draft after a few hours of rest. Normally what is not seen clear now, it will be later.

10. Why QRA is worthy?

QRA provides us with a systematic and deep knowledge of the problems faced by the process under study. In this sense QRA can help us to overcome one of the main troubles derived of the application of GMP, their being seen as a kind of "tables of the law". The consideration of GMP as the golden principles supposed to provide the answer for every problem, has often led to a passive attitude. We had a problem and then we looked up in

GMP. Now, QRA means that we analyze and understand well our problem. It is not necessary to say that this is extremely important and worthy.

QRA is also a key element in the introduction of improvement in the pharmaceutical industry. Without the concept of QRA, improvement would not be understood. In the past improvement didn't exist, because we didn't accept risk. Now when we recognize risk and accept that it is inevitable we can also say that improvement exists and that it consists of the reduction of this risk.

11. Some common problems solved by using QRA

QRA is a powerful instrument, which may help to solve typical problems faced by the technicians working in a pharmaceutical laboratory and in this sense the ICH Q9 guideline provides many examples regarding the potential applications of QRA "in [4]". It is however worthwhile to study in detail three very common problems: documentation, inspections/audits and handling of active products.

11.1. Documentation

"Which documents do I need?" or "tell me which documents I have to prepare and I will do it" are much heard questions. And this is so because GMP does not provide an answer. An impossible answer, anyway, because documentation is closely related to every particular situation and this is why GMP just mentions the documents directly associated with the products (specifications, formulae, processing instructions, etc.) and some general documents (sampling, testing, release, etc.). Documents have to be established on a case by case base, even if in the end all laboratories finish by having about the same documents and the basic difference relies on their extension and on how information is organized and grouped.

Documentation is used to attain two basic goals, information on how to perform operations, in general, and confidence in the control of critical operations, in particular. The first determines basically the number of documents to be written, whereas the second focuses on their contents.

Write a list of the processes of your company and then prepare flow charts for each of them. By doing this a number of operations will appear. Procedures have to describe these operations. There is no rule of thumb about how many are needed. Neither very long nor very short procedures are practical. Consequently a certain level of grouping is necessary and each company has to decide how to do it. Factors which might help in this decision are basically related to logistics, organization and personnel.

Once you have these flow charts perform on them QRA using a very basic tool such as PHA. This will allow detecting hazards and will oblige thinking of the related control measures. And then, these control measures will have to appear in the documents describing the operations.

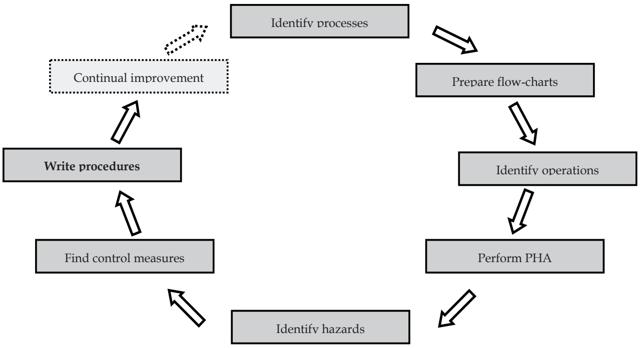


Figure 14. The documentation virtuous circle

11.2. Inspections / Audits

Inspections and audits worry always personnel. They remind them of the student days and consequently they ask themselves again "how can I pass them successfully"? In the school the obvious answer would have been "learn well the lessons of the program", but in the pharmaceutical industry the program is less obvious and "lessons" are not so clearly stated. They have to take into account GMP, of course, but also other things (GLP, regulations, guidelines, unwritten expectations, company standards, etc.). Here again QRA can help us. As summarized in the figure below our approach should focus in showing that we know our hazards and we keep them mastered.

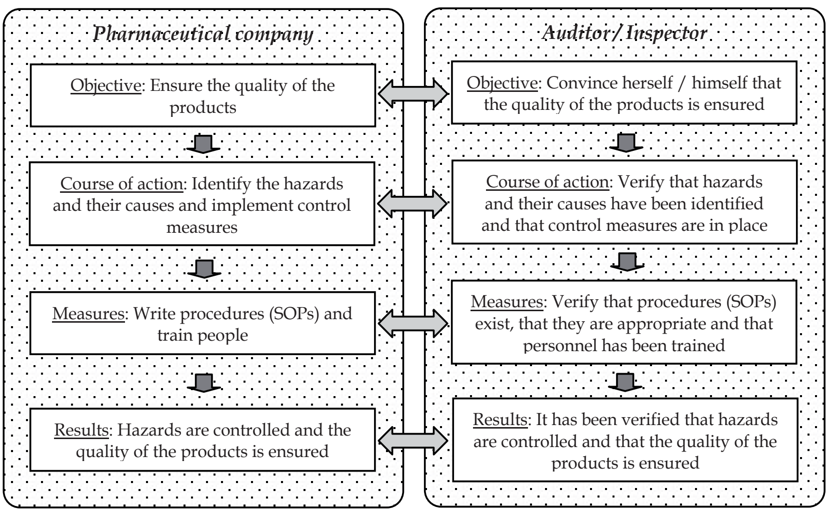


Figure 15. How to face an inspection/audit

11.3. Handling of active products

GMP establishes the need for dedicated facilities in some cases (production of penicillin or live microorganisms) or requires separation for certain products (antibiotics, hormones, cytotoxics, etc.). Most products are however manufactured in multiproduct facilities, where products share utilities and equipment, after performing the required validations. As products and processes can be very varied it is evident that the orientation provided by GMP can't be very detailed. This is why a QRA approach is very useful.

Among both extremes, a dedicated building (with separated equipment and utilities) and a multiproduct facility, there are less radical solutions, such as, for instance, separated areas (rooms or set of rooms with their own air-locks and changing rooms), provided with separate equipment and utilities, within a multiproduct unit. Also, when this is feasible, a simpler way of preventing cross-contamination might be the use of specific parts of equipment (e.g., filters, sieves), instead of dedicated equipment.

Another approach would be segregating the process from its environment by using isolation technology (i.e. production equipment within isolators or closed equipment instead of separating the rooms).

Campaign manufacturing, which certainly requires appropriate validation, is also a way to prevent cross-contamination.

The application of quality risk assessment to the products and to the processes to be performed within a pharmaceutical unit, allows for their rational design.

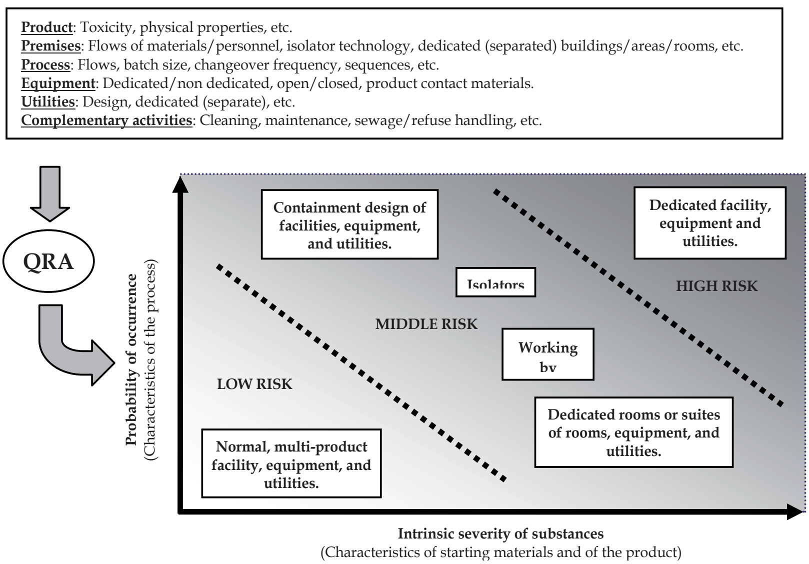


Figure 16. Handling of active products: QRA approach

The appraisal of the hazard can be performed by using primary hazard analysis (PHA) as it was previously described. In this case from the five categories of hazards mentioned only two have to be taken into account: environmental contamination and cross-contamination.

Hazard	Cause	Preventive measure	Comments
Environmental contamination	Release of dust, particles, droplets, aerosols, effluents and waste	Manufacturing areas are isolated	Clean-rooms are isolated from the environment
		Products are isolated	Product containment inside closed systems
		Air handling	HEPA-filtered air
		Pressure differentials	Negative pressure impedes the diffusion of particles
		Effluent treatment	Cleaning effluents are treated before being released to the sewer
		Waste control	Wastes are disposed of in an adequate manner
		Control of dust	Release of dust is kept under control (closed systems / dust extraction)
			Personal protective equipment (PPE)

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 16. Handling of active products – QRA – Environmental contamination

Although in both cases the causes are similar, the array of control measures is very varied and concerns different systems.

Hazard	Cause	Preventive measure	Comments
Cross-contamination	Release of dust, droplets, particles, aerosols, during operations	Manufacturing areas are separated	Different pharmaceutical forms are completely separated
		Products are physically separated	Only one product at a time
		Environments are separated	HVAC system provides environment cleanliness (classification) and separation (pressure differentials)
		Access of personnel across specific changing rooms	There is an specific change room for each area
		Transfers across air-locks	All materials and products are transferred across air-locks
		Air treatment	HEPA-filtered air
		Differential pressure / Airflows	HVAC system provides environment cleanliness (classification) and separation (pressure differentials)
		Use of closed systems	Production is performed in closed systems.
		Temporal separation of products	Campaign working allows for a simplification in cleaning.
	Residues on equipment	"Sanitary" design and construction	Clean-rooms follow GMP-design
		Cleaning / sanitation	Validated cleaning procedures
	Dust / particles on clothing	Use of specific clothing	Clothing is related to the type of operation, which is carried out

Note: The contents of this table are just given as an example; they don't intend to represent any real situation.

Table 17. Handling of active products – QRA – Cross-contamination

Author details

Jordi Botet
JBF - GMP Consultant, Spain

12. References

[1] FDA (2004) Pharmaceutical CGMPS for the 21st century - A risk-based approach. Final report.
[2] ICH (2005) Quality risk management. ICH Harmonised Tripartite GuidelineGuideline Q9.
[3] ICH (2008) Pharmaceutical quality system. ICH Harmonised Tripartite GuidelineQ10.
[4] ICH (2009) Pharmaceutical development. ICH Harmonised Tripartite GuidelineQ8 (R2).
[5] ICH (2011). Development and Manufacture of Drug Substances (Chemic al Entities and Biotechnological/Biological Entities). Draft Consensus Guideline Q11.
[6] Vesper, J. L. (2006) Risk assessment and risk management in the pharmaceutical industry. Clear and simple. PDA/DHI.
[7] CEI/*IEC (1985)* Techniques d'analyse de la fiabilité des systèmes. Procédure d'analyse des modes de défaillance et de leurs effets (AMDE) / Analysis techniques for system reliability. Procedure for failure mode and effects analysis (FMEA). Norme internationale / International standard 608121985.
[8] WHO (2003) Hazard and risk analysis in pharmaceutical products. Technical Report Series, No. 908. Annex 7.
[9] CEI/IEC (1990) Analyse par arbre de panne (AAP) / Fault tree analysis (FTA). Norme internationale / International standard 1025
[10] FDA (2004) Risk-Based Method for Prioritizing CGMP Inspections of Pharmaceutical Manufacturing Sites — A Pilot Risk Ranking Model.

Chapter 18

Project and Enterprise Risk Management at the California Department of Transportation

Pedro Maria-Sanchez

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51442

1. Introduction

A better understanding of risk management processes and practices within a government agency is crucial for enhancing the project delivery process and for implementing formally risk management. This chapter outlines the whole implementation process carried out with the risk management team formed from different functional units and backgrounds. In addition, a discussion is held over the critical steps and aspects for performing project and enterprise risk management in the real world.

Risk management is not new for the transportation industry in the United States, specifically in highway projects. The California Department of Transportation (Caltrans), a leading authority in public transportation projects in the US, has used basic project management principles along its statewide Districts offices. Risk management has been part of the project management menu; nevertheless its application was limited only to developing a risk register and a qualitative analysis at the most. The Office of Statewide Project Management Improvement (OSPM), has developed a Project Risk Management handbook which is a guide for project managers at Caltrans for using risk management. Unfortunately, the latest version of the manual which is from 2007, did not included a detail explanation of the benefits for performing quantitative risk analysis while determining the risks impacts into the project objectives, in terms of cost and time. The term quantitative risk analysis is merely described, lacking a sound description of the tools and methodologies that have been in place and use in the industry for many years and even with other government agencies around the world. Cost overruns caused by a lack of using risk management in the practice for infrastructure and transportation projects, has been mentioned in the literature for many years. However, only few examples of how risk management can be use in the real life are available, including how can a risk team be formed and how to educate the team for performing a sound and trusted risk management exercise.

2. Risk management planning

As any other process in project management, risk management has to be planned in order to forecast the total effort required by the project team for developing the full scope of risk management. The California Department of Transportation (Caltrans), developed a Project Risk Management Handbook which is being used as a reference for planning the steps for applying risk management into specific projects. The purpose of the study, roles and responsibilities, the scope of the Risk Register, risk identification, analysis methods, implementation period, schedule and budget allocation need to be defined with the plan. Special attention should be placed into the human resource (method selection) aspect and in identifying the right phase of the project to initiate with the study. The roles of the Project Manager (PM) and the Risk Manager (RM) are critical for developing a realistic implementation plan (Figure 1). In addition, before starting working wit the RMT, the PM and RM should ensure that important project data is available. For example the project report, cost estimate, project plan, etc.

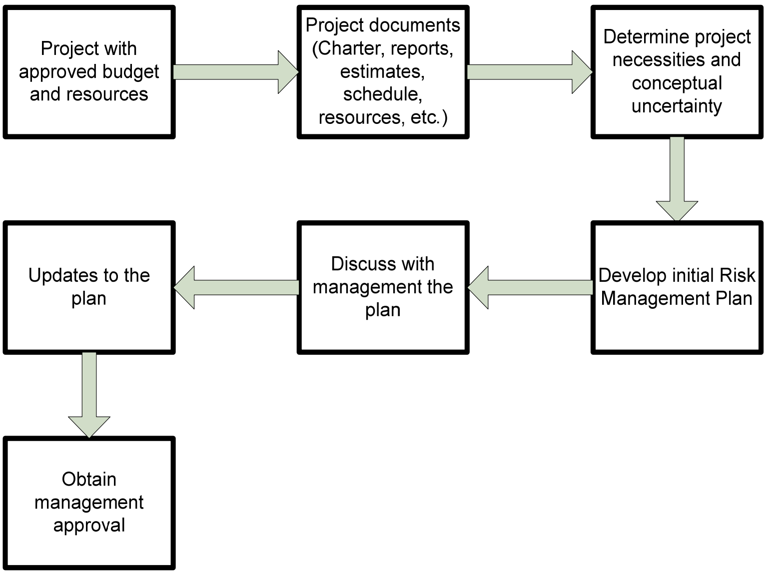


Figure 1. Planning the risk management plan

It is ideal to have the project charter for developing the risk management plan, since in the charter it is possible to identify critical information about the project like scope, conceptual cost estimate, delivery milestones, conceptual risks, stakeholders, etc.

3. Risk management training and education

Caltrans's Project Risk Management Handbook (California Department of Transportation [Caltrans], 2007) is the reference for performing training to the project risk management team. It covers the basics of risk management applied to transportation projects. Nevertheless, additional knowledge is provided to the team members for assessing properly the risks and opportunities during the qualitative and quantitative analysis. Although there are considerable resources for learning about risk management, Caltrans has adopted this process into its project development process (Figure 2). This approach has assisted for providing on the job training for the Project Development Team (PDT).

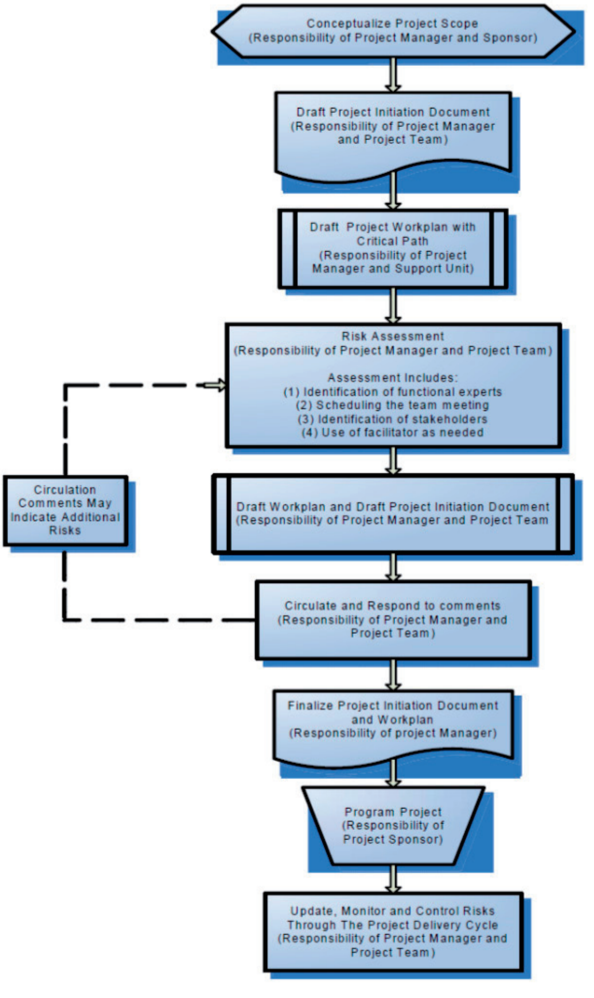


Figure 2. Risk management process flow diagram (Caltrans, 2007)

It is important to notice, that the risk assessment is responsibility of the PM and the project team. Nevertheless, it is recommended to use whenever is possible a RM. The RM is a neutral element of the project team and can reduce the bias, which can seriously affect the outcome of the risk management study.

4. Risk manager role

The relevance that a risk manager plays during the implementation phase is crucial for the success of the study. The RM as a risk expert should be able to lead, coordinate, educate, explain, convince, propose, monitor and evaluate the entire process; plus he or she needs to be able to have experience in leading teams from different backgrounds and coming from different functional units and agencies. Vose (2008) enounces the following as the characteristics of the risk analysts: creative thinkers, confident, modest, thick-skinned, communicators, pragmatic, able to conceptualize, curious, good at mathematics, a feel for numbers, finishers, cynical, pedantic, careful, social and neutral. The reality is that we not always can find individuals that have all the virtues mentioned before, therefore; we need to select the most critical characteristics that a risk manager should have. Definitely a risk manager should be a good communicator, must have an analytical mind and needs to be able to think outside the box. The skills of a risk manager are somehow related to the project manager's, in the sense of managing and controlling. However, the risk manager needs to deal with risk assessment that in the quantitative arena requires analytical modelling skills that the project manager is usually not trained for.

It is a demanding list and indicates; that risk analysis should be performed by people who have a proven track record of doing risk management for several projects ideally. It is also rather unlikely to find these skills in one person: the best risk analyst units with which we work are composed of a number of individuals with complementary skills and strengths Vose (2008).

The Washington State Department of Transportation ([WSDOT], 2010) in their Guidelines for CRA-CEVP Workshops, enlisted some of the risk manager responsibility: risk elicitation, risk modelling, cost validation/review, lead meetings, and provides reports and Develops or implements workshops on topics such as project definition, and risk identification and management. The California Department of Transportation (Caltrans, 2007) describes the risk manager (Risk Officer) responsibilities as risk management planning, identification, qualitative and quantitative analysis, risk response and risk monitoring and control. See Figure 3.

It is very important to take into consideration that the person selected or named to become the risk manager, should be someone that in addition of having the minimum capabilities, should be a neutral team element that can guide the risk team towards a non biased risk assessment of the project.

Although the position of RM in public agencies and private companies is rather new; currently the importance of having such an expert formally within the organization structure has become more formal and demanded. It has been realized that the benefits of having such an expert are much higher than the losses caused by not properly assessing and managing risks.

Process Tasks	Role					
	Sponsor	Deputy District Director, Program and Project Management	Project Manager	Project Manager Support/ Risk Officer	Project Team	Risk Owner
Risk management planning	C	C	R, A	S	S	
Risk identification	C	C	R	S	S	
Qualitative risk analysis			R	S	S	
Quantitative risk analysis (As applicable)			R	S	S	
Risk response planning	C	C	R, A	S	S	R
Risk monitoring and control			R	S	S	R

Figure 3. Risk management key responsibilities (Caltrans, 2007)

5. Project stakeholders

Due to the nature of the transportation projects, several stakeholders are usually involved. In particular, there is a closed communication between other local, regional and federal agencies which need to be involved and provide assessments and feedback. Caltrans as a state agency, deal internally as well with several stakeholders coming from the different functional units or divisions.

For the above reasons, it is extremely important to develop project communication plans, so the right stakeholders can be identified, together with investigating the best ways of communication with them critical project information for decision making. Figure 4 shows a typical stakeholder analysis used at Caltrans, which is part of the project communication plan.

As can be seen, the stakeholder analysis aid in selecting those stakeholders that could play a relevant role in the project. For the PM, is important to know which is the prefer way for communicating with each stakeholder, since he needs to keep them informed at all times. The frequency is another factor which needs to be addressed carefully, some people like to be informed about every change or movement in the project and others, only want to receive updates related to considerable changes in scope, time or cost for example.

6. Subject matter experts

The Subject Matter Experts (SMEs) have two objectives along the risk management study; firstly they provide specialized knowledge about specific project risks and uncertainties and

Name	Function	Contact Information	Preferred Method of Communication	Goals on this project	Frequency
Project Sponsor	Progress/Status Report	email	email	Budget	Monthly
Project Manager	Progress/Status Meetings and Report. Information sharing and issue resolution	email	email	Time, cost, quality	Monthly
Project Manager Assistance	Progress/Status Meetings and Report. Information sharing and issue resolution	email	email	Time, cost, quality	As occurs

Table 1. Stakeholder analysis (Caltrans, 2007)

secondly; in lieu of having a set of data available, they can fill the gaps of insufficient data while defining the uncertainty of a project variable. In practice and with transportation projects, usually two or three variable values are asked to the SMEs. For other type of projects or industry this approach may be different for example if the project is very unique and complex, like an energy plant or even an oil platform.

If for example the goal is to work with two values for defining the uncertainty behaviour for a particular risk variable, then the probability distribution called Uniform Distribution is used, which assumes all the values between the minimum and the maximum values have the same probability of occurring. In this case, a worst and best case scenarios are asked from the SMEs for defining a minimum and a maximum value of the risk. The other common probability distribution function use in this field is the triangular distribution function. This function assumes that there are three different values assigned to a variable of risk; a minimum, a most likely and a maximum. It is difficult to define which function is better for obtaining the data from the SMEs, since it will depend of their experience, the type of project and the project data available at the moment of the study. In general, the uniform distribution is the most popular since it is rather easy to estimate only two values.

A key part for soliciting the information for the SMEs, are the questions asked by the Risk Manager or Risk Facilitator. In some cases, the SMEs are reluctant to participate and optional methods should be place on the table by the risk manager for getting the opinion needed from the SME. An alternative and perhaps a highly recommended approach is to have one on one interviews with the SMEs for asking the questions and getting the answers. However, the recommendation is to have this information during the meetings, so the rest of the team can discuss and in some cases even do they can challenge the SMEs opinion, which is a very productive and healthy action for the project. The downside of conducting these interviews is that they are rather time consuming.

In some cases, we can get different SMEs opinions for the same variable or risk as mentioned by Vose (2008). Experts will sometimes produce profoundly different probability distribution estimates of a parameter. This is usually because the experts have estimated different things, made differing assumptions or have different sets of information on which to base their opinion. However, occasionally two or more experts simply genuinely disagree.

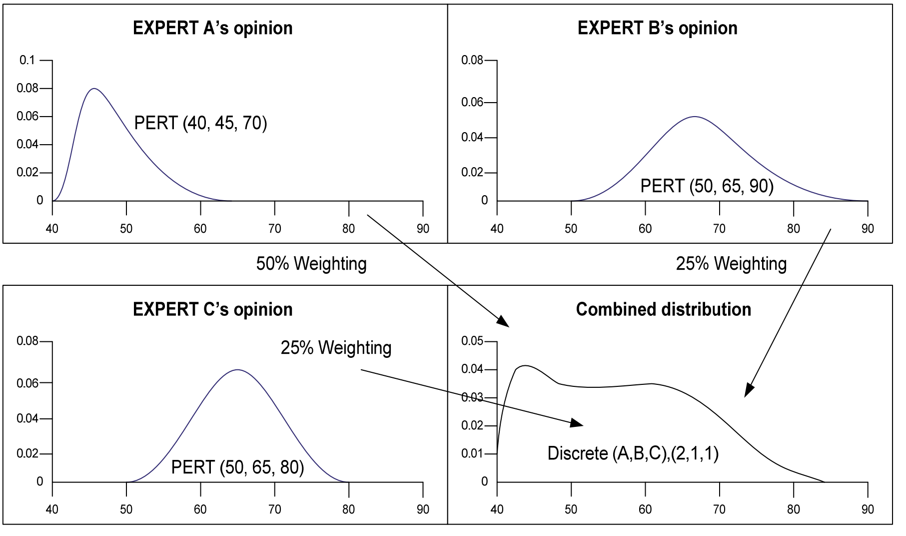


Figure 4. Combining three dissimilar expert opinions (Vose, 2008)

As can be observed, from the input contribution to the risk model from the SMEs is a combination distribution. This is not a common situation but there is always a chance for encountering these types of challenges. Especially for complex projects where a considerable expertise is required. Figure 4 illustrates an example of combining three differing opinions, but where expert A is given twice the emphasis of the owing to the greater experience of that expert (Vose, 2008).

It is relevant to notice that not all the SMEs are willing to participate actively in a risk management exercise at the first time, especially if they have no been exposed before into one. When SMEs are first asked to provide probabilistic estimates, they usually won't be particularly good at it because it is a new way of thinking (Vose, 2008).

For overcoming this and in order to provide the SMEs some sort of support in the field of the probabilistic risk management, we held a sort of training and educational meeting with all the RMT and SMEs.

It can be noted from Figure 5 that the expert's knowledge can extend beyond the knowledge base into the absolute truth area as a result of creativity and imagination of the expert.

Therefore, the intersection of the expert's knowledge with the ignorance space outside the knowledge base can be viewed as a measure of creativity and imagination. Another expert (i.e., Expert B) would have her/his own ellipses that might overlap with the ellipses of Expert A, and might overlap with other region by varying magnitudes (Ayyub, 2000).

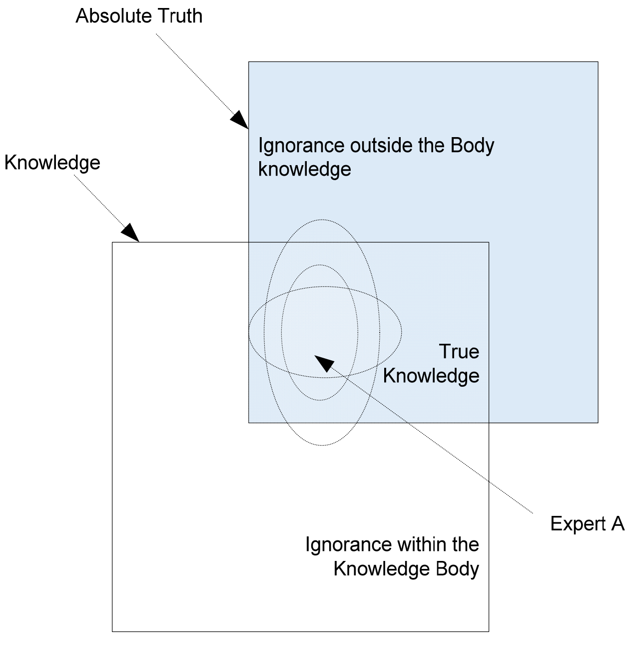


Figure 5. Knowledge and Ignorance of Humans (Ayyub, 2000),

SME`s are crucial for giving credible data for the risk model. If SMEs are not part of the risk management study, the results will not be trusted, causing this a failure of the process, the PM and the team.

7. Risk management team elicitation

The selection of members for the Risk Management Team (RMT), is not an easy task at all. A depth analysis of background and education was carried on for selecting the potential members of the team. The analysis, brainstorming, experience, background of the RMT members is critical not only for the risk management identification; it is for all the process including the monitoring and control. Eliciting the risk team members can be such a critical milestone for success or failure for the rest of the risk management process and the project itself. At the end, risk management is an input-output process that if wrong data or knowledge is feed into, then the results expected, most likely would be trustless or would included biases which at the end affect seriously the integrity and best practices in risk management.

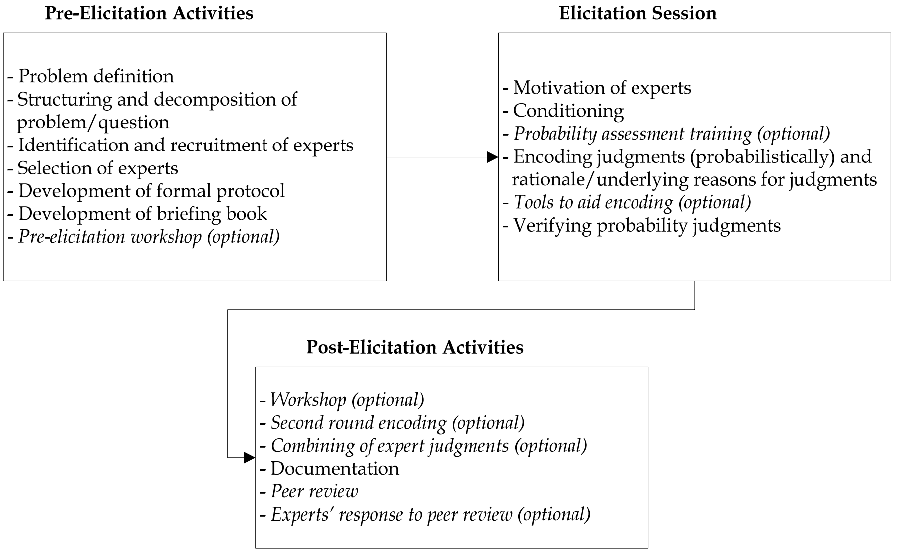


Figure 6. Overview of the Expert Elicitation Process (USEPA, 2011)

Expert Elicitation (EE) is a multidisciplinary process (Figure 6) that can inform decisions by characterizing uncertainty and filling data gaps where traditional scientific research is not feasible or data are not yet available. Although there are informal and nonprobabilistic EE methods for obtaining expert judgment. The goal of an EE is to characterize, to the degree possible, each expert's beliefs (typically expressed as probabilities) about relationships, quantities, events, or parameters of interest. The EE process uses expert knowledge, synthesized with experiences and judgments, to produce probabilities about their confidence in that knowledge. Experts derive judgments from the available body of evidence, including a wide range of data and information ranging from direct empirical evidence to theoretical insights. Even if direct empirical data were available on the item of interest, such measurements would not necessarily capture the full range of uncertainty. EE allows experts to use their scientific judgment transparently to interpret available empirical data and theory (Ayyub, 2011).

As mentioned by Vose (2008), it is usually impossible to obtain data from which to determine the uncertainty of all the variables within the model accurately, for a number of reasons:

- The data have simply never been collected in the past
- The data are too expensive to obtain
- Past data are no longer relevant (new technology, changes in political or commercial environment, etc)
- The data are sparse, requiring expert opinion "to fill in the holes"
- The area being modelled is new

The uncertainty in subjective estimates has two components: the inherent randomness of the variable itself and the uncertainty arising from the expert's lack of knowledge of the parameters that describe that variability.

In cases where there are no subject matter experts available, it is recommended to even hire external experts for obtaining the uncertainty ranges per variable or risk. Caltrans's practices towards project management and the project delivery process use a Project Development Team (PDT) which changes alongside the phases of the project. These teams are rather big in size and usually integrate stakeholders, subject matter experts and individuals coming from the same or different functional units (right of way, environmental, design, construction, etc.). Firstly, the downside of using such a team like the PDT is the size since it could be complicated for the RM to facilitate the meetings for obtaining better results towards assessing the uncertainty. In addition, the team members should feel free to talk about risk and in some cases having more than one member coming from the same division of functional unit, can cause some limitations for discussion and brainstorming.

In general, risk teams are rather small in compare with the traditional project teams. As a general rule, to invite a representative from each department or division is recommended. This approach definitely will help the whole process and will assist the RM to maintain the team focused in talking only about risk.

8. Risk identification, analysis and response

The whole process of risk management was implemented for three different projects. Regardless of the project scope, cost, schedule, location, type of funding, etc; the same steps were followed in order to determine the overall risk cost 8cost contingency). Figure 7 shows

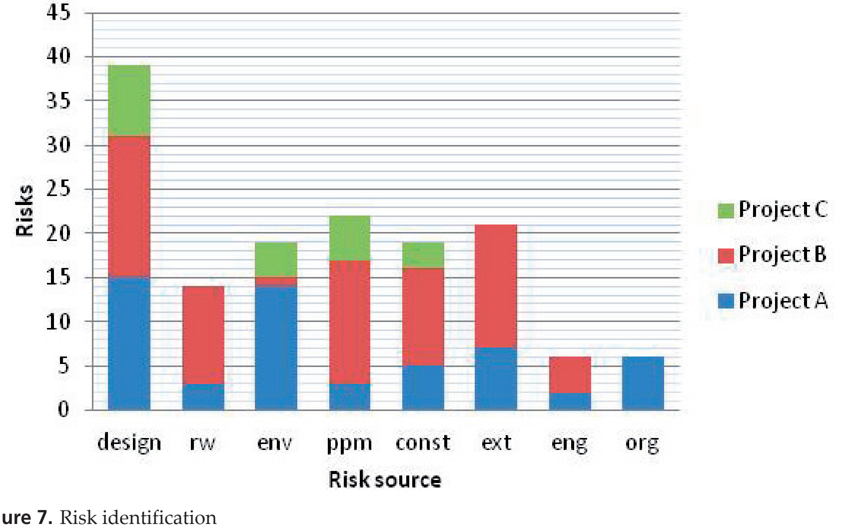


Figure 7. Risk identification

the risks identified for each project; it can be observed the source of risk, which in Caltrans relates to the functional units (e.g. environmental, right of way, design, project management, construction etc.)

Project A was a Historical Bridge, Project B a Direct Access Ramp (DAR) with a Transit Station (TS) and Project C a standard bridge. The main goal of the risk analysis for each project was to determine the cost risk contingency associated to each project risk register. Figure 10 shows the critical risks obtained for each project once the qualitative analysis was performed. The qualitative assessment matrix (Figure 9) used for ranking the risk criticality, forms part of Caltrans standard process for risk management.

The risk sources indicated in Figure 7 are: design, right of way (rw), environmental (env), program project management (ppm), construction (const), exterior (ext), engineering (eng) and organizational (org).

Option 1: PxI Matrix for Significant Focus on High and Very High Impacts (Non-linear Impact Scoring)					
Probability	Threats				
5	5	10	20	40	80
4	4	8	16	32	64
3	3	6	12	24	48
2	2	4	8	16	32
1	1	2	4	8	16
	1	2	4	8	16
	Impact on Selected Objective				

Figure 8. Risk matrix (Caltrans, 2007)

The RM was responsible for educating the RMT along each phase of the risk management implementation. For the qualitative risk analysis in particular, it was useful to provide tangle examples to the team before and during the meeting.

As can be observed, the critical risks are only a few ones in compare with all the risks identified at the initiation phase. The purpose of the qualitative risk analysis was to select those risks that represent a major negative or positive impact into the project objectives. A critical part for performing the qualitative analysis was to define the probability and impact ranges. The impacts were defined in terms of cost only for these projects. In addition, construction, engineering and organizational risks were not part of the quantitative risk analysis, since there were not scored as critical in the qualitative risk analysis.

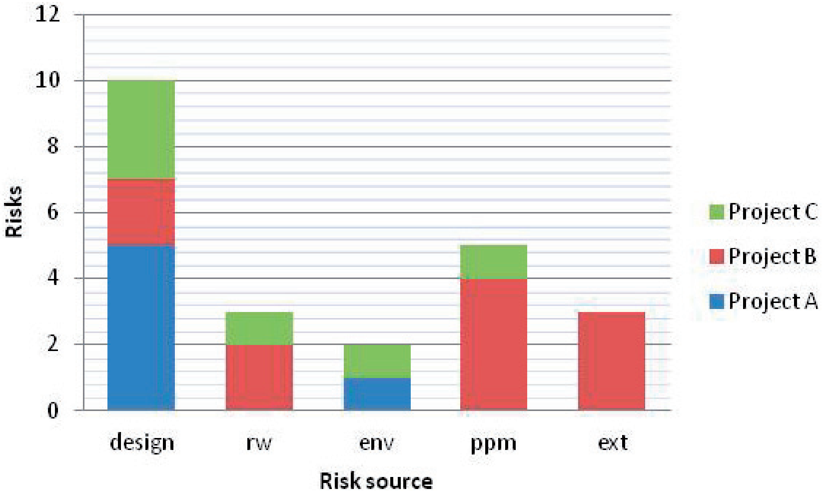


Figure 9. Qualitative risk analysis

The output of the qualitative risk analysis formed the basis or inputs for conducting the quantitative risk analysis. A set of cost ranges where developed for each project for matching the impact value selected. A model was built for running the Monte Carlo Simulation technique for obtaining the risk cost contingency for each project. Figure 10 shows the contingency values obtained per project. It is important to notice that instead of assigning a randomly selected contingency percent, these values are based directly on the critical risks identified with the Risk Registers.

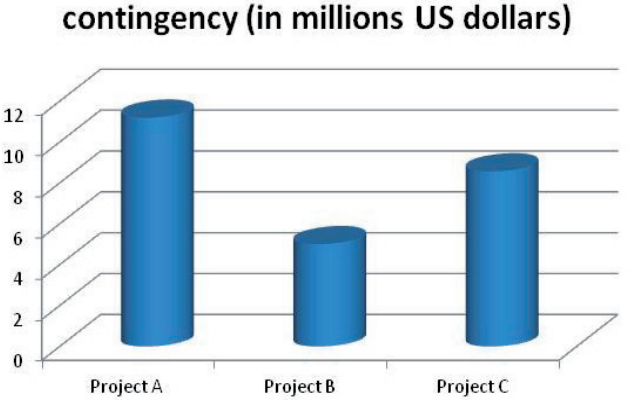


Figure 10. Quantitative risk analysis

The cost contingency per project assessed with the quantitative analysis provides a substantial support for justifying the overall project uncertainty. It was a common practice at Caltrans to assign a merely flat rate for risk contingency, without referring to specific project risks or by justifying the percentage upon a formal risk management study.

For each project, a set of risk responses were developed by the RMT and placed with the Risk Register. A Risk Owner was named for implementing the responses. It is recommended to match the risk owner with the source of the risk. In other words, if the risk has a source in construction, then the best suitable risk owner should come from that division.

The whole risk management implementation, including the risk identification, analysis and response was conducted in three meetings with the RMT. This is a standard set at Caltrans. The time and effort invested from the RM is not included since most of this work is done outside the meetings.

9. Risk management meetings

It was not common at all at Caltrans to have risk management meetings for the PDT members and even for executives. What is a reality is that those meetings are critical for performing in depth thoughts about what can go wrong with the project. The most critical of the meetings is the planning meeting; which usually includes educating the RMT. Questions like: why a risk management exercise is needed? What could be the possible project risk contingency? What are the restrictions? What is the overall project uncertainty rate? Need answers for properly doing a risk management study.

The following aspects are quite relevant in order to keep an order and sense within the risk management planning meeting (Vose, 2008):

- Rank the questions that need answering from “critical” down to “interesting”
- Discuss with the risk analysis the form of the answer
- Explain what arguments will be based on these outputs
- Explain whether the risk analysis has to sit within a framework
- Explain the target audience
- Discuss any possible hostile reactions
- Figure out a timeline
- Figure out the priority level
- Decide on how regularly the decision-maker and risk analysis will meet

It was interesting to notice along the meetings that when the meetings were held, the invitees were already prepared just to talk about risk. In some cases, critical project issues were discovered thanks to the risk discussions.

The risk management meetings were properly planned, one for the identification, analysis and results. No meeting took more than two hours and instead of meeting minutes, the Risk Register was used as the deliverable for discussion and follow up.

10. Risk monitoring and control

Risk monitoring and control has been mentioned as one of the most common failures of risk management. In part because the follow up process is usually forgotten by the project and risk manager and as results, there is no comparison between the baseline and actual Risk Registers or risk results. Risk metrics are fundamental for determining and assessing how risk management is contributing or enhancing the project delivery process. If the benefits of risk management are tangible and can be promoted with management and executives, the chances for formalizing the risk management process are very high. Otherwise, there will be less interest and support for keeping such a program.

Figure 11 shows the risk contingency behaviour for a given project. As can be notice, the contingency is dynamic thought the project delivery process. For example, in can be assumed that the baseline risk contingency was the one marked in grey colour. Then, after the risks responses implementation, risks are mitigated and the contingency is reduced (red colour). Nevertheless, the monitoring and control process continues and a new risk arises, increasing again the project contingency (blue colour).

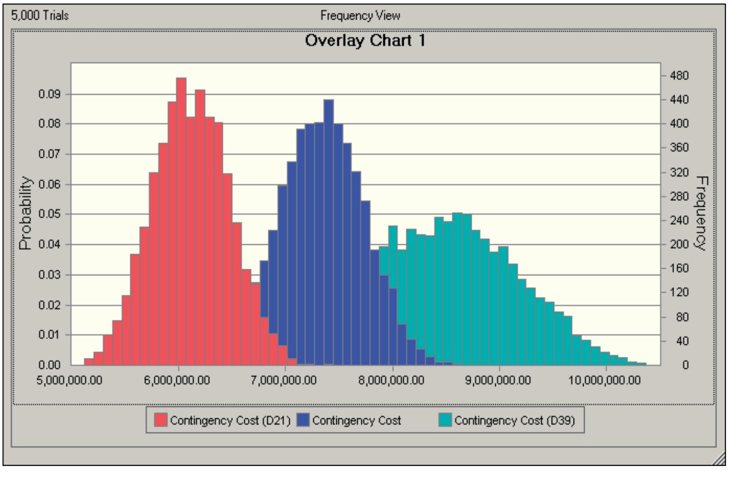


Figure 11. Risk contingency behaviour

Caltrans has implemented risk monitoring through the project development team meetings, where the project manager and the risk manager address any changes within the current risks or document new risks. In addition, an alternative way for risk monitoring has been through email. In this case, the risk manager is in charge of contacting the risk owners for updates and feedback regarding their risks.

Risk status reports have been developed with the purpose of maintaining informed the executive though a proper risk management communication. These reports are developed by the RM and are updated every time there is a change with the Risk Register.

Risk monitoring and control must be maintain through the project life cycle until the closeout phase, were the lessons learned can include feedback from risk management. The PM and RM should put extra care in keeping a constant review of the baseline Risk Register with the purpose of actively managing the project contingency and for assessing the effectiveness of the risk responses.

11. Risk management system

A risk management system could be considered the ultimate tool for managing risks for a portfolio of projects. In might be a powerful instrument for communication, monitoring and control. In addition, the basis for a future risk management database could be established.

In practical terms, for having a risk management system it is necessary to have developed a sufficient number of risk management studies and to have performed formal training to project managers. By having a set of projects in which risk management was previously implemented, the system can be fulfilled with data, including lessons learned. This data can provide an additional support for developing new risk management plans. A system can represent a considerable advantage for executives and project managers since can provide risk management status reports for a set of projects for specific data dates. The project manager can view, edit and direct actions to the team while using the system. Alternately, the RM can use the system for the risk management implementation. The following figure shows a view of Caltrans risk management system.

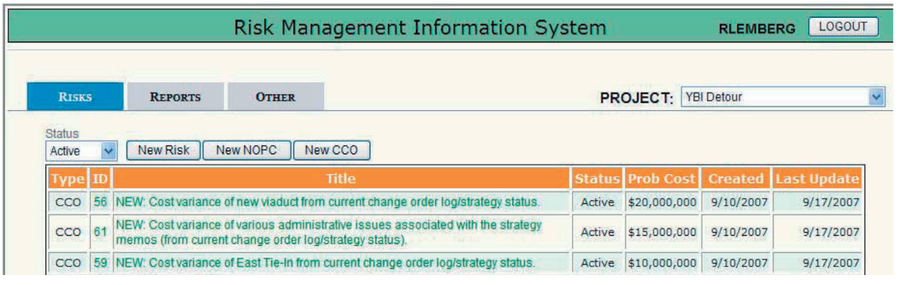


Figure 12. Caltrans Risk Management System

The RMT can view all the information for their project(s). In addition, some members for example the project manager, the design manager or the risk manager can have rights for edition. Usually with these systems, electronic alerts or messages are sent via email to keep all the team informed. As well, reports can be generated for supporting the decision making process.

Although a risk management system is a great tool for supporting the decision making along the project delivery process. It is recommended first to start risk management with education, a pilot project and training. After several studies, the role of a system can be justified for enhancing the overall process of risk management and its communication.

12. Caltrans enterprise risk management

Although the term of enterprise risk management (ERM) is not new, Caltrans started looking into its current process for doing business. The nature of risk management is intimately related to Caltrans functional units. Program Project Management, Construction, Environmental, Design, Right of Way and Surveys are the most representative functional divisions existing at Caltrans. Each division plays a role with the project delivery process;

usually representatives of each division formed the project development team (PDT) which has the responsibility of carried on the project since the initial phase until the closeout.

Caltrans has four major risks. These are project risk, program risk, operations risk, and organizational risk. Taken together, management of these components constitutes a Department-wide Risk Management Strategy. In some organizations, this is known as "Enterprise Risk Management." Each of these risk components relate directly to the different types of the work performed by the Department's functional areas. For example, project risk is related to the project delivery process; those things that, if they occurred, could impact a project's schedule, cost, or scope. Project Delivery staff can assess and take intelligent risks in delivery because taking intelligent risks fosters innovation and responsible decision-making. But, at the same time, Caltrans needs to follow its project-related processes and controls to manage that risk. Caltrans must monitor the quality of our performance over time, and evaluate any deficiencies and adjust our processes accordingly. Actively incorporating the concept of Risk Management into Caltrans's programs and activities is a dynamic process and must be constantly evaluated and adjusted to meet the strategic goals of the Department.

According to the Minesota Department of Transportation (MnDOT, 2012), ERM is a risk-based approach to managing an enterprise, incorporating concepts of internal control, planning and budgeting. It addresses the needs of various stakeholders who want to understand the broad spectrum of risks facing complex organizations so they can be appropriately managed.

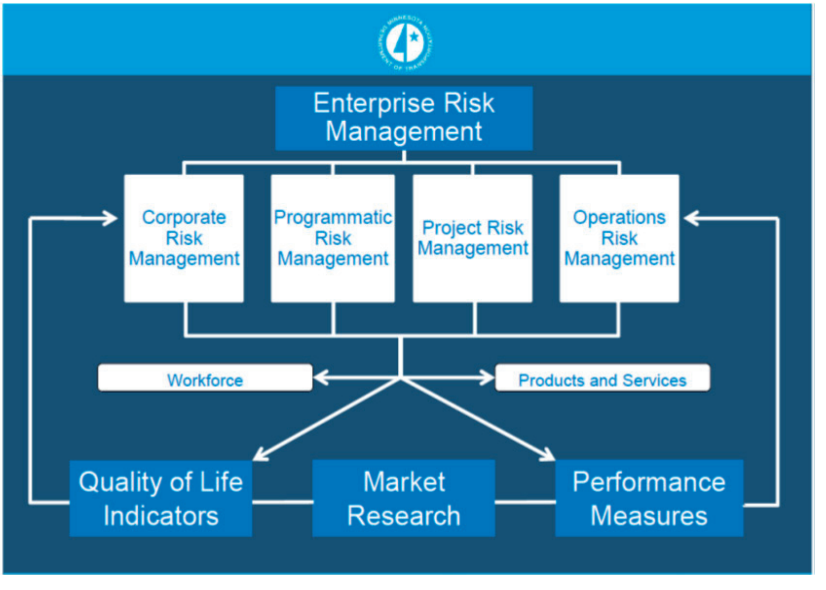


Figure 13. Enterprise Risk Management System (MnDOT, 2012)

The ERM at the Minesota Department of Transportation (Figure 13) is a clear example of how risk management can evolve for an organization. It is important to notice that outside of the common risk sources (corporate, programmatic, project and operations); other variables such the quality of life indicators, market research and performance measures are included. These variables have a direct impact into the risk sources, which at the end could influence the results and benefits of risk management.

Risk management has to be implemented for projects or within projects, but this is only the first step. Risk management means a change of doing business. For that reason, the culture of implementing Risk management should be brought by the executives and the company's policies. Risk management has evolved into the "Enterprise Risk Management" (see Figure 14).

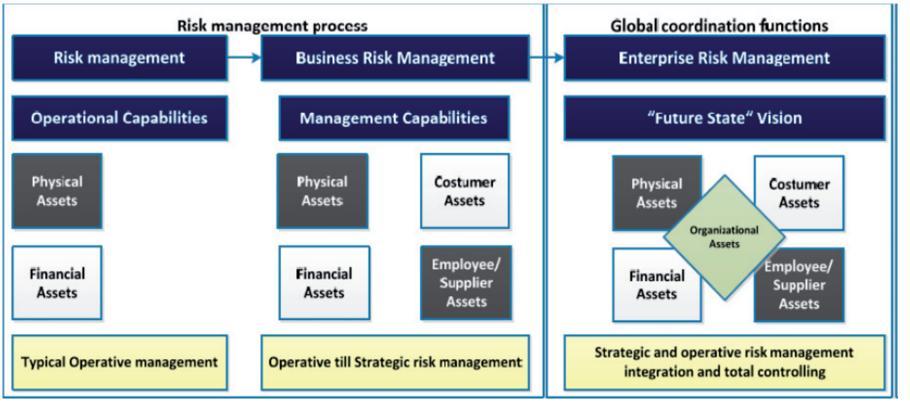


Figure 14. Risk Management evolution (Protiviti, 2006)

This new methodology explains that there are different philosophies about risk management with their own methods and focus; these are normally classified in four different types.

- the Risk Silo Management (Operational Risks),
- the Integrated Risk Management (Economic and Capital Risks),
- the Risk and Value Management (Management and Performance) and
- the Strategic Risk Management (Senior Management and Strategical Risks),

ERM is in charge to unify all this philosophies and confer the determination of one whole risk in a corporation. One of the most important topics handled by the ERM is the determination of the Risk Appetite. The Risk Appetite is: "the quantum of risk that the firm is willing to accept within its overall capacity" (Barfield, 2008).

In order to propitiate the ERM in every company, the standard "ISO 31000:2009 Risk management - Principles and guidelines" is a comprehensive guide of how companies should implement a formal risk management process. It sets out principles, a framework and a process for the management of risks that are applicable to any type of organization in public or private sector. It does not mandate a "one size fits all" approach, but rather emphasizes the fact that the management of risk must be tailored to the specific needs and structure of the particular organization. It depends at the end of the company's desire to be competitive and willing to manage proactively risks and opportunities.

Caltrans` goal towards ERM is related directly to its current position as a leader in the United States in the field of transportation projects. Caltrans expects to implement ERM in a near future with the intention of reinforcing its project and business processes.

13. Conclusions

Caltrans has evolved considerably in the past five years in the field of project risk management. The most common project risk management techniques used in the private sector are currently part of Caltrans project delivery process. The implementation of project risk management has assisted Caltrans executives and project managers in assessing properly the project contingency cost based upon specific identified risks. Enterprise Risk Management is still new in Caltrans. Nevertheless, management is taking currently formal steps in implementing it through all the state of California with the intention of managing and controlling not only project risks. The goal is to standardise the best practices in risk management from an enterprise perspective.

Author details

Pedro Maria-Sanchez
California Department of Transportation, District 11 San Diego, USA

14. References

Ayyub, B. (2000). Methods for Expert-Opinion Elicitation of Probabilities and Consequences for Corps Facilities. IWR Report 00-R-10. Prepared for U.S. Army Corps of Engineers Institute for Water Resources, Alexandria, VA. Available from http://www.iwr.usace.army.mil/docs/iwrreports/00-R-101.pdf

Barfield, R. (2008). Private Waterhouse Coopers. Risk appetite - How Hungry are you. 07. 07 2008, Available from http://www.pwc.com/uk/eng/about/svcs/vs/pwc_risk-appetite.pdf

Caltrans. (2007). Project Risk Management Handbook, Threats and Opportunities, (2nd Ed.), Office of Statewide Project Management Improvement, Sacramento, CA, USA, May 2, 2007

MnDOT. (2012). Enterprise Risk Management. 2000-2012, Available from http://www.dot.state.mn.us/riskmanagement/index.html

Protiviti. (2006). Guide to Enterprise Risk Management - Frequently asked questions. 24. 06 2008, Available from http://www.protiviti.de/portal/site/pro-us

USEPA. (2011). Expert Elicitation Task Force White Paper. August 2011, Available from http://www.epa.gov/stpc/pdfs/ee-white-paper-final.pdf

Vose, D. (2008). Risk Analysis, A quantitative guide, (3rd Ed.), John Wiley & Sons, ISBN 978-0-470-51284-5, West Sussex, PO19 8SQ, England

WSDOT. (2010). Project Risk Management, Guidance for WSDOT Projects, Olympia, WA, USA, July, 2010

Chapter 19

Risk Management in Construction Projects

Nerija Banaitiene and Audrius Banaitis

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51460

1. Introduction

The financial and economic crisis has had an adverse impact on the Lithuania's economy and construction industry. The GDP of Lithuania grew slightly in 2010, in contrast to a decrease of 14.7% in 2009. Lithuania's GDP increased from 1.3% in 2010 to 4.6% in 2011. Annual GDP growth decreased from its highest point of 6.7%, reached in the third quarter, to 4.4% in the last quarter of 2011 [1,2]. Some industries, such as construction; trade, transport and communications; and the industry sectors were most affected by the crisis. In 2010, the gross value added within the construction sector decreased by 43.3%, and in the trade, transport and communications sector – by 16.6%. In 2011, a positive change in the gross value added was observed in all groups of economic activities. The largest growth in the gross value added was observed in enterprises engaging in construction (by 15%) and trade, transport and communication services (7.3%) [1,3]. The construction sector, one of the engines of economic growth in Lithuania over the last decade, is now facing with serious challenges as companies' closures, rising unemployment, and postponed or even cancelled investments. These events also have changed the clients' and construction companies' behaviour. A reduced demand and shortage of orders dramatically increased a competition between companies of the construction sector. This increased pressure to improve quality, productivity and reduce costs, and the need for project strategies and management that can appropriately and effectively manage project risk.

Risk management is one of the nine knowledge areas propagated by the Project Management Institute [4]. Furthermore, risk management in the construction project management context is a comprehensive and systematic way of identifying, analyzing and responding to risks to achieve the project objectives [5,6]. The benefits of the risk management process include identifying and analyzing risks, and improvement of construction project management processes and effective use of resources.

Construction projects can be extremely complex and fraught with uncertainty. Risk and uncertainty can potentially have damaging consequences for the construction projects [7,8].

Therefore nowadays, the risk analysis and management continue to be a major feature of the project management of construction projects in an attempt to deal effectively with uncertainty and unexpected events and to achieve project success.

Construction projects are always unique and risks raise from a number of the different sources [9,10]. Construction projects are inherently complex and dynamic, and involving multiple feedback processes [11,12]. A lot of participants – individuals and organisations are actively involved in the construction project, and they interests may be positively or negatively affected as a result of the project execution or project completion [4]. Different participants with different experience and skills usually have different expectations and interests [13]. This naturally creates problems and confusion for even the most experienced project managers and contractors.

Cost of risk is a concept many construction companies have never thought about despite the fact that it is one of the largest expense items [14]. Risk management helps the key project participants – client, contractor or developer, consultant, and supplier – to meet their commitments and minimize negative impacts on construction project performance in relation to cost, time and quality objectives. Traditionally, practitioners have tended to associate construction project success with these three aspects of time, cost and quality outcomes.

The current economic downturn and challenges in a highly competitive Lithuania's construction sector require contractors to manage risks by themselves. This paper reports the research that aims to examine the risk analysis and risk management practices in the Lithuanian construction companies.

2. Literature review

In today's post-crisis economy effective risk management is a critical component of any winning management strategy. Risk management is one of the nine knowledge areas propagated by the Project Management Institute (PMI). The PMBOK® Guide recognises nine knowledge areas typical of almost all projects. The nine knowledge areas are [4]:

1. Project integration management.
2. Project scope management.
3. Project time management.
4. Project cost management.
5. Project quality management.
6. Project human resource management.
7. Project communications management.
8. Project risk management.
9. Project procurement management.

Although these knowledge areas are all equally important from a project manager's point of view, in practice a project manager might determine the key areas which will have the greatest impact on the outcome of the project.

Each PMI knowledge area in itself contains some or all of the project management processes. For example, project risk management includes [4]:

- Risk management planning;
- Risk identification;
- Qualitative risk analysis;
- Quantitative risk analysis;
- Risk response planning;
- Risk monitoring and control.

Risk management is probably the most difficult aspect of project management. A project manager must be able to recognise and identify the root causes of risks and to trace these causes through the project to their consequences. Furthermore, risk management in the construction project management context is a comprehensive and systematic way of identifying, analyzing and responding to risks to achieve the project objectives [5,6]. The use of risk management from the early stages of a project, where major decisions such as choice of alignment and selection of construction methods can be influenced, is essential [15]. The benefits of the risk management process include identifying and analyzing risks, and improvement of construction project management processes and effective use of resources.

The construction industry is heterogeneous and enormously complex. There are several major classifications of construction that differ markedly from one another: housing, non-residential building, heavy, highway, utility, and industrial [16]. Construction projects include new construction, renovation, and demolition for both residential and non-residential projects, as well as public works projects, such as streets, roads, highways, utility plants, bridges, tunnels, and overpasses. The success parameters for any project are in time completion, within specific budget and requisite performance (technical requirement). The main barriers for their achievement are the change in the project environment. The problem multiplies with the size of the project as uncertainties in project outcome increase with size [17,18]. Large construction projects are exposed to uncertain environment because of such factors as planning, design and construction complexity, presence of various interest groups (owner, consultants, contractors, suppliers, etc.), resources (manpower, materials, equipment, and funds) availability, environmental factors, the economic and political environment and statutory regulations.

Construction projects can be unpredictable. Managing risks in construction projects has been recognized as a very important process in order to achieve project objectives in terms of time, cost, quality, safety and environmental sustainability [19]. Project risk management is an iterative process: the process is beneficial when is implemented in a systematic manner throughout the lifecycle of a construction project, from the planning stage to completion.

In the European Union construction is the sector most at risk of accidents, with more than 1300 people being killed in construction accidents every year. Worldwide, construction workers are three times more likely to be killed and twice as likely to be injured as workers in other occupations. The costs of these accidents are immense to the individual, to the employer and to society. They can amount to an appreciable proportion of the contract price [20].

Construction activities in Lithuania provided employment to an estimated 93.7 thousand persons in 2011, while an annual turnover in excess of EUR 1.91 billion [21]. Construction is one of Lithuania's largest industries. Unfortunately it has also the occupational health and safety problems. More construction workers are killed, injured or suffer ill-health than in any other industry. In 2011, 13 construction workers killed whilst at work, compared to 7 industrial workers and 4 agricultural workers. In comparison with 2010, the number of fatal accidents in construction enterprises increased by more than 2 times, i.e. from 6 to 13 cases has been reported [22].

The risk analysis and management techniques have been described in detail by many authors [23-27]. A typical risk management process includes the following key steps [28]:

- Risk identification;
- Risk assessment;
- Risk mitigation;
- Risk monitoring.

Risk identification is the first and perhaps the most important step in the risk management process, as it attempts to identify the source and type of risks. It includes the recognition of potential risk event conditions in the construction project and the clarification of risk responsibilities [29]. Risk identification develops the basis for the next steps: analysis and control of risk management. Corrects risk identification ensures risk management effectiveness. Carbone and Tippett [30] stated that the identification and mitigation of project risks are crucial steps in managing successful projects.

The PMBOK® Guide [4] defines a project risk as "an uncertain event or condition that, if it occurs, has a positive or negative effect on at least one project objective". There are many possible risks which could lead to the failure of the construction project, and through the project, it is very important what risk factors are acting simultaneously. As stated by Raz et al. [31], too many project risks as undesirable events may cause construction project delays, excessive spending, unsatisfactory project results or even total failure.

Many approaches on risk classification have been suggested in the literature for effective construction project risk management. Tah and Carr [32] categorized risks into two groups in accordance with the nature of the risks, i.e. external and internal risks. Combining the fuzzy logic and a work breakdown structure, the authors grouped risks into six subsets: local, global, economic, physical, political and technological change. According to Wang et al. [33], the classification of the risks depends mainly upon whether the project is local or international. The internal risks are relevant to all projects irrespective of whether they are local or international. International projects tend to be subjected to the external risk such as unawareness of the social conditions, economic and political scenarios, unknown and new procedural formalities, regulatory framework and governing authority, etc.

According the PMBOK® Guide [4], the risks are categorized into such groups: technical, external, organizational, environmental, or project management. Some categories of risk that affect a construction project are similar to risks for other investment projects, whether it is an investment in common stocks or government bonds, and some are specific to construction. The risk identification process would have highlighted risks that may be considered by project management to be more significant and selected for further analysis [34]. Risk identification is an iterative process because new risks may become known as the project progresses through its life cycle and previously-identified risks may drop out [35]. Construction projects carry complex risks for all involved—including owners, consultants, contractors, and suppliers—that can increase when construction takes place near an active facility or congested area. Risks include geological or pollution-related conditions, interference with ongoing operations, construction accidents, as well as design and construction faults that may negatively impact the project both construction and when the project is complete.

Generally two broad categories, namely, qualitative and quantitative analysis are distinguished in literature on risk assessment. A qualitative analysis allows the key risk factors to be identified. Risk factors may be identified through a data-driven (quantitative) methodology or qualitative process such as interviews, brainstorming, and checklists. It is considered as an evaluation process which involves description of each risk and its impacts or the subjective labelling of risk (high/medium/low) in terms of both risk impact and probability of its occurrence [19]. Qualitative risk analysis assesses the impact and likelihood of the identified risks and develops prioritized lists of the risks for further analysis or direct mitigation. Carr and Tah [36] introduced a hierarchical risk breakdown structure (HRBS), and the HRBS represents a formal model for qualitative risk assessment. Quantitative analysis involves more sophisticated techniques and methods to investigate and analyze construction project risks. Quantitative risk analysis attempts to estimate the frequency of risks and the magnitude of their consequences by different methods such as the decision tree analysis, the cost risk analysis, and Monte Carlo simulation [37]. The application of the quantitative risk analysis allows the construction project exposure to be modelled, and quantifies the probability of occurrence of the identified risk factors as well as their potential impact.

Various risk management tools are available, but unfortunately they are not suitable for many industries, organizations and projects [38]. Although today's organizations appreciate the benefits of managing risks in construction projects, formal risk analysis and management techniques are rarely used due to lack of knowledge and to doubts on the suitability of these techniques for construction projects.

There are four alternative strategies – risk avoidance, risk transfer, risk mitigation, and risk acceptance, for treating risks in a construction project. As stated by Hillson [39], risk mitigation and risk response development is often the weakest part of the risk management process. The proper management of risks requires that they be identified and allocated in a well-defined manner. This can only be achieved if contracting parties comprehend their risk responsibilities, risk event conditions, and risk handling capabilities [40].

Before the crisis (2004-2008), due to a lack of contractors' responsibilities and control in various steps of a project's development, the time and quality performance levels of construction projects in the Lithuania were generally inadequate or even poor. In construction projects, many parties are involved such as owner, consultant, contractor, sub-contractor, and supplier. Each party has its own risks. Risk transfer means the shift of risk responsibility to another party either by insurance or by contract. Wang and Chou [29] reported that contractors usually use three methods to transfer risk in construction projects:

- through insurance to insurance companies;
- through subcontracting to subcontractor;
- through modifying the contract terms and conditions to client or other parties.

Construction projects can be managed using various risk management tools and techniques. Ahmed et al. [23] reviewed techniques that can be used for development of risk management tools for engineering projects. Techniques for context establishment, risk identification, risk assessment and treatment were provided. Application of risk management tools depends on the nature of the project, organization's policy, project management strategy, risk attitude of the project team members, and availability of the re-sources [12]. A risk assessor model (RAM) presented by Jannadi and Almishari [41] was developed to determine risk scores for various construction activities. The model provides an acceptability level for the risks and determines a quantitative justification for the proposed remedy.

Risks and uncertainties, involved in construction projects, cause cost overrun, schedule delay and lack of quality during the progression of the projects and at their end [28,29,42]. As stated by Baloi and Price [43], poor cost performance of construction projects seems to be the norm rather than the exception, and both clients and contractors suffer significant financial losses due to cost overruns.

Oyegoke et al. [44] discusses the problems of managing risk and uncertainty in construction project due to the owner dissatisfaction in project outcome and dynamism within agile construction environment. The authors identified some areas in supply chain processes which are prone to greater risks and uncertainty and propose an agile management principle based on the concept of integration and fragmentation in product development and execution processes respectively.

Many authors have reviewed problems on time performance in construction projects [43,45,46]. Aibinu and Odenyinka [46] investigated and assessed the causes of delays in building projects in Nigeria. The nine factor categories evaluated include: client-, contractor-, quantity surveyor-, architect-, structural engineer-, services engineer-, supplier-, and subcontractor-caused delays, and external factors (i.e. delays not caused by the project participants). Finally, ten overall delay factors were identified, namely: contractors' financial difficulties, client' cash flow problems, architects' incomplete drawings, subcontractors' slow mobilization, equipment break-down and maintenance problems, suppliers; late delivery of ordered materials, incomplete structural drawings, contractors' planning and scheduling problems, price escalation, and subcontractors' financial difficulties. The authors pointed the poor risk management as one of the principal delay factors and concluded that actions and inactions of construction project participants contribute to overall project delays.

According to Baloi and Price [43], the construction contractors highlight that delay in payments is common both in private and public projects, with the public sector being the worse defaulter. Moreover, most types of contracts presume compensation clauses for delay in payments, but clients rarely agree to pay the interests due to the contract. Nasir et al. [47] analysed schedule risks and developed a comprehensive construction schedule risk model is referred to as Evaluating Risk in Construction–Schedule Model (ERIC-S). The ERIC-S model provides decision support to project owners, consultants, and researchers as a project delay prediction tool. Similarly, the Cost-Time-Risk diagram (CTR) proposed by Aramvareekul and Seider [48] helps project managers consider project risk issues while monitoring and controlling their project schedule and cost performance in one diagram.

The performance by the project management team highly influences the success of a construction project. Some of the incidental risks associated with poor project management performance are [49]:

- Unclear or unattainable project objectives;
- Poor scoping;
- Poor estimation;
- Budget based on incomplete data;
- Contractual problems;
- Insurance problems;
- Delays;
- Quality concerns;
- Insufficient time for testing.

Many authors have recognized the value of trust within the project business. Lewicki and Bunker [50] emphasize that trust is a critical success element to most business, professional, and employment relationships. Trust is argued to improve the inter-organizational relationships among principal actors in project development, such as owners, contractors, and suppliers [51]. According by Krane et al. [52] trust between project owners and project managers is crucial for project success.

In business relations, as stated by Kaklauskas et al. [53], the global economic crisis brought about distrust of other stakeholders. Trust reinforces the relationships of the critical stakeholder that often determine the success of a project [51,54,55]. Ward and Chapman [56] concluded that stakeholders are a major source of uncertainty in construction projects. Smyth et al. [57] note that trust provides an important resource for creating greater probability and certainty. Wilkinson [58] found that project management companies need to overcome problems in their relationships with other professionals on the project team and with the client. For the success of construction projects, there is a need for alignment of the project owners’ interests and the project management team's interests and trust between them.

Construction projects are tendered and executed under different contract systems and payment methods [59]. According by Zaghloul and Hartman [60], there is no possibility to eliminate all the risks associated with a specific project. All that can be done is to regulate the risk allocated to different parties and then to properly manage the risk. Chapman and Ward [61] argue that the contract choice decisions are central to both stakeholder management and the management of risk and uncertainty. The authors proposed an integrated approach based on a balanced incentive and risk sharing (BIARS) approach to contracting as well as a best practice approach to risk management in terms of the whole project life cycle.

Contractors generally aim to make an acceptable range of profit margin. Profit margins in the industry have been low for most contractors on projects in recent years. Correct understanding and allocation of risk helps for contractors to avoid erosion of the profit margin. Ökmen and Öztas [62] proposed a new simulation-based model – the correlated cost risk analysis model (CCRAM) – to analyse the construction costs under uncertainty when the costs and risk-factors are correlated. The CCRAM model captures the correlation between the costs and risk-factors indirectly and qualitatively. Baloi and Price [63] determined the most critical risk factors affecting construction cost performance. The authors stated that global risk factors pose more challenges to contractors, which are less familiar with them. The authors introduced a fuzzy decision framework for a systematic modelling, analysis and management of global risk factors affecting construction cost performance from contractor's perspective and at a project level. Similarly, Ismail et al. [64] provide a 'Level-Severity-Probability' approach to determine the critical risk source and factors. Fuzzy logic is used in the proposed methodology for evaluation of the risk level, severity and probability. As stated by Zeng et al. [65], the application of fuzzy reasoning techniques provides an effective tool to handle the uncertainties and subjectivities arising in the construction project.

The review of the literature revealed a wide range of risk types and sources in construction projects, and that various risk management methods and techniques can be employed in the management of construction projects in order to control potential risks.

3. Methods and data

The aims of the research were: first, to identify contractors' opinion on the significance of the construction projects risks; and second, to explore the risk analysis and risk management practices in the Lithuanian construction companies.

The initial survey was distributed during February through March 2008. A second, similar questionnaire was distributed during February through March 2009. A questionnaire containing three sections was developed to facilitate data collection. The first section includes the respondents' opinion on the risk factor in terms of its probability and impact to overall construction project success. The second section includes the respondents' opinion on the risk consequences for construction project performance measures as well as the risk assessment and response practices. The third section aims to collect the background information of the respondents, e.g. their age, gender, position, education, work experience and professional background.

The questionnaire of first survey was distributed either personally or via e-mail to 40 members of top and middle management in the construction companies. A sample of 40 practitioners received the questionnaire and 38 valid questionnaires were returned for analysis with a response rate of 95%. The second questionnaire was distributed either personally or via e-mail to 35 members of top and middle management in the construction companies. Of the 35 questionnaires distributed in the second survey, 35 were returned, but 5 were incompletely completed and so were excluded from the data analysis. The response rate was 86%.

In both surveys, the baseline characteristics of the respondents were relatively similar. Of the 38 respondents in the first study, site managers comprise 29%, project managers 26%, other position senior managers 21%, civil engineers 16%, and designing engineers 8%. Of the 30 respondents in the second study, site managers, project managers, and other position senior managers comprise 80%.

The Likert scale was selected to obtain the probability of the risk factors in construction project that are identified in the literature review. A 5-point Likert scale was adopted, where 1 represented "rare", 2 "occasional", 3 "somewhat frequent", 4 "frequent", and 5 "very frequent". Likewise, the Likert scale was selected to obtain the impact of the risk factors in construction project that are identified in the literature review. A 5-point Likert scale was adopted, where 1 represented "very low", 2 "low", 3 "moderate", 4 "high", and 5 "very high".

The latest survey was carried out in December 2010-January 2011. In a subsequent study, the role of risk factors at a project level was addressed. A questionnaire containing three sections was developed to facilitate data collection. The first section includes the respondents' opinion on the risk consequences for construction project performance measures as well as the risk assessment and response practices. The second section includes the respondents' opinion on the risk factor in terms of its probability and impact to overall construction project success. The third section aims to collect the background information of the respondents, e.g. their age, gender, position, education, work experience and professional background.

The third questionnaire was distributed either personally or via e–mail to 23 members of top and middle management in the construction company and 23 valid questionnaires were returned for analysis in time to be included in the analysis (100% overall response rate). Of the 23 respondents in the study, site managers, project managers, and other position senior managers comprise 91%.

In all surveys, the majority of the respondents have more than 15 years' experience in construction/project management or working knowledge of construction/project management activities. Based on work experience and employment position, it was inferred that the respondents have adequate knowledge of the activities associated with construction project risk. This makes them as reliable and credible sources of information which is crucial to satisfy the research goal. The procedure, findings, and relevant discussion of the analyses are detailed in the following section.

4. Results

As outlined in Section 2, risk factors on construction projects can be split into two major groups:

1. Internal risks, which fall within the control of clients, consultants and contractors.
2. External risks, which include risk elements that are not in the control of key stakeholders.

The potential risk sub-factors were adapted from studies by Chapman and Ward [25], Tah and Carr [32], Perera et al. [40], Pinto et al. [51], Baloi and Price [63], Kartam and Kartam [66], Lahdenperä [67], Majamaa et al. [68], Mbachu and Nkado [69], Mitkus and Trinkūnienė [70], and Yang et al. [71].

In order to illustrate the respondents' opinions regarding the importance of analysed risk factors, an average was calculated for each factor. Next, the Kendall coefficient of concordance W [72,73] was calculated to test the reliability of the responses, and significance testing was based on the Chi-square distribution at the 1% significance level. The W coefficients were calculated for each defined group of risk factors created by the analysis perspectives.

In both surveys, the respondents agree as regards the external risks impact and probability. The respondents agree as regards the external risks impact, what can be judged by values W=0.183; χ^2=34.7 (α=0.01), in the first survey; W=0.10; χ^2=12.4 (α=0.01), in the second survey. The identified external risks according to their potential effect on construction project objectives were ranked. In the first survey, the top three important external risks identified are:

1. Natural forces;
2. Inflation and interest rate;
3. Fiscal policy.

In the second survey, the top three important external risks identified are:

1. Fiscal policy;
2. Natural forces;
3. Political controls.

Probability assessment of risks of the external project constrains is reflected in Fig. 1. Impact assessment of risks of the external project constrains is reflected in Fig. 2.

The risk management perceivers are the project participants, and a contractor is any entity which has the power to influence project decision making directly. Related to experience, only 11% of the respondents affirmed that they have experience in risk management. Most of them are project manager and have more than 15 years' experience; it proofs that the relationship between risk perception and experience of respondents. And even 34% of the respondents affirmed that they have no experience in risk management, while 55% of the respondents affirmed that they do not have enough experience in risk management. And

97% of the respondents answered that risks must be managed at the early stages of the construction project.

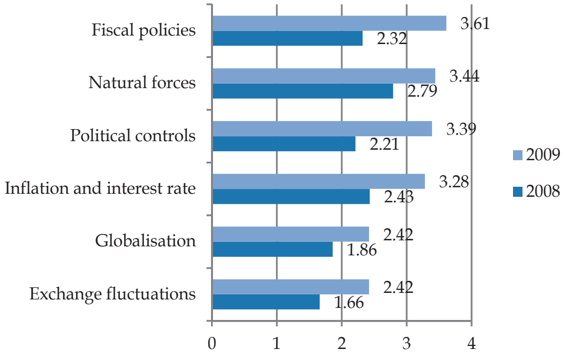


Figure 1. Probability assessment of external project risks

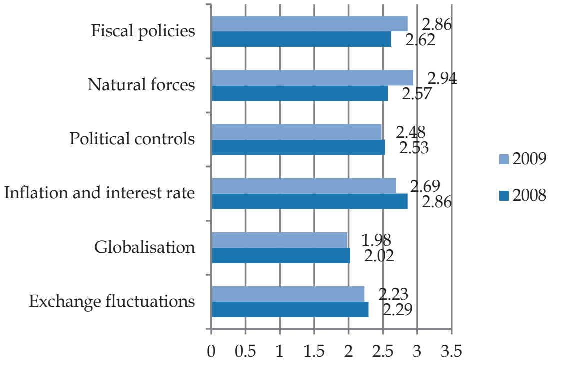


Figure 2. Impact assessment of external project risks

In terms of the sources and providers of the data and information required in the risk analysis, the most frequently used technique is experiential or documented knowledge analysis with 92% of the respondents' agreement in the first survey, and 93% of the respondents' agreement in the second survey (Fig. 3). And the project documentation reviews, project team brainstorming, and analysis of other information resources are frequently used in the risk assessment.

Comparison between the two surveys in terms of risk analysis showed a decrease in reviews of project documentation, from 63% in the first survey to 47% in the second survey, as well as greater use of experts' judgement, from 26% in the first survey to 43% in the second survey, and project team brain-storming, from 45% in the first survey to 53% in the second survey, in the risk assessment.

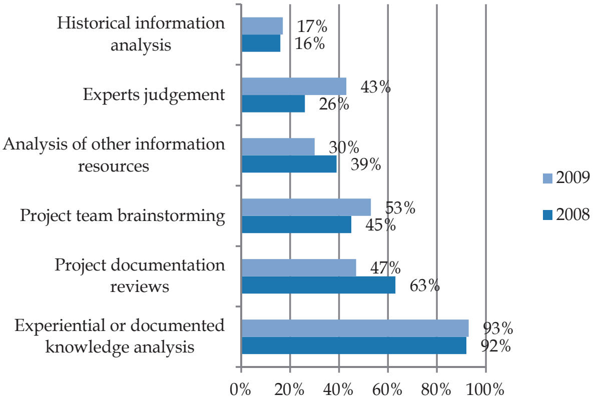


Figure 3. Risk analysis practices in construction projects

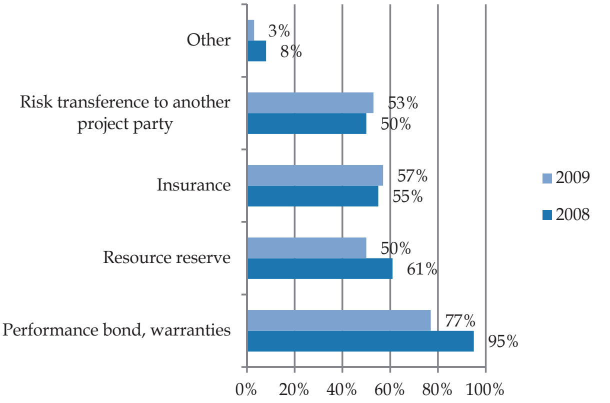


Figure 4. Risk response techniques employed for construction projects

In terms of the risk response tools and techniques, the most frequently used tool is performance bonds and warranties with 95% of the respondents' agreement in the first survey, and 77% of the respondents' agreement in the second survey (Fig. 4). And the some resource reservation, insurance, and risk transference to another project party are frequently used risk response techniques.

Comparison between the two surveys in terms of risk response tools and techniques showed a decrease of performance bond and warranties, from 95% in the first survey to 77% in the second survey, and resource reservation, from 61% in the first survey to 50% in the latter survey; as well as greater use of risk transference to another party, from 50% in the first survey to 53% in the second survey, and insurance, from 55% in the first survey to 57% in the latter survey, for the risk responses.

In last survey (2010-2011), the respondents agree as regards the project level risks impact, what can be judged by values W=0.54; χ^2=51.3 (α=0.01). As regards the assessment of the project level risks probability, respondents also agree what can be judged by values W=0.51; χ^2=48.5 (α=0.01). The identified project level risks according to their potential effect on construction project objectives were ranked. The top three important categories of internal risks identified are:

1. Construction risks;
2. Design risks;
3. Project management risks.

Overall assessment of risks of the internal project risks is reflected in Fig. 5. Risk priority is utilized during response planning and risk monitoring. It is critical to understand the priority for each risk as it allows the project team to properly understand the relative importance of each risk.

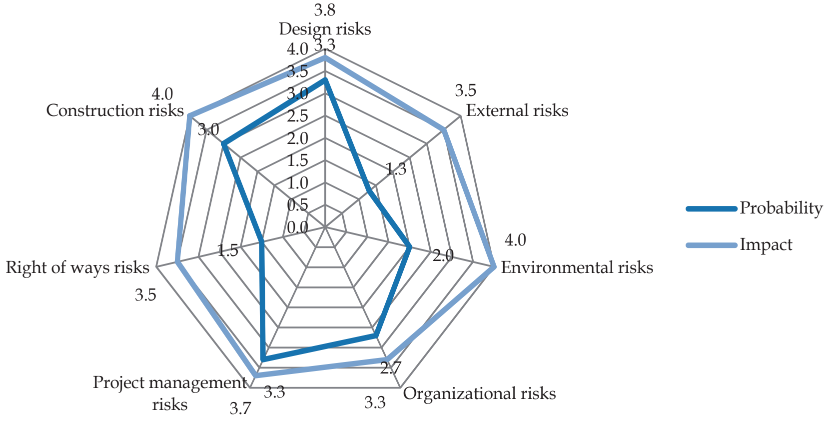


Figure 5. Assessment of project level risk categories

Table 1 shows that the controllable risk sources as identified in the study could be further broken down into seven sub-categories: design risks, external risks, environmental risks, organizational risks, project management risks, right of way risks, and construction risks which fall within the control of the project team.

	Categories	Likelihood 1 (rare)-5 (very frequent)	Impact 1 (very low)-5 (very high)
Design risks			
D1	Design errors and omissions	4	5
D2	Design process takes longer than anticipated	3	4
D3	Stakeholders request late changes	3	3
D4	Failure to carry out the works in accordance with the contract	3	3
External risks			
Ex1	New stakeholders emerge and request changes	2	4
Ex2	Public objections	1	3
Ex3	Laws and local standards change	1	3
Ex4	Tax change	1	4
Environmental risks			
En1	Environmental analysis incomplete	2	4
En2	New alternatives required to avoid, mitigate or minimize environmental impact	2	4
Organizational risks			
O1	Inexperienced workforce and staff turnover	3	3
O2	Delayed deliveries	3	3
O3	Lack of protection on a construction site	2	4
Project management risks			
PM1	Failure to comply with contractual quality requirements	3	4
PM2	Scheduling errors, contractor delays	4	4
PM3	Project team conflicts	3	3
Right of way risks			
R1	Expired temporary construction permits	1	4
R2	Contradictions in the construction documents	2	3
Construction risks			
C1	Construction cost overruns	4	4
C2	Technology changes	2	4

Table 1. Risk categories

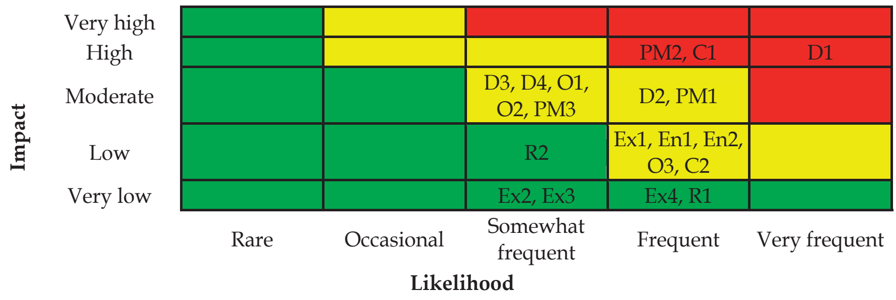


Figure 6. Risk matrix

Once the risks and probabilities are determined, the risk score can be calculated. Risk score is detailed in Table 1. The probability and impact matrix (Fig. 6) illustrates a risk rating assignment for individual risk factors in the identified risks categories. The risk matrix shows the combination of impact and probability that in turn yield a risk priority (shown by the red, yellow, and green colour). Qualitative risk analysis can lead to further analysis in quantitative risk analysis or directly to risk response planning.

Twenty risk factors were established to be significant under the internal risks categories. Under the design risk category, design errors/omissions and design process delays were the most frequently mentioned risk factors attributed to the contractors. Under the project management risk category, scheduling errors and failure to comply with contractual quality requirements were the most frequently mentioned risk factors. Under the construction risk category, construction cost overruns and technology changes were the most frequently mentioned risk factors attributed to the contractors. Respondents believed that these risk events are responsible for poor quality of work, delays and associated losses. Risks with high impact and high probability, such as D1 (design errors and omissions), C1 (construction cost overruns), and PM2 (scheduling errors, contractor delays) are required further analysis, including quantification, and aggressive risk management.

5. Conclusions

An effective risk management process encourages the construction company to identify and quantify risks and to consider risk containment and risk reduction policies. Construction companies that manage risk effectively and efficiently enjoy financial savings, and greater productivity, improved success rates of new projects and better decision making.

Risk management in the construction project management context is a comprehensive and systematic way of identifying, analyzing and responding to risks to achieve the project objectives. The research results show that the Lithuanian construction company significantly differ from the construction companies in foreign countries in the adoption of risk management practices. To management the risk effectively and efficiently, the contractor must understand risk responsibilities, risk event conditions, risk preference, and risk management capabilities.

The lack of experience makes it very difficult to change Lithuanian contractors' attitude towards risk management. Nevertheless, the construction companies need to include risk as an integral part of their project management. In our view, the use of risk management in the Lithuanian construction companies is low to moderate, with little differences between the types, sizes and risk tolerance of the organizations, and experience and risk tolerance of the individual respondents.

Qualitative methods of risk assessment are used in construction companies most frequently, ahead of quantitative methods. In construction project risk management, risks may be compared by placing them on a matrix of risk impact against a probability. Mitigation options are then derived from predefined limits to ensure the risk tolerance and appetite of the construction company.

The risk management framework for construction projects can be improved by combining qualitative and quantitative methodologies to risk analysis.

Author details

Nerija Banaitiene* and Audrius Banaitis
Department of Construction Economics and Property Management, Faculty of Civil Engineering, Vilnius Gediminas Technical University, Vilnius, Lithuania

6. References

[1] Statistics Lithuania. GDP by production, by expenditure, by income approach, IV Q 2011 (revised) and I Q 2012 (first estimate) [Internet]. Statistics Lithuania; 2005 [updated 2012 April 30; cited 2012 May 1]. Available from: http://www.stat.gov.lt/en/pages/view/?id=1867
[2] Global Finance. Lithuania: Country Economic Reports & GDP data [Internet]. Global Finance; 2011 [cited 2012 April 1]. Available from: http://www.gfmag.com/gdp-data-country-reports/231-lithuania-gdp-country-report.html#axzz1tdm8vIK2
[3] Statistics Lithuania. Gross domestic product first estimate [Internet]. Statistics Lithuania; 2005 [updated 2012 April 30; cited 2012 May 1]. Available from: http://www.stat.gov.lt/en/news/view?id=10755
[4] Project Management Institute. Guide to the project management body of knowledge (PMBOK® Guide). 4th ed. Newtown Square: Project Management Institute; 2008.
[5] Institution of Civil Engineers and the Actuarial Profession. Risk analysis and management for projects (RAMP). 2nd ed. Institution of Civil Engineers and the Actuarial Profession. London: Thomas Telford Ltd; 2005.
[6] Project Management Institute. Construction extension to the PMBOK® Guide. 3rd ed. Newtown Square: Project Management Institute; 2007.

* Corresponding Author

[7] Flanagan R, Norman G, Chapman R. Risk management and construction. 2nd ed. Oxford: Blackwell Pub; 2006.

[8] Mills A. A systematic approach to risk management for construction. St Surv 2001;19(5):245–252.

[9] Oyegoke AS. Construction industry overview in the UK, US, Japan and Finland: a comparative analysis. J Constr Res 2006;7(1/2):13–31.

[10] Pheng LS, Chuan QT. Environmental factors and work performance of project managers in the construction industry. Int J Project Manage 2006;24(1):4–37.

[11] Sterman JD. System dynamics modeling for project management [Internet]. Sloan School of Management, Massachusetts Institute of Technology; 1992 [cited 2012 May 1]. Available from: http://web.mit.edu/jsterman/www/SDG/project.pdf

[12] Uher TE, Loosemore M. Essentials of construction project management. Sidney: University of New South Wales Press; 2004.

[13] Dey PK, Ogunlana SO. Selection and application of risk management tools and techniques for build-operate-transfer projects. Ind Manage Data Syst 2004;104(4):334–346.

[14] Cavignac J. Managing risk in a construction company [Internet]. Construction Business Owner 2009; November [cited 2012 March 10]. Available from: http://www.constructionbusinessowner.com/topics/insurance/construction-insurance/managing-risk-construction-company

[15] Eskesen SD, Tengborg P, Kampmann J, Veicherts TH. Guidelines for tunnelling risk management, International Tunnelling Association, Working Group No. 2-1. Tunn Undergr Sp Tech 2004;19(3):217–237.

[16] Keoki Sears S, Sears GA, Clough RH. Construction project management – A practical guide to field construction management. 5th ed. Hoboken: John Wiley & Sons; 2008.

[17] Dey PK. Issues and challenges of managing projects in India: A case study. In: Budhwar PS, Varma A, editors. Doing business in India: Building research-based practice. New York: Routledge; 2011.

[18] Dey PK, Ogunlana SO. Risk based decision support system for effective implementation of projects. Int J Risk Assess Manage 2002;3(2/3/4): 189–204.

[19] Zou PXW, Zhang G, Wang J. Understanding the key risks in construction projects in China. Int J Project Manage 2007;25(6):601–614.

[20] European Agency for Safety and Health at Work. Report – Prevention of risks in construction in practice [Internet]. Luxembourg: Office for Official Publications of the European Communities; 2004 [cited 2012 March 1]. Available from: http://osha.europa.eu/en/publications/reports/108/view

[21] Statistics Lithuania. Construction statistics [Internet]. Statistics Lithuania; 2005 [updated 2012 February 14; cited 2012 May 1]. Available from: http://www.stat.gov.lt/en/news/view/?id=10706

[22] State Labour Inspectorate of the Republic of Lithuania. Annual report on the state on employees' health and safety and labour law enforcement in Lithuanian companies, institutions and organizations [Internet]. Vilnius: State Labour Inspectorate of the

Republic of Lithuania; 2011 [cited 2012 May 1]. Available from: http://www.vdi.lt/index.php?1716170122

[23] Ahmed A, Kayis B, Amornsawadwatana S. A review of techniques for risk management in projects. Benchmark Int J 2007;14(1):22–36.

[24] Cretu O, Stewart RB, Berends T. Risk management for design and construction (RSMeans). Hoboken: John Wiley & Sons; 2011.

[25] Chapman C, Ward S. Project risk management: Process, techniques and insights. 2nd ed. Chichester: John Wiley & Sons; 2003.

[26] Klemetti A. Risk management in construction project networks. Helsinki University of Technology, Laboratory of Industrial Management, Report 2006/2, Espoo; 2006.

[27] Smith NJ, Merna T, Jobling P. Managing risk: in construction projects. 2nd ed. Oxford: Blackwell Publishing; 2006.

[28] Wysocki RK. Effective project management: traditional, agile, extreme. Indianapolis: John Wiley & Sons; 2009.

[29] Wang MT, Chou HY. Risk allocation and risk handling of highway projects in Taiwan. J Manage Eng 2003;19(2):60–68.

[30] Carbone TA, Tippett DD. Project risk management using the project risk FMEA. Eng Manage J 2004;16(4):28–35.

[31] Raz Z, Shenhar AJ, Dvir D. Risk management, project success and technological uncertainty. R&D Manage 2002;32(2):101–109.

[32] Tah JHM; Carr V. A proposal for construction project risk assessment using fuzzy logic. Constr Manage Econ 2000;18(4):491–500.

[33] Wang SQ, Dulaimi MF, Aguria MY.): Risk management framework for construction projects in developing countries. Constr Manage Econ 2004;22(3):237–252.

[34] Adams FK. Construction contract risk management: a study of practices in the United Kingdom. Cost Eng 2008;50(1):22–33.

[35] Caltrans Office of Statewide Project Management Improvement. Project risk management handbook: Threats and opportunities. 2nd ed. [Internet]. Sacramento: Caltrans; 2007 [cited 2012 May 1]. Available from: http://www.dot.ca.gov/hq/projmgmt/documents/prmhb/caltrans_project_risk_management_handbook_20070502.pdf

[36] Carr V, Tah JHM. A fuzzy approach to construction project risk assessment and analysis: construction project risk management system. Adv Eng Softw 2001;32(10–11):847–857.

[37] Modarres M. Risk analysis in engineering – techniques, tools, and trends. 1st ed. Boka Raton: CRC Press; 2006.

[38] Zwikael O, Sadeh A. Planning effort as an effective risk management tool. J Oper Manag 2007; 25(4):755–767.

[39] Hillson D. Developing effective risk responses. In Proceedings of the 30th Annual Project Management Institute Seminars & Symposium, 10-16 October, 1999, Sylva. NC: Project Management Institute; 2000.

[40] Perera BAKS, Dhanasinghe I, Rameezdeen R. Risk management in road construction: the case of Sri Lanka. Int J St Prop Manage 2009;13(2):87–102.

[41] Jannadi OA, Almishari S. Risk assessment in construction. J Constr Eng Managet 2003;129(5): 492–500.
[42] Simu K. Risk management in small construction projects [Licentiate Dissertation]. Department of Civil and Environmental Engineering, Luleå: LTU; 2006.
[43] Baloi D, Price ADF. Evaluation of global risk factors affecting cost performance in Mozambique [Internet]. London: RICS Foundation, COBRA; 2001 [cited 2012 March 11]. Available from:
http://www.rics.org/site/download_feed.aspx?fileID=2483&fileExtension=PDF
[44] Oyegoke AS, Khalfan MMA, McDermott P, Dickinson M. Managing risk and uncertainty in an agile construction environment: application of agile building specialist model. Int J Agile Syst Manage 2008;3(3–4):248–262.
[45] Assaf SA, Al-Hejji S. Causes of delay in large construction projects. Int J Proj Manage 2006;24(4):349–357.
[46] Aibinu AA, Odeyinka HA. Construction delays and their causative factors in Nigeria. J Constr Eng Manage 2006;132(7):667–677.
[47] Nasir D, McCabe B, Hartono L. Evaluating risk in construction-schedule model (ERIC-S): construction schedule risk model. J Constr Eng Manage 2003;129(5):518–527.
[48] Aramvareekul P, Seider DJ. Cost-time-risk diagram: project planning and management. Cost Eng 2006;48(11):12–18.
[49] European Commission. FWC sector competitiveness studies N° B1/ENTR/06/054 – Sustainable competitiveness of the construction sector. Final report [Internet]. ECORYS SCS Group; November 2011 [cited 2012 January 21]. Available from: http://ec.europa.eu/enterprise/sectors/construction/files/compet/sustainable_competitiveness/ecorys-final-report_en.pdf
[50] Lewicki RJ, Bunker BB. Developing and maintaining trust in work relationships. In: Kramer RM, Tyler TR, editors. Trust in organizations: frontiers of theory and research. Thousand Oaks: Sage Publications; 1996.
[51] Pinto JK, Slevin DP, English B. Trust in projects: An empirical assessment of owner/contractor relation-ships., Int J Proj Manage 2009; 27(6):638–648.
[52] Krane HP, Olsson NOE, Rolstadås A. How project manager–project owner interaction can work within and influence project risk management. Proj Manage J 2012;43(2):54–67.
[53] Kaklauskas A, Zavadskas EK, Bagdonavicius A, Kelpsiene L, Bardauskiene D, Kutut V. Conceptual modelling of construction and real estate crisis with emphasis on comparative qualitative aspects description. Transformations in Business & Economics 2010;9,1(19): 42–61.
[54] Chan APC, Chan DWM, Ho KSK. Partnering in construction: critical study of problems for implementation. J Manage Eng 2003;19:126–135.
[55] Brewer G, Strahorn S. Trust and the Project Management Body of Knowledge. Eng Constr Arch Manage 2012;19(3):286–305.
[56] Ward S, Chapman C. Stakeholders and uncertainty management in projects. Constr Manage Econ 2008;26(6):563–577.

[57] Smyth H, Gustafsson M, Ganskau E. The value of trust in project business. Int J Proj Manage 2010;28(2): 117–129.
[58] Wilkinson S. An analysis of the problems faced by project management companies managing construction projects. Eng Constr Arch Manage 2001;8(3):160–170.
[59] Öztas A, Ökmen Ö. Risk analysis in fixed-price design-build construction projects. Build Environ 2004;39(2):229–237.
[60] Zaghloul R, Hartman F. Construction contracts: the cost of mistrust. Int J Proj Manage 2003;21(6):419–424.
[61] Chapman C, Ward S. Developing and implementing a balanced incentive and risk sharing contract. Constr Manage Econ 2008;26(6):659–669.
[62] Ökmen Ö, Öztas A. Construction cost analysis under uncertainty with correlated cost risk analysis mode. Constr Manage Econ 2010;28(2):203–212.
[63] Baloi D, Price ADF. Modeling global risk factors affecting construction cost performance. Int J Proj Manage 2003;21(4):261–269.
[64] Ismail A, Abbas MA, Zamri BC. Approach to analyze risk factors for construction projects utilizing fuzzy logic. J Appl Sci 2008,8(20):3738–3742.
[65] Zeng J, An M, Smith NJ. Application of a fuzzy based decision making methodology to construction project risk assessment. Int J Proj Manage 2007;25(6):589–600.
[66] Kartam NA, Kartam SA. Risk and its management in the Kuwaiti construction industry: a contractors' perspective. Int J Proj Manage 2001;9(6):325–335.
[67] Lahdenperä P. Phased multi-target areal development competitions: algorithms for competitor allocation. Int J St Prop Manage 2009;13(1):1–22.
[68] Majamaa W, Junnila S, Doloi H, Niemistö E. End-user oriented public-private partnerships in real estate industry. Int J St Prop Manage 2008;12(1):1–17.
[69] Mbachu J, Nkado R. Factors constraining successful building project implementation in South Africa. Constr Manage Econ 2007;25(1):39–54.
[70] Mitkus S, Trinkuniene E. Reasoned decisions in construction contracts evaluation. Technol Econ Dev Ec 2008;14(3):402–416.
[71] Yang J, Shen GQ, Ho M, Drew DS, Chan APC. Exploring critical success factors for stakeholder management in construction projects. J Civil Eng Manage 2009;15(4):337–348.
[72] Savić S, Vučković L. Forming of reliability model in order to evaluate the operator's activity risk. Series: Working and Living Environmental Protection 2004;2(4):259–265.
[73] Zavadskas EK, Kaklauskas A, Banaitiene N. Multiple criteria analysis of a building life cycle. Vilnius: Technika; 2001. (in Lithuanian)

Chapter 20

Sources of Risk and Risk Management Strategies: The Case of Smallholder Farmers in a Developing Economy

Satit Aditto, Christopher Gan and Gilbert V. Nartea

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50392

1. Introduction

Risk in agriculture is pervasive and complex, especially in agricultural production.[1, 2] Farmers confront a variety of yields, unstable output and input prices and radical changes in production technology as inherent in their farming operations. These affect the fluctuation in farm profitability from season to season and from one year to another.[3, 4] The sources of risk and level of its severity can vary according to the farming systems, geographic location, weather conditions, supporting government policies and farm types. Risk is a major concern in developing countries where farmers have imperfect information to forecast things such as farm input prices, product prices, and weather conditions, that might impact the farms in the future.[2, 5, 6] The types and severity of risks that farmers face differ from place to place. Incorporating and understanding the effects of risk at the farm level will benefit policy makers who develop appropriate strategies that can help farmers survive the numerous risks they confront.

Sources of risk in agriculture are classified into *business risk* and *financial risk*.[1, 7] Business risks can be classified further into a) production or yield risk, b) marketing or price risk, c) institution, policy, and legal risk, d) human or personal risk, and e) technological risk. On the other hand, financial risk occurs when farmers borrow to finance farm activities as farmers often face variations in interest rates on borrowed funds, inadequacy of cash flow for debt payments and changes in credit terms and conditions.[8, 9]

For several decades, agricultural production in Thailand has faced many risks such as variability in yields, product-prices and cost of inputs.[10-12] Thai farmers typically grow their crops in rain-fed conditions due to poor irrigation systems.[13] The annual rainfall fluctuates widely each year, and pests, diseases and poor soil fertility affect the yields of cash crops in

Thailand. In addition, agricultural commodity prices rise and fall annually depending on the demand and supply in both local and international markets, which are out of the farmer's control. Similarly, the costs of farm inputs also vary each year and may negatively affect farm production costs.

Agriculture contributes approximately 7.86 per cent to Thailand's GDP and 8.98 per cent to exports in 2008.(14, 15) However, large numbers of farmers in rural Thailand still live below the poverty line. In 2007, Thai farm households earned an average income of 129,236 baht/year (US$ 3,692) but only 39 per cent or 50,370 baht/year (US$ 1,439) is from farm activities.(16) Thai farmers are basically smallholders and the national farm size is approximately 7.72 acres.(17) Most farmers have limited diversification potential, face resource problems, environmental variability, lack of soil fertility and water shortages especially smallholder farmers in the north-east region.(18) In addition, smallholder farmers in Thailand also face various sources of risk that vary both seasonally and annually.

Knowledge of the characteristics of risks that influence smallholder farmers is the key to developing appropriate strategies to deal with risks. However, empirical studies on farmers' responses to risks and how risk affects farmers' income, especially in rural Thailand are limited. The aim of this chapter is to examine the sources of risk for smallholder farmers in the central and northeast regions of Thailand and their risk management strategies. We will also relate the farmers' socioeconomic characteristics to their perceived sources of risk and their favoured risk management strategies to gain a deeper understanding of their choices.

2. Sources of risk and risk management strategies on farm

There is much literature on risk sources that impact farming operations and their risk management strategies. Flaten et al. argued that the assessment of farmers' perceptions and how they respond to risk are very important because this can describe the decision making behaviour of farmers when faced with risky situations.(19) Similarly, Hardaker et al. states that *"the welfare of the farm family and the survival of farm business may depend on how well farming risks are managed"*.(1)

The lack of relevant information on farmers' risk perceptions and their risk behaviour present a challenging task for policy makers and researchers who want to create a proper risk management system to help farmers.(19, 20) Extant literature shows that there is no agreement about the most appropriate methods to describe sources of risk and risk responses on farms. However, the Likert-scale rating method has been regularly applied in previous research. In most of those studies, the respondents were asked to rate the sources of risk that affected their farm and the risk management strategies they used on a five-point scale (where 1 is not particularly important and 5 is highly important).

Boggess, Anaman, and Hanson examined farmers' awareness of risk in crop and livestock production in northern Florida and southern Alabama.(21) The respondents were asked to define risk and then to rank the sources of risk and risk management strategies based on how important each risk was to their farm. The results showed that most respondents defined risk as the probability of a negative outcome. The respondents ranked rainfall variability, pests and diseases, and crop price variability as the primary sources of risk for crop production. Livestock price and weather variability and livestock diseases were perceived as important sources of risk for livestock production.

Patrick, Wilson, Barry, Boggess and Young studied farmer attitudes towards risk and risk management among mixed crop and livestock farmers in the US.[22] A total of 149 farmers in 12 states were interviewed. The respondents were grouped into five types of farm; mixed farming; cotton; corn, soybean and hogs; small grain and ranch. The results showed that changes in weather, output price and input costs were rated as the three most important sources of risk in both crop and livestock production.

A nationwide mail survey was used to examine the sources of risk and the risk management strategies of New Zealand farmers by Martin[23]. The survey covered eight types of farm including sheep and beef, dairy, deer, pip fruit, kiwifruit, cropping, vegetables and flowers. The results showed that marketing risk (such as change in product prices and change in input costs) was ranked as a very important source of risk by all farmers. Conversely, production risks (such as rainfall variability, weather, and pests and diseases) were regarded differently depending on geographical location, farm type and product.

Pellegrino studied rice farmers' perceptions of the sources of risk and risk management responses in Argentina.[9] Using size of the respondents' farms as large, medium, and small farms, the author argued that a farmer's awareness of the sources of risk varied depending on farm size. The small size farm group tended to have a higher awareness of production risks than the other two groups.

Meuwissen, Huirne, and Hardaker identified price and production risks as the most important sources of risk for livestock farmers in the Netherlands.[24] An insurance scheme was rated as the appropriate strategy to manage risk. Flaten et al. compared risk perception and the risk responses of conventional and organic dairy farmers in Norway.[19] The results revealed that the institutional (such as government support policies) and marketing risks were classified as the principal sources of risk for the organic dairy farmers. The authors ranked production cost variability and animal welfare policy as the greatest worries for conventional dairy farmers.

Hall, Knight, Coble, Baquet and Patrick found severe drought and meat price variability as the primary sources of risk perceived amongst cattle farmers in Texas and Nebraska.[25] In a recent study, large-scale South African sugarcane farmers perceived land reform regulations, labour legislation and crop price variability as the three most important risk factors.[20]

In terms of risk management strategies, Boggess et al. and Patrick et al. reported that 'placing of investments', 'obtaining market information' and 'enterprise diversification' were the most important strategies that the sampled crop and livestock farmers use to handle risk in the US.[21, 22] Meuwissen et al. found that 'cost of production' and 'insurance schemes' were regarded as important risk strategies among livestock farmers in the Netherlands.[24] Similarly, Flaten et al. noted that organic and conventional dairy farmers in

Norway perceived 'increasing farm liquidity', 'disease prevention', 'buying farm insurance' and 'cost of production' as the most important strategies used to deal with risk on their farms.[19] On the other hand, New Zealand farmers used a mix of risk management strategies to reduce risk. The strategies varied among the groups of farmers depending on the nature of the product, market structure and conditions, farmer characteristics, dynamic risk adjustment considerations and the regulatory situation.[23]

Despite the fact that the evaluation of farmers' risk perceptions and risk management responses are essential to better understand their risk behaviour and managerial decisions, few studies have explicitly investigated awareness of risk among Thai farmers. Akasinha, Ngamsomsuk, Thongngam, Sinchaikul and Ngamsomsuk examined risk perceptions among rice farmers in Payao and Lampang provinces in the northern region.[26] In their study, the Participatory Risk Mapping (PRM) technique was used to elicit sources of risk. The authors' results showed that rice farmers in Payao faced five major sources of risk including 'outbreak of rice disease', 'insects causing damage to rice', 'high input costs', 'flooding', and 'shortage of water supply'. Farmers in Lampang typically faced 'drought', 'insects causing damage to rice', 'low output prices', 'pests', and 'high input costs'.

3. Data and methodology

The sources of risk and their preferred risk management strategies are obtained from face-to-face interviews of 800 farmers, 400 each in the central and northeast regions of Thailand. The central and north-east regions differ in terms of resources, economic development and income distribution. The central region has a farming area of 8.61 million acres or 19.2 per cent of the total farming area. In 2007, the average monthly income per farm in this region is 15,271 baht.[16] The central region is known as the 'rice bowl of Thailand' and more than half of the country's irrigation systems are located in this region known for wet-rice cultivation.[27] In contrast, the north-east region is defined as the 'poorest region' with a long dry season and an annual rainfall that fluctuates widely each year.[13, 18] Approximately 45 per cent of the total farming area in Thailand is located in this region. In 2007, the average monthly income per farm in this region is 8,344 baht.[16]

A smallholder farmer is defined as a farmer who has a farming area less than 30 rai (4.8 ha). Purposive random sampling was employed to classify a particular group of respondents from a certain portion of the population. The sample selection process is as follows. First, the provinces in each region were separated into two main groups: (a) the provinces with large and medium irrigation systems and (b) the provinces in the rain-fed area. Second, purposive sampling was employed to select smallholder farmers in each group. This procedure ensured that the sample covered smallholder farmers of both the irrigated and rain-fed areas in the central and north-east regions.

The information on the sources of risk and risk management strategies perceptions obtained from the respondents using a five-point Likert scale were analyzed in two steps. First, exploratory factor analysis (EFA) was used to capture the information on the interrelationships among the set of variables. This technique enabled the researcher to manage and reduce the number of original variables into a smaller group of new correlation dimensions (factors), which are linear combinations of the original variables.[28, 29] The Kaiser-Meyer-Oklin (KMO) method measured the appropriateness for factor analysis of both data sets. The KMO index varies from 0 to 1, with results of 0.6 or greater suitable for factor analysis. The latent root criterion (eigenvalue > 1) was estimated to identify how many factors in each data set to extract. After the number of factors had been identified, the orthogonal (varimax) rotational method was performed in order to minimize the number of variables that have high loadings on each factor. A factor loading of ± 0.4 was employed as a cut off criterion to determine the inter correlation among the original variables. In addition, Cronbach Alpha was employed to evaluate the internal consistency of each factor.[28]

The relationships between the socioeconomic variables and the perception of risk sources and risk management strategies of the smallholder farmers were also analyzed. Multiple regression was employed to evaluate the influence of farm and farmer characteristics on the smallholder farmers' risk perception and risk management responses. Diagnostic tests were carried out to verify that there was no violation of the multiple regression assumptions. The model specification for the farmer's perception of risk source with socioeconomic variables is postulated as follows:

$$S_i = b_0 + b_1 AGE + b_2 GEN + b_3 EDU + b_4 EXP + b_5 OFFW + b_6 FSIZ + \\ + b_7 INCM + b_8 LOC + b_9 FINC + b_{10} AHIN + b_{11} HSIZ + e \tag{1}$$

The model for risk management responses with socioeconomic variables is given as follows:

$$R_i = b_0 + b_1 AGE + b_2 GEN + b_3 EDU + b_4 EXP + b_5 OFFW + b_6 FSIZ \\ + b_7 INCM + b_8 LOC + b_9 FINC + b_{10} AHIN + b_{11} HSIZ + e \tag{2}$$

where:
Si is source of risk i (from factor analysis); Ri is risk management strategy *i* (from factor analysis); *AGE*= 1, if the respondent's age is over 40 years old, 0 otherwise; *GEN*= 1, if the respondent is male, 0 if female;*EDU* = 1, if the highest education of the respondent is high school and higher, 0 if primary school education or less; *EXP*= 1, if the farming experience is over 30 years, 0 otherwise; *OFFW*= 1, if the respondent has off-farm work, 0 if no off-farm work; *FSIZ* is farm size; *INCM* is net farm income; *LOC*= 1, if the respondent's farm is located in central region, 0 if a farm located in north-east region; *FINC*= 1, if farm has a loan, 0 if farm without a loan; *AHIN*= 1, if the annual household income greater than 90,001 baht, 0 otherwise; *HSIZ* is household size; and e is error term.

4. Results and discussion

4.1. Socioeconomic characteristics of the farmers

The household and farm characteristics of the central and north-east region farmers are presented in Table 1. Table 1 shows that except for gender, household size and finance used for the farm business, central and north-east region farmers generally differ in terms of personal and farm characteristics, and income distribution. The age group distribution indicates that the majority of the farmers in both regions were over 40 years old. Around 40 per cent of the north-east region farmers were over 60 years old, whereas 42 per cent of the central region farmers were between 41-50 years old. The age distribution between the farmers in both regions was significantly different with the north-east region farmers more likely to be older than the central region farmers. Nearly half of the farmers in the north-east had been involved in agricultural work for over 40 years which implies that younger farmers are rare especially in the north-east. This may be a result of the rural-to-urban migration problems in Thailand.

Around 75 per cent of the farmers in both regions graduated with a primary education and about three per cent were illiterate. The result indicates that the central region farmers had higher levels of education than the north-east farmers ($P<0.01$). Mustafa argued that the educational level of farmers affected their decision making capacity.[30] A higher educated farmer was expected to perform better than an uneducated farmer in terms of management skills and farm resource allocation to maximize farm profitability.

The average farm size of the farmers in the central region was 21.40 rai (3.42 ha) of which 30 per cent was self-lease operated. In contrast, farmers in the north-east had an average farm size of 14.80 rai (2.37 ha) of which 90 per cent was self owned. This result indicates that the central region farmers hold average farm sizes larger than north-east farmers ($P<0.01$). This is consistent with the Office of Agricultural Economics who reported that farmers in the central region usually had an average farm size larger than the north-east farmers.[16]

The results for the average net farm income between the farmers in the central and north-east regions were statistically significant at the one per cent level. This result indicated that the average net farm income of the central farmers was larger than for the north-east farmers. In 2008, the central farmers had an average net farm income of 166,445.05 baht/household, whereas the average net farm income of the north-east farmers was only 42,632.80 baht/household.

In addition, approximately 63 per cent of the central region farmers worked off-farm, which was significantly more than for the north-east farmers ($P<0.01$). The results also showed that central farmers had significantly higher annual household incomes than north-east farmers.

In terms of farmer access to credit, nearly 70 per cent of the farmers in the central and north-east regions had loans and nearly half of them borrowed from the Bank of Agriculture and Agricultural Cooperatives. In addition, eight per cent of the farmers used their own savings to operate their farm business. Only about four per cent had loans from commercial banks. The majority of the farmers obtained short-term loans (see Table 2). This finding supports Limsombunchai, who argued that smallholder farmers in rural Thailand lacked investment funds due to a credit accessibility barrier.[31]

Nearly 50 per cent of the farmers had small debts. Further, 30 per cent of the farmers in the north-east had outstanding debts of less than 30,000 baht during the 2008 crop year. Similarly, 27 per cent of the farmers in the central region had debts between 31,000-50,000 baht. An average of 72.6 per cent of the loans were used in operating the farm business, such as purchasing farm equipment, seeds and fertilizers, but the balance was spent on the farmer's personal and household consumption, for example, food and clothing.

Item	*Unit*	*Region*		*Overall (n=800)*	*Test of differrence* [a]
		Central (n=400)	*North-east (n=400)*		
Gender	%				0.66
Male		73.3	75.8	74.5	
Female		26.8	24.3	25.5	
Age group	%				67.14***
Less than 30 years old		1.5	0.5	1.0	
31-40 years old		10.3	7.3	8.8	
41-50 years old		42.0	22.3	32.1	
51-60 years old		30.0	30.5	30.3	
Over 60 years old		16.3	39.5	27.9	
Marital status	%				12.52***
Single/Never married		4.0	2.0	3.0	
Married		87.5	86.3	86.9	
De factor relationship		0.8	4.3	2.5	
Divorced/separated		7.8	7.5	7.6	
Highest education	%				17.79***
Illiterate		3.3	2.0	2.6	
Primary school		69.5	81.8	75.6	
Secondary school		23.5	14.0	18.8	
Vocational training		2.3	0.8	1.5	
Bachelor degree		1.5	1.5	1.5	
Farming experience	%				105.69***
Less than 10 years		12.8	6.5	9.6	
11-20 years		29.3	10.0	19.6	
21-30 years		22.5	16.0	19.3	
31-40 years		19.5	22.8	21.1	
Over 40 years		16.0	44.8	30.4	
Household size	Persons	4.36	4.28	4.32	-0.66
Total farm size	rai [b]	21.40	14.80	18.09	-10.10***
Land ownership status	%				168.93***
Owner-self operated		64.8	89.5	77.1	
Lease-self operated		29.3	2.0	15.6	
Tenant		0	8.5	4.3	
Other		6.0	0	3.0	
Finance farm business	%				0.15
Yes		69.3	68.0	68.6	
Average net farm income [c]	baht	166,450	42,632	104,541	-19.26***
Working off-farm	%				43.29***
Yes		63.3	40.0	51.6	
Annual household income	%				113.16***
Less than 10,000 baht		0	1.3	0.6	

Item	Unit	Region		Overall (n=800)	Test of differrence [a]
		Central (n=400)	North-east (n=400)		
10,001-30,000 baht		0.8	14.3	7.5	
30,001-50,000 baht		5.0	16.3	10.6	
50,001-70,000 baht		11.0	15.8	13.4	
70,001-90,000 baht		11.5	11.0	11.3	
More than 90,001 baht		71.8	41.5	56.6	

[a] Test of differences of the central and north-east household and farm characteristics based on chi-square and independent *t* test; * significant at 10%, ** significant at 5% and *** significant at 1%
[b] 1 rai = 0.16 ha. [c] Net farm income is based on the 2008 crop year.
Source: Field survey, 2009

Table 1. Household and farm characteristics of the farmers in central and north-east Thailand

Item	Region		Overall (n=800)
	Central (n=400)	North-east (n=400)	
Sources of finance [a]			
Bank of Agriculture and Agricultural Cooperative	57.5	34.3	44.6
Cooperatives	23.7	15.9	19.4
Village funds	11.8	25.4	19.4
Personal funds	3.2	12.4	8.3
Commercial bank	6.8	0.3	3.9
Duration of credit			
Less than 1 year	65.0	72.8	68.9
Greater than 3 years	6.9	20.6	13.7
Outstanding loan debt			
Under 30,000 baht	14.4	29.4	21.9
31,000-50,000 baht	27.4	23.2	25.3
Over 91,000 baht	13.7	21.0	17.3
Average percentage of loan used			
On-farm activities	79.8	65.1	72.6
Household expenses	20.1	34.7	27.4

[a] Multiple responses
Source: Field survey, 2009

Table 2. Financial background of the farmers in central and north-east Thailand

4.2. Farmers' perceptions of sources of risk and risk management strategies

4.2.1. Sources of risk

The mean scores of each source of risk were ranked and the standard deviation (SD) was used to indicate the variation in the ratings. Independent sample t-test was employed to compare mean score differences between the farmers in the central and north-east regions.

Table 3 summarizes the results of the most important perceived sources of risk for the farmers in the central and north-east regions. The table shows that marketing risks associated with 'unexpected variability of input prices' and 'unexpected variability of product prices' had the highest and second highest mean scores for sources risk, respectively rated by the farmers in both regions. The SDs of both sources of risk in each group were less than one and this indicates that those sources of risk gained a high level of consensus among the farmers in both regions.(24)

Source of risk	*Overall (n=800)*			*Central (n=400)*			*North-east (n=400)*			*Test of diff.* [b]
	Mean [a]	*SD*	*Rank*	*Mean* [a]	*SD*	*Rank*	*Mean* [a]	*SD*	*Rank*	
Unexpected variability of input prices	4.22	0.910	*(1)*	4.09	0.901	*(1)*	4.34	0.904	*(1)*	3.92***
Unexpected variability of product prices	3.82	0.926	*(2)*	3.83	0.861	*(2)*	3.82	0.988	*(2)*	-0.11
Diseases and pests that affect plants and animals	3.52	1.153	*(3)*	3.70	1.014	*(3)*	3.34	1.252	*(3)*	-4.47***
Changes in Thailand's economic and political situation	3.48	1.080	*(4)*	3.44	0.992	*(4)*	3.53	1.161	*(4)*	1.28
Unexpected variability of yields	3.47	0.946	*(5)*	3.58	0.965	*(5)*	3.36	0.915	*(5)*	-3.35***
Changes in national government laws and policies	3.38	1.090	*(6)*	3.38	1.024	*(6)*	3.39	1.154	*(6)*	0.16
Natural disasters such as heat, fire, flood, storm	3.38	1.345	*(7)*	3.47	1.092	*(7)*	3.29	1.554	*(7)*	-1.92*
Changes in the world economic and political situation	3.30	1.097	*(8)*	3.27	1.029	*(8)*	3.32	1.161	*(8)*	0.71
Excess rainfall	3.27	1.293	*(9)*	3.59	1.017	*(9)*	2.95	1.453	*(9)*	-7.16***
Deficiency in rainfall causing drought	3.11	1.441	*(10)*	3.09	1.372	*(10)*	3.13	1.508	*(10)*	0.44
Problems with hired labour	3.02	1.259		2.95	1.161		3.10	1.347		1.72*
High level of debt	2.84	1.075		2.90	1.052		2.77	1.095		-1.75*
Accidents or problems with health	2.74	1.145		2.56	1.007		2.91	1.245		4.34***
Changes in interest rates	2.73	1.106		2.86	1.054		2.60	1.144		-3.28***

Source of risk	Overall (n=800)			Central (n=400)			North-east (n=400)			Test of diff. [b]
	Mean [a]	SD	Rank	Mean [a]	SD	Rank	Mean [a]	SD	Rank	
Changes in technology and breeding	2.52	1.089		2.49	0.952		2.55	1.211		0.75
Changes in land prices	2.47	1.222		2.56	1.241		2.38	1.198		-2.03**
Risk from theft	2.19	1.179		2.57	1.144		1.82	1.094		-9.44***
Changes in family situation such as marital status, inheritances, etc.	1.98	1.032		2.11	0.966		1.85	1.081		-3.52***
Being unable to meet contracting obligations	1.82	1.046		2.13	1.038		1.52	0.965		-8.50***

[a] Likert scale is used from 1 (not important) to 5 (extremely important).
[b] The mean scores of central and north-east farmers are significantly difference at *$P<0.1$, **$P<0.05$ and ***$P<0.01$ based on independent samples *t* test.
Source: Field survey, 2009

Table 3. Ranking of perceptions of sources of risk by sampled farmers in central and north-east Thailand

The survey results showed that the uncertainty of input prices and product prices have become increasingly worrying among smallholder farmers in the central and north-east regions. This is probably due to the fact that both sources of risk are out of the farmers' control but directly affect their farm incomes. The prices of the major cash crops in Thailand, such as rice, cassava and sugarcane, are unstable; they depend on supply and demand in both local and international markets. Similarly, the average prices of the major farm inputs such as fertilizer NPK 16-20-0, which is widely used by rice farmers, fluctuated from 9,485 baht/tonne in 2006 to a peak of 19,386 baht/tonne in 2008 and then dropped to 16,199 baht/tonne in 2009.(16)

This finding is consistent with those of Patrick et al., Martin, and Flaten et al. who argued that marketing risks associated with the variability of product and input prices were the most important sources of risk considered by the farmers in their respective study areas.(19, 22, 23)

The production risks related to 'diseases and pests affecting plants and animals', 'excess rainfall' and 'natural disasters such as floods' were ranked third, fourth and sixth, among the farmers in the central region with mean scores of 3.70, 3.59 and 3.47, respectively. The results reflect the heavy floods that inundated the central provinces during September 2008. Following this incident, 100,000 rai (16,000 ha.) of farmland in the central region were damaged.(32)

Institutional risks related to 'changes in Thailand's economic and political situation' and 'changes in national government laws and policies' were ranked third and fourth, among the north-east region farmers, respectively. This finding revealed that smallholder farmers were concerned about the effect of the political conflicts in Thailand on their farm operation.

'Unexpected variability of yields' was ranked the fifth most important source of risk in both regions. In addition, the financial risks associated with 'changes in interest rates' and 'high levels of debt' were considered as 'quite important' by all farmers.

Sources of risk that obtained low mean scores included 'changes in technology and breeding', 'changes in land prices', 'risk from theft', 'changes in the situation of farm families' and 'unable to meet contracting obligations'.

Comparisons of risk perception between the farmers in the central and north-east regions showed significant differences in most sources of risk. This interesting finding might be attributable to the fact that sources of risk vary depending on the farm's geographical condition, farm type, the environmental impact and the country's political and economic situation. Evidently, the small farm business may be affected in different ways by changes in these sources of risk.

4.2.2. Risk management strategies

Table 4 summarizes the results of the perceptions of risk management strategies elicited from the farmers in the central and north-east regions. Production and financial strategies were considered more important managerial responses to risk than marketing strategies by the farmers in both regions.

Among the production strategies perceived by the central region farmers, 'purchase farm machinery to replace labour' was the most important with an average rating of 3.45. Nearly 60 per cent of central region farmers reported using this strategy to cope with hired agricultural labour problems on their farms. From the survey, farm machinery, such as hand tractors and four-wheel tractors, was widely used among the central region farmers. This reflects the imbalance problem between agricultural and industrial labour forces in Thailand. This finding supports Ahmad and Isvilanonda who argued that the rural labour force preferred to work in the industrial sector more than in the agricultural sector due to the gap in wage rates.(33) This may be caused by the lack of agricultural labour especially in the central region, which has many factories located there.

'Storing feed and/or seed reserves' and 'have a farm reservoir for water supplies in dry season' showed significant differences in importance between the farmers in the central and north-east regions ($P<0.01$). North-east farmers perceived the importance of these two production strategies higher than central region farmers. They rated 'storing feed and/or seed reserves' as the most important production strategies and 'having a farm reservoir for water supplies in dry season' was ranked third with mean ratings of 3.61 and 3.47, respectively. Over 80 per cent of the north-east farmers preferred 'storing feed and/or seed reserves' in managing their small farm operations and approximately 65 per cent of them preferred using the 'having a farm reservoir for water supplies in dry season' strategy on their farm. This indicates that the north-east farmers were confronted with the variability of input prices and severe droughts.

'Having diversified crop, animal or other enterprises' and 'planting several varieties of crops' were the least important production strategies for both groups. The north-east farmers considered these two production strategies as 'important' but the central region farmers rated them as 'quite important', which is statistically significant different (P<0.01). The results indicated that the lack of farm resources may affect the diversification performance of the farmers in both groups.

Financial strategies associated with 'holding cash and easily converted cash assets' and 'working off farm to supplement household income' were considered 'important' by the farmers in the central and north-east regions. Approximately 60 per cent of the farmers in both regions reported that they used these two financial strategies. However, the north-east farmers perceived the importance of 'holding cash and easily converted cash assets' significantly higher than the central region farmers. In addition, 'reduce debt level' was given greater importance by the north-east farmers, whereas 'investing in non-farm businesses' was more important among the central region farmers. In terms of marketing strategies, north-east farmers assigned significantly greater rating scores than central region farmers to 'obtaining market information', 'spread sale over several time period' and 'selection of crop and/or animal varieties with low price variability'.

Source of risk	*Overall (n=800)*			*Central (n=400)*			*North-east (n=400)*			*Test of diff.* [c]
	Mean [a]	*%* [b]	*Rank*	*Mean* [a]	*%* [b]	*Rank*	*Mean* [a]	*%* [b]	*Rank*	
Production strategies:										
Purchase farm machinery to replace labour	3.44	61.6	*(1)*	3.45	58.8	*(1)*	3.43	64.5	*(5)*	-0.26
Storing feed and/or seed reserves	3.40	60.9	*(3)*	3.20	40.8	*(6)*	3.61	81.0	*(1)*	5.49***
Apply pests and diseases program	3.23	53.9	*(7)*	3.26	53.8	*(4)*	3.19	54.0	*(9)*	-0.89
Have a farm reservoir	3.06	47.9	*(10)*	2.65	35.5		3.47	60.3	*(3)*	9.40***
Having diversified crop, animal or other enterprises	2.94	33.4		2.84	26.0		3.05	40.8		2.65***
Planting several varieties of crops	2.86	30.0		2.71	19.5		3.01	40.5		3.64***
Marketing strategies:										
Obtaining market information	3.27	65.3	*(5)*	3.09	51.8	*(7)*	3.46	78.8	*(4)*	4.89***
Spreading sale over several time periods	3.19	41.6	*(8)*	3.01	31.5	*(9)*	3.39	51.8	*(6)*	4.48***
Selection of crop and/or animal varieties with low price variability	2.70	24.8		2.61	21.0		2.79	28.5		2.46**

Source of risk	Overall (n=800)			Central (n=400)			North-east (n=400)			Test of diff. [c]
	Mean [a]	% [b]	Rank	Mean [a]	% [b]	Rank	Mean [a]	% [b]	Rank	
Use forward contracts	2.13	12.4		2.32	12.3		1.95	12.5		-4.59***
Financial strategies:										
Holding cash	3.41	64.8	*(2)*	3.31	60.0	*(3)*	3.52	69.5	*(2)*	2.98***
Working off farm	3.28	63.3	*(4)*	3.33	68.8	*(2)*	3.24	57.8	*(8)*	-1.07
Reduce debt level	3.27	60.0	*(6)*	3.20	48.5	*(5)*	3.33	71.5	*(7)*	1.73*
Leasing farm machinery	3.13	48.9	*(9)*	3.08	38.5	*(8)*	3.17	59.3	*(10)*	1.17
Investing in non-farm businesses	2.64	31.3		2.92	42.3		2.36	20.3		-6.30***
Miscellaneous strategies:										
Able to adjust quickly to weather, price and other adverse factors	3.02	42.6		2.98	42.0	*(10)*	3.06	43.3		1.18

[a] Likert scale is used from 1 (not important) to 5 (extremely important).
[b] The percentage of farmers using each risk management strategy.
[c] The mean scores of central and north-east farmers are significantly difference at *$P<0.1$, **$P<0.05$ and ***$P<0.01$ based on independent samples *t* test
Source: Field survey, 2009

Table 4. Ranking of perceptions of risk management strategies by sampled farmers in central and north-east Thailand

'Use forward contracts' was the least important marketing strategy considered by most central and north-east regions farmers. Only 10 per cent of the farmers in both regions had used this strategy to manage risk. This suggests that the agricultural production under forward contracts in Thailand is still in its developmental stages and is not popular among the smallholder farmers in rural areas. However, the central region farmers perceived the importance of this marketing strategy significantly more than the north-east farmers (P<0.01) with the mean scores of 2.32 and 1.95, respectively.

The perceptions of risk responses between the farmers in the central and north-east regions were statistically different in many strategies similar to their perceived sources of risk (see Table 4). The findings from the survey revealed that the smallholder farmers in both regions used a mix of risk strategies to manage and reduce the sources of risk they are confronted with. The findings support Martin, who argued that the farmers' selection criteria for risk management strategies varied depending on farm type, climatic conditions, marketing factors and agricultural rules and regulations.(23)

4.3. Factor analysis

In this section, the results of the factor analysis of sources of risk and risk management strategies are discussed. Exploratory factor analysis with varimax orthogonal rotation was applied to the data using SPSS version 15. Exploratory factor analysis is used to reduce the number of sources of risk and risk management strategies for each group of farmers.

4.3.1. Sources of risk

The rotated factor loadings of risk sources for all farmers in the central and north-east regions, obtained from the principal component analysis and a varimax orthogonal rotation, are discussed in this section. The KMO measure of data sufficiency was 0.779 and the Bartlett's Test of Sphericity achieved statistical significance (χ^2 = 4927.58, P<0.01), both indicating that the data set was appropriate for factor analysis. However, the preliminary results indicated three sources of risk including 'accidents or problems with health', 'deficiency rainfall' and 'changes in technology or breeding' should be eliminated from the factor analysis because of their low communalities (<0.40).[28] Following this, iteration of varimax orthogonal rotation was performed.

The results are presented in Table 5. Latent root criteria (eigenvalues > 1) were specified for six factors (AS1-6) from the 16 sources of risk variables for all farmers in both regions. These six factors can explain almost 71.2 per cent of the total variance. The Cronbach's Alpha values for factors AS1-5 ranged from 0.671 to 0.899, which exceeded the minimum requirement of 0.6. This demonstrates an adequate reliability among those factors. However, the alpha value was somewhat lower (0.426) for factor AS6. Factors AS1-6 can be labelled in accordance with the significant loading variables that were obtained for each factor and explained as follows:

Factor AS1: this factor is named 'economic and political' because of the relatively high loadings on the sources of risk variables with the changes in Thailand and the world economic and political situations and changes in the government laws and policies that affected the small farm operations.

Factor AS2: this factor incorporates a number of sources of risk related to the farm business environment, including risk from being unable to meet contracting obligations, problems with hired labour, theft and changes in land prices. Moreover, risk from changes in family situation (also as personal risk) loaded highly on this factor. Therefore, this factor is named 'personal and farm business environment'.

Factor AS3: this factor consists of the significant loading of 'excess rainfall' and 'natural disaster'. Factor AS3 is labelled 'natural disaster'.

Factor AS4: this factor can be interpreted as the 'financial situation' because of the high factor loadings on the changes in interest rates and high level of debt.

Factor AS5: this factor is related to the risk from unexpected variability in yields and the unpredictable product prices. Thus, this factor is classified as 'yields and product prices'.

Factor AS6: this factor is labelled 'input prices' because of the highest factor loading of the unexpected variability in input prices in this factor.

Source of risk	Factors [a]						Commu-nality
	AS1	*AS2*	*AS3*	*AS4*	*AS5*	*AS6*	
Changes in Thailand's economic and political situation	**0.923**	0.091	0.005	0.092	0.134	0.053	0.890
Changes in the world economic and political situation	**0.875**	0.064	0.066	0.164	0.030	0.050	0.804
Changes in national government laws and policies	**0.833**	0.220	0.003	0.048	0.179	0.094	0.786
Changes in family situation	0.087	**0.748**	0.097	0.079	0.126	-0.176	0.629
Being unable to meet contracting obligations	0.009	**0.747**	0.121	0.285	0.042	-0.082	0.663
Risk from theft	0.107	**0.700**	0.078	0.203	0.151	0.108	0.583
Problems with hired labour and contractors	0.132	**0.616**	-0.170	-0.147	-0.127	0.427	0.646
Changes in land prices	0.315	**0.559**	-0.014	0.242	0.107	0.087	0.489
Excess rainfall	0.018	0.050	**0.895**	0.086	0.085	-0.039	0.821
Natural disasters	0.033	0.077	**0.862**	-0.056	-0.007	0.190	0.789
Changes in interest rates	0.119	0.261	-0.024	**0.827**	0.065	0.162	0.797
High level of debt	0.169	0.220	0.070	**0.825**	0.064	0.010	0.768
Unexpected variability of yields	0.141	0.103	0.053	0.071	**0.846**	0.017	0.755
Unexpected variability of product prices	0.131	0.122	0.033	0.046	**0.823**	0.135	0.730
Unexpected variability of input prices	0.077	-0.094	-0.014	0.064	0.115	**0.852**	0.758
Diseases and pests that affect plants and animals	0.073	0.104	0.329	0.135	0.071	**0.579**	0.483
Eigenvalues	4.35	1.83	1.71	1.22	1.21	1.07	
Per cent of total variance explained	27.17	11.46	10.70	7.61	7.55	6.69	
Cumulative per cent of the variance explained	27.17	38.63	49.33	56.95	64.49	71.19	
Cronbach's Alpha	0.889	0.743	0.776	0.763	0.671	0.426	
Number of variables	3	5	2	2	2	2	

[a] Factors AS1-6 are labelled as AS1=economic and political, AS2=personal and farm business environment, AS3=natural disaster, AS4=financial situation, AS5=yields and product prices and AS6=input prices.
'Accidents or problems with health', 'deficiency in rainfall causing drought' and 'changes in technology and breeding' are deleted from the analysis due to these sources of risk have low communalities.
Factor loadings for an absolute value greater than 0.4 are in **bold**.
Source: Field survey, 2009

Table 5. Varimax rotated factor loadings of sources of risk for all sampled in Thailand farmers (n=800)

4.3.2. Risk management strategies

Factor analysis was employed to reduce the risk strategy categories as perceived by the farmers in both the central and north-east regions. The KMO measure of data sufficiency was 0.887. In addition, Bartlett's Test of Sphericity was statistically significant at the one per cent level ($\chi^2 = 3301$). This indicates that the data were suitable for factor analysis.

The first iteration of factor analysis resulted in the removal of 'able to adjust quickly to weather, price and other adverse factors' and 'purchase farm machinery to replace of labour', because these variables exhibited low communalities. Following this, the second rotation was performed with 14 risk strategies.

The final results of the varimax rotated factor loadings for each risk strategy are documented in Table 6. Factor analysis grouped the 14 risk management strategies into four factors. These four factors explained almost 58.33 per cent of the variance.

With regard to reliability, the Cronbach's Alpha values for factors AR1-3 were 0.742, 0.711 and 0.642, respectively. The alpha value for factor AR4 was 0.596, which is very close to the minimum cut-off level of 0.6. The factors AR1-4 can be named according to each factor structure as follows:

Factor one (AR1): this factor has a relatively high loading of the risk strategy variables related to 'apply pests and diseases programme', 'storing feed and/or seed reserves', 'have a farm reservoir for water supplies in dry season', 'spreading sale over several time period' and 'obtaining market information on prices forecast and trends'. This factor is named 'farm production and marketing management'.

Factor two (AR2): this factor is described as 'diversification' because there were significant loadings of risk strategy variables related to 'having diversified crop, animal or other enterprises', 'planting several varieties of crops' and 'selection of crop and/or animal varieties with low price variability'.

Factor three (AR3): this factor is loaded highly on 'investing in non-farm investment/business' and 'working off farm to supplement net farm income', which represent the influence of off-farm income. Thus, factor three is named 'off-farm income'.

Factor four (AR4): this factor is interpreted as 'financial management', which is concerned with 'reduce debt level', 'leasing farm machinery rather than owning them' and 'holding cash and easily converted cash assets'.

Risk management strategy	*Factors* [a]				*Communality*
	AR1	*AR2*	*AR3*	*AR4*	
Apply pests and diseases program	**0.655**	-0.035	0.318	0.047	0.533
Storing feed and/or seed reserves	**0.651**	0.162	-0.025	0.339	0.565
Have a farm reservoir for water supplies in dry season	**0.641**	0.288	0.022	0.031	0.495
Spreading sale over several time period	**0.618**	0.301	0.183	0.159	0.531

Risk management strategy	*Factors* [a]				*Communality*
	AR1	*AR2*	*AR3*	*AR4*	
Obtaining market information on prices forecast and trends	**0.505**	0.363	0.259	0.280	0.532
Having diversified crop, animal or other enterprises	0.211	**0.796**	-0.030	0.147	0.700
Planting several varieties of crops	0.252	**0.742**	0.093	0.095	0.632
Selection of crop and/or animal varieties with low price variability	0.387	**0.505**	0.345	-0.039	0.525
Investing in non-farm investment/business	0.172	-0.001	**0.807**	0.124	0.696
Working off farm to supplement net farm income	0.341	0.058	**0.711**	0.143	0.646
Use forward contracts	-0.121	**0.441**	**0.590**	0.076	0.563
Reduce debt level	0.094	0.117	0.061	**0.787**	0.645
Leasing farm machinery rather than owning them	0.164	-0.023	0.111	**0.715**	0.551
Holding cash and easily converted cash assets	0.117	**0.440**	0.177	**0.559**	0.552
Eigenvalues	4.69	1.28	1.19	1.01	
Per cent of total variance explained	33.48	9.14	8.48	7.24	
Cumulative per cent of the variance explained	33.48	42.62	51.09	58.33	
Cronbach's Alpha	0.742	0.711	0.642	0.596	
Number of variables	5	3	3	3	

[a] Factors AR1-4 labelled as AR1=farm production and marketing management, AR2=diversification, AR3=off-farm income and AR4=financial management.
'Able to adjust quickly to weather, price and other adverse factors' and 'purchase farm machinery to replace of labour' are deleted from the analysis due to these risk management strategies have low communalities.
Factor loadings for an absolute values greater than 0.4 are in **bold**.
Source: Field survey, 2009

Table 6. Varimax rotated factor loadings of risk management strategies for all farmers sampled inThailand (n=800)

However, in factor AR3, factor analysis grouped the 'use forward contracts' variable, which is unrelated to the definition of this factor. Therefore, the 'use forward contracts' variable was deleted from factor AR3 and the Cronbach Alpha coefficient slightly improved from 0.642 to 0.697. This result illustrated that factor AR3 had a stronger internal consistency after 'use forward contracts' variable was deleted.

4.4. The association between the farmers' characteristics and source of risk and management perception of risks

Multiple regression analysis was employed to investigate the relationship between the farmers' socioeconomic characteristics and the perceptions of sources of risk and risk management strategy components obtained from the factor analysis. The summated scales of sources of risk and risk strategy factors of each group of farmers were summed up and averaged based on the relevant variables in each factor structure and their internal consistency. Before performing multiple regression analysis, all models were assessed for normality, linearity, multicollinearity and homoscedasticity to ensure the appropriateness of the equations.[29]

4.4.1. Sources of risk

Table 7 shows the relationship between all farmers' socioeconomic status and the different perceptions of sources of risk components. Models 1-4 are statistically significant at the one per cent level. However, the coefficients of determination (R^2) of most of the models are low. This result is consistent with the findings of Flaten et al. and Meuwissen et al. who found low explanatory power of regression models between the perceptions of sources of risk and risk strategies with the farmers' characteristics.[19, 24] Both authors argued that the lower R^2 in the regression models implies that the farmers' perceptions of sources of risk and risk strategies differed from farmer to farmer.

Gender is negatively related to the 'personal and farm business environment' and 'natural disaster' risks on farm. This implies that female heads of farm households are likely to perceive these sources of risk as significantly more important than male household heads. Similarly, the age of farmers and farm size are negatively related to the 'natural disaster' risk, which means young farmers and farmers who have smaller farm sizes tended to perceive 'natural disaster' as a higher on-farm source of risk. This finding may be attributable to the severe floods across Thailand in 2008.

The highest educational level is positively related to the 'personal and farm business environment' risk, which indicates that more educated farmers perceived this source of risk as significantly more important in farming. The reason is because the more educated farmers realized that the family farm situation and the changes in farm business environment, such as high labour wages and relatively high prices of agricultural land, may indirectly affect their farm operations.

The number of years in farming is negatively related to the 'economic and political' risk perceptions. However, the annual household income and the size of farm household exhibited a positive relationship with this source of risk. This result suggests that less experienced farmers, farmers who have higher annual household income and farmers with larger household size tended to perceive risk related to 'economic and political' as highly important. This finding may have resulted from the instability of Thailand political situation since September 2006.

Farm business finance is positively related to the 'financial situations' risk factor and is statistically significant at the one per cent level. This suggests that farmers who have loans are more likely to pay more attention to the changes to their farm financial situation, such as interest rates and level of debt. In addition, farm business finance is positively related to the 'natural disaster' risk factor. This implies that farmers who have loans perceived this source of risk as highly important. This may be due to the 'natural disaster' risk damaging their farm crops, which results in insecurity of their farm income and debt repayment capacity.

Risks related to the 'economic and political' and 'personal and farm business environment' were perceived as highly important by farmers who had off-farm work. This suggests that farmers who have off-farm work are very concerned about those risks that can disrupt their off-farm income.

With regard to the farm location variable, the regression result showed a strong relationship with more than half of the risk factors. Farmers in the central region perceive the 'personal and farm business environment', 'natural disaster' and 'financial situation' as more important risk factors than north-east farmers; north-east farmers are more concerned about 'economic and political' risk. This finding suggests that the sources of risk on small-holding farms differ significantly between these two regions.

Independent variables	*Risk source components* [b]				
	AS1	*AS2*	*AS3*	*AS4*	*AS5*
Constant	3.170***	1.943***	3.287***	2.466***	3.619***
Age [c]	-0.039	-0.079	-0.306**	-0.056	-0.118
Gender [d]	-0.024	-0.199***	-0.182*	-0.063	-0.056
Highest education [e]	0.068	0.233***	0.123	0.122	0.123
Farming experiences [f]	-0.139*	0.024	0.134	-0.098	0.013
Off-farm work [g]	0.135*	0.281***	0.037	0.067	0.092
Farm size	-0.003	0.005	-0.011**	-0.004	0.001
Net farm income	-2.37E-07	-9.81E-07***	1.35E-06**	-6.90E-07	-2.77E07
Farm location [h]	-0.166*	0.301***	0.313***	0.196**	0.079
Finance farm business [i]	0.028	-0.038	0.294***	0.408***	0.027
Annual household income [j]	0.231***	0.068	0.009	0.130	0.100
Household size	0.063***	0.051***	0.008	0.023	0.001
R^2	0.034***	0.124***	0.064***	0.061***	0.021

[a] Variables and models significant at *P<0.1, **P<0.05 and ***P<0.01;

[b] Factors AS1-6 are labelled as AS1=economic and political, AS2=personal and farm business environment, AS3=natural disaster, AS4=financial situation, AS5=yields and product prices and AS6=input prices;

[c] 1, if the farmer's age over 40 years old, 0 otherwise;

[d] 1, if the farmer is male, 0 if female;

[e] 1, if the highest education of the farmer is high school and higher, 0 if primary school education or less;

[f] 1, if the farming experience over 30 years, 0 otherwise;

[g] 1, if the farmer has off-farm work, 0 if no off-farm work;

[h] 1, if farmer's farm is located in central region, 0 if a farm located in north-east region;

[i] 1, if farm has a loan, 0 if farm without a loan; and

[j] 1, if household income greater than 90,001 baht and 0 represent otherwise.

Source: Field survey, 2009

Table 7. Multivariate regression of the source of risk components and household and farm characteristics of all sampled Thai farmers (n=800) [a]

4.4.2. Risk management strategies

Table 8 summarizes the multiple regression models of the risk management strategy components and the socioeconomic variables for all farmers. The goodness-of-fit coefficients of all models were rather low, except for model three where the coefficient explained around 27 per cent of the variation of the dependent variable. Models 1-4 are statistically significant ($P< 0.01$).The age variable is insignificant in relation to the risk strategy components of all farmers.

Gender was negatively related to 'off-farm income', which means that female household heads perceived this risk strategy as more important than male household heads. The reason is because the female farmers or wives can easily find off-farm work, such as weaving and/or handicrafts that are widely found throughout the north-east region, to supplement their household income.

The highest educational level was positively related to the 'farm production and marketing management', 'diversification' and 'off-farm income' risk strategies. This implies that the more educated farmers perceived these risk management strategies as highly important. This finding is similar to that of Mustafa who argued that the more educated farmers performed better in managing their farm business compared with less educated farmers.[30]

The length of farming experience was negatively related to the 'farm production and marketing management', 'diversification' and 'financial management' risk strategies. This suggests that less experienced farmers were more likely to be interested in employing these strategies to manage risk on their farms than the more experienced farmers.

Off-farm work was positively related to all four risk strategy components. These relationships may be due to the farmers who have off-farm work to enhance their farm income; they are willing to adopt such strategies to improve and maintain their farm income. Similarly, the net farm income coefficient shows a negative relationship with all four risk strategy components. This suggests that the farmers who have a lower net farm income believe that these risk strategies can help to increase their farm income.

Farm size was positively related to the 'diversification' strategy. Farmers with larger farms perceived a diversification strategy as highly important. It should be noted that farm size is one of the constraints to diversification, that is, farmers with a small holding have limited ability to diversify their farm activities.[33]

Farmers who had higher annual household incomes perceived the 'financial management' strategy as highly important. In contrast, they perceived the 'diversification' strategy as less important than farmers who had lower annual income. In addition, risk management strategies related to 'farm production and marketing management' and 'off-farm income' were perceived as less important by the farmers who had loans. Farmers with larger households perceived 'farm production and marketing management' as slightly more important than smaller household farmers.

The farm location coefficient was negatively related to 'farm production and marketing management', 'diversification' and 'financial management' risk strategies. This may imply that farmers in the north-east region perceived these risk strategies as more important than the central region farmers. This is because most north-east farmers are poorer.

Independent variables	*Risk strategy components* [b]			
	AR1	*AR2*	*AR3*	*AR4*
Constant	3.310***	2.956***	2.523***	3.428***
Age [c]	0.054	0.124	0.003	-0.002
Gender [d]	-0.019	-0.107	-0.136*	-0.047
Highest education [e]	0.258***	0.167**	0.378***	0.110
Farming experiences [f]	-0.132**	-0.238***	-0.100	-0.121*
Off-farm work [g]	0.249***	0.227***	0.944***	0.150**
Farm size	0.001	0.015***	0.003	-0.004
Net farm income	-1.11E-06***	-1.98E-06***	-7.67E-07*	-7.32E-07**
Farm location [h]	-0.383***	-0.143*	0.092	-0.160**
Finance farm business [i]	-0.126**	-0.039	-0.202***	-0.026
Annual household income [j]	0.023	-0.275***	0.054	0.158**
Household size	0.033*	0.002	0.026	-0.001
R^2	0.146***	0.138***	0.267***	0.053***

[a] Variables and models significant at $^{*}P<0.1$, $^{**}P<0.05$ and $^{***}P<0.01$;
[b] Factors AR1-4 are labelled as AR1=farm production and marketing management, AR2=diversification, AR3=off-farm income and AR4=financial management;
[c] 1, if the farmer's age over 40 years old, 0 otherwise;
[d] 1, if farmer is male, 0 if female;
[e] 1, if the highest education of the farmer is high school and higher, 0 if primary school education or less;
[f] 1, if the farming experience over 30 years, 0 otherwise;
[g] 1, if the farmer has off-farm work, 0 if no off-farm work;
[h] 1, if the farmer's farm is located in central region, 0 if a farm located in north-east region;
[i] 1, if farm has a loan, 0 if farm without a loan; and
[j] 1, if household income greater than 90,001 baht and 0 represent otherwise.
Source: Field survey, 2009

Table 8. Multivariate regression of the risk strategy components and household and farm characteristics of all sampled Thai farmers (n=800) [a]

5. Implication of the results

Farmers in both regions perceived 'unexpected variability of input prices' as the most important sources of risk on the farm. In addition to the prices of chemical fertilizer, the increase in wage rates and higher land rental rates are the main factors that pushed the farm production costs upward. Over the past decade, the intervention of the Thai government in agricultural input policies had actually declined. The distribution of chemical fertilizers at reduced cost was the only scheme that the government organized to assist poor rural farmers. However, this scheme has recently been terminated due to limited government budget and this consequently reduced opportunities for the farmers to control production costs.

The results of the sources of risk perceptions, showed that 'unexpected variability of product prices' was the second most important source of risk among the central and north east region farmers. The Thai government operated a pledging scheme for the major cash crops such as wet rice, dry rice, cassava and maize.[34] This scheme aimed to help farmers when commodity market prices fluctuated early in the harvesting season. However, the pledging scheme has been widely debated among policy experts, especially for rice.[35-37] The advantage of the rice pledging scheme is that farmers can obtain low-interest loans from the government when they decided to pledge their rice to the Bank of Agriculture and Agricultural Cooperatives (BAAC) at the pledging prices and the rice will be transferred to storage at the Public Warehouse Organisation. The government allowed the farmers to redeem and sell their rice in the market when market prices increased above the pledging prices. The pledging price was set by a government announcement and generally the pledging period is approximately five to seven months each year.[38] Conversely, some economists argued that the pledging scheme would have long-term negative impacts on the efficiency of the country's rice market and it seems that the management of the scheme is shaped by political forces.[35] The pledging scheme persuaded farmers to increase their production, but the quality of the products was frequently ignored.[39] Some economists also suggested that the government should discontinue this highly-interventional price policy and should encourage farmers to sell their products using futures contracts to reduce the risk of price and income volatilities.[40, 41] This challenged policy makers to create mechanisms to stabilize agricultural prices at levels that are economically reasonable for both farmers and consumers. In addition, the effects of price policies such as the pledging scheme should be assessed cautiously to improve the effectiveness of the scheme. Direct access to futures trading markets may perhaps be too complicated for smallholder farmers in Thailand. Hence, government agencies such as Ministry of Agriculture and Cooperatives, Ministry of Commerce and The Agricultural Futures Exchange of Thailand should develop strategies that would increase small farmers' access to the futures market.

The development of a national agricultural crop insurance scheme should be one of the Thai government's priorities. Crop insurance is, theoretically, an efficient instrument in managing risks and can facilitate efforts to protect farmers from either the loss of their crops or farm income caused by natural disasters or drops in commodity prices. To date, a new crop insurance scheme for Thai farmers that has been operated by BAAC since 2008 is still in the pilot project stage.[42] The government expects this crop insurance scheme will continue to develop to cover all farmers and crops countrywide in the near future.[43] In addition, there are some obstacles that policy makers should consider for the successful implementation of the crop insurance schemes.[1, 42, 44]

First, the crop insurance scheme itself should not be too complicated because it could lead to high administrative costs for the scheme. Second, the appropriate insurance premiums and coverage accessibility under the scheme for each crop must be carefully considered. Low premiums may not always cover all the losses from the large-scale disasters, but the high insurance premiums will lead to increased farm production costs. Lastly, the government should promote the benefits of crop insurance schemes that could increase farmers' understanding and participation.

Strengthening the role of farmer groups or cooperatives should be considered as part of agricultural risk reduction policies in Thailand. This is because farmers' groups or cooperatives can help farmers to improve their negotiating power. Higher product prices and lower input prices can then be achieved more easily due to economies of scale.(1)

Author details

Satit Aditto

Department of Agricultural Economics, Faculty of Agriculture, Khon Kaen University, Thailand

Christopher Gan and Gilbert V. Nartea

Department of Accounting, Economics and Finance, Faculty of Commerce, Lincoln University, Christchurch, New Zealand

6. References

[1] Hardaker JB, Huirne RBM, Anderson JR, Lien G. Coping with risk in agriculture. Cambridge, MA: CABI Pub.; 2004.

[2] Hazell PBR, Norton RD. Mathematical programming for economic analysis in agriculture. New York: Macmillan; 1986.

[3] Dunn JW. Farm level risk analysis for Kansas farmers [doctoral thesis]. Manhattan, Kansas: Kansas State University; 2002.

[4] Hossain S, Mustapha NHN, Chen LT. A quadratic application in farm planning under uncertainty. International Journal of Social Economics. 2002;29(4):282-98.

[5] Nyikal RA, Kosura WO. Risk preference and optimal enterprise combinations in Kahuro division of Murang'a district, Kenya. Agricultural Economics. 2005;32(2):131-40.

[6] Pannell DJ, Malcolm B, Kingwell RS. Are we risking too much? Perspectives on risk in farm modeling. Agricultural Economics. 2000;23(1):69-78.

[7] Harwood J, Heifner R, Coble K, Perry J, Somwaru A. Managing risk in farming: concepts, research, and analysis. [report]. In press 1999.

[8] Shadbolt NM, Martin SK. Farm management in New Zealand. Auckland: Oxford University Press; 2005.

[9] Pellegrino JM. Risk management in agriculture: Argentine evidence of perceived sources of risk, risk management strategies and risk efficiency in rice farming [master thesis]. Lincoln, New Zealand: Lincoln University; 1999.

[10] Katikarn K. Risk and uncertainty of farmers in the central plain of Thailand [doctoral thesis]. Lexington, Kentucky: University of Kentucky; 1981.

[11] Patamakitsakul S. Thailand agriculture in the 10th National Economic and Social Development Plan (2007-2011) (in Thai)2006.

[12] Sayaphan C. Crop production planing under risk situations in Pitsanulok province crop year 1998/99 (in Thai) [master thesis]. Bangkok, Thailand: Kasetsart University; 2001.

[13] Kermel-Torrès D. Atlas of Thailand: spatial structures and development. Paris: IRD Editions; 2004.

[14] Bank of Thailand [BOT]. Thailand key economic indicators. 2012 [cited 2012 1 June]; Available from: http://www.bot.or.th/English/Statistics/Indicators/Pages/index.aspx.

[15] Ministry of Finance [MOF]. Value of exports. 2012 [cited 2012 1 June]; Available from: http://dwfoc.mof.go.th/.

[16] Office of Agricultural Economics [OAE]. Agricultural statistics of Thailand. 2012 [cited 2012 1 June]; Available from:
http://www.oae.go.th.

[17] National Statistical Office of Thailand [NSO]. The 2003 agriculture census. 2006 [cited 2006 26 October]; Available from:
http://web.nso.go.th/en/census/agricult/cen_agri03-1.htm.

[18] Lovelace GW, Subhadhira S, Simaraks S. Rapid rural appraisal in northeast Thailand: case studies. [working paper]. In press 1988.

[19] Flaten O, Lien G, Koesling M, Valle PS, Ebbesvik M. Comparing risk perceptions and risk management in organic and conventional dairy farming: empirical results from Norway. Livestock Production Science. 2005;95(1-2):11-25.

[20] Nicol RM, Ortmann GF, Ferrer SR. Perceptions of key business and financial risks by large-scale sugarcane farmers in KwaZulu-Natal in a dynamic socio-political environment. Agrekon. 2007;46(3):351-70.

[21] Boggess WG, Anaman KA, Hanson GD. Importance, causes, and management responses to farm risks: evidence from Florida and Alabama. Southern Journal of Agricultural Economics. 1985;17(2):105-16.

[22] Patrick GR, Wilson PN, Barry PJ, Boggess WG, Young DL. Risk perceptions and management responses: producer-generated hypotheses for risk modeling. Southern Journal of Agricultural Economics. 1985;17(2):231-8.

[23] Martin S. Risk management strategies in New Zealand agriculture and horticulture. Review of Marketing and Agricultural Economics. 1996;64(1):31-44.

[24] Meuwissen MPM, Huirne RBM, Hardaker JB. Risk and risk management: an empirical analysis of Dutch livestock farmers. Livestock Production Science. 2001;69(1):43-53.

[25] Hall DC, Knight TO, Coble KH, Baquet AE, Patrick GF. Analysis of beef producers' risk management perceptions and desire for further risk management education. Review of Agricultural Economics. 2003;25(2):430-48.

[26] Akasinha B, Ngamsomsuk K, Thongngam K, Sinchaikul S, Ngamsomsuk V, editors. Risk perceptions among rice farmers (in Thai). Multiple Cropping Center annual seminar; 2006; Chiangmai, Thailand.

[27] Sirisup S, Kammeier HD. Government policy and farmers' decision making: the agricultural diversification programme for the Chao Phraya river basin (1993-1995). [working paper]. In press 2003.

[28] Hair JF. Multivariate data analysis. 6th ed. New Jersey: Pearson Prentice Hall; 2006.
[29] Pallant J. SPSS survival manual: a step by step guide to data analysis using SPSS for Windows. 3rd ed. New York: McGraw Hill/Open University Press; 2007.
[30] Mustafa RH. Risk management in the rain-fed sector of Sudan: case study, Gedaref area Eastern Sudan [doctoral thesis]. Giessen, Germany: Justus Liebig University; 2006.
[31] Limsombunchai V. Rural financing in Thailand [doctoral thesis]. Lincoln, New Zealand: Lincoln University; 2006.
[32] IRINNews. Thailand: Flash floods continue to batter farmlands. 2008 [cited 2009 24 November]; Available from: http://www.irinnews.org/PrintReport.aspx?ReportId=80360.
[33] Ahmad A, Isvilanonda S. Rural poverty and agricultural diversification in Thailand. [Working Papers]. In press 2003.
[34] Department of Internal Trade. Pledging schemes. 2010 [cited 2010 10 October]; Available from: http://www.dit.go.th/contentmain.asp?typeid=8&catid=133.
[35] Forssell S. Rice policy in Thailand: policy making and recent developments. Minor Field Study Series. [working paper]. In press 2009.
[36] Hayami Y. An emerging agricultural problem in high-performing Asian economies. Policy Research Working Paper. [working paper]. In press 2007.
[37] Kajisa K, Akiyama T. The evolution of rice price policies over four decades: Thailand, Indonesia and the Philippines. [working paper]. In press 2004.
[38] Isvilanonda S, Bunyasiri I, editors. Food security in Thailand: status, rural poor vulnerability and some policy options. agricultural and food policy reforms: food security from the perspective of Asian small-scale farmers, 24-28 August 2009; 2009; Seoul, South Korea.
[39] Chantanusornsiri W. Crop insurance programe not expected to hit rice exports. Bangkok Post. 2009 September 18.
[40] Arunmas P. Ammar slams 'corrupt' rice scheme. Bangkok Post. 2008 November 21.
[41] Pratruangkrai P. Rice pledging scheme putting exports at risk. The Nation. 2009 February 23.
[42] Yimlamai A. Agricultural weather index insurance in Thailand. 2010 [cited 2010 19 December]; Available from: http://www.foodsecurity-asiapacific.com/wp-content/uploads/2010/10/pdf/Session5/3%20A%20Yimlamai-Agricultural%20Weather%20Index%20Insurance%20in%20Thailand.pdf.
[43] Commodity Online. Thailand to extend crop insurance plan. 2010 [cited 2010 20 December]; Available from: http://www.commodityonline.com/crops-weather/Thailand-to-extend-crop-insurance-plan-2010-03-09-26289-3-1.html.

[44] Abada J. Micro and area-based schemes: project and program issues. In: Ortiz ID, editor. Social protection in Asia and the Pacific: Asian Development Bank; 2001. p. 408-27.

Section 5

Climate Risk Management

Chapter 21

Improving Climate Risk Management at Local Level – Techniques, Case Studies, Good Practices and Guidelines for World Meteorological Organization Members

R. Martínez, D. Hemming, L. Malone, N. Bermudez,
G. Cockfield, A. Diongue, J. Hansen, A. Hildebrand, K. Ingram,
G. Jakeman, M. Kadi, G. R. McGregor, S. Mushtaq, P. Rao, R. Pulwarty,
O. Ndiaye, G. Srinivasan, Eh. Seck, N. White and R. Zougmoré

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51554

1. Introduction

Climate can be viewed in a number of ways. As a constraint or setting, climate provides the broad boundary conditions within which a range of ecosystems services function. Climate may be considered a determinant in that it may be part of a causal chain of direct or indirect events leading to a particular impact or outcome. Climate can also be viewed as a resource; this notion implies that climate has a value, could be managed and manipulated and by extension could be allocated. More often than not, climate is viewed as a hazard. Whichever of these views of climate is adopted, it is clear that climate has a close relationship with nature and society and therefore climate variability and change may pose a range of risks for environments, societies and economies.

As our understanding of the climate system and our ability to predict it into the future have improved, and as society has become more aware of the possible costs and benefits of managing (including adapting to) climate risks (see Box 1. 'Climate risk definition'), individuals, communities and organisations are seeking suitable information, tools and techniques to enable appropriate management decisions to be made. These need to be accessible, dependable, usable, credible, authoritative, responsive, flexible and sustainable.

The generic process of applying such information to climate risk decision making, including identification, assessment and prioritization of the risks followed by a coordinated and sustainable application of resources to reduce, monitor and control the probability and or impact of detrimental effects, is known broadly as Climate Risk Management (CRM).

Box 1. Climate risk definition

There are many definitions of 'risk' depending on the application and context. Based on the Intergovernmental Panel on Climate Change definition of disaster risk (IPCC, 2012), climate risk can be defined qualitatively as the likelihood of unfavourable impacts occurring as a result of severe climate events interacting with vulnerable environmental, social, economic, political or cultural conditions.

It can also be defined more quantitatively, as the product of the probability of a given climate event occurring and the adverse consequences of this. As such, climate risk originates from a dynamic combination of climate hazards (e.g. extent and duration of extreme temperatures or rainfall) and the vulnerabilities (propensity or predisposition to be adversely affected) of exposed elements (e.g. communities, economic or societal sectors or ecosystems).

There are a number of challenges for CRM. Climate risks pose what has been termed a wicked problem[1]. Wicked problems do not have set solutions; instead, greater understanding of the wicked problem and partial solutions to the problem evolve iteratively within the social contexts of the scientists/analysts, intermediary agents, and end users. As well, there is no commonly accepted methodology for assessing and prioritising climate risks, identifying key thresholds in these risks, or for considering what are important criteria for managing these risks.

Current challenges for climate risk decision makers include identification of and assessing timely, reliable and appropriate climate risk information and then using that information to make well informed decisions. These are not simple processes given the complexity of the social and institutional mechanisms, the multiple potential sources of climate information (not all of which is consistent or authorative or easy to understand), and the multi-faceted nature of information, that are often involved in such decisions. Appropriate approaches and principles should be adopted to foster collaborations among climate risk information users and providers, and enable the implementation of effective management actions.

The World Meteorological Organization (WMO) and partnering agencies have recognized the needs of users of all kinds for relevant, actionable climate information for CRM and are taking steps to address those needs. Thousands of scientists and decision makers from climate and other disciplines met at World Climate Conference-3 (WCC-3, 31 August to 4

September, 2009, Geneva, Switzerland) to discuss the issues, and concluded that a Global Framework for Climate Services (GFCS) was needed to organize and provide users with the climate information, products and services appropriate to their requirements. Following WCC-3, a major report (WMO-No 1065) on the GFCS outlined, *inter alia*, current capabilities around the world, the needs and opportunities for climate services and recommendations. The GFCS is designed to mainstream climate science into decision making at all levels and help ensure that every country and every climate-sensitive sector of society is well equipped to access and apply relevant climate information, enabling an adjustment of planning and decisions to optimize the given situation. The application of climate services must therefore involve close interaction between all stakeholders including the providers and the users, and requires concerted multi-disciplinary efforts[2]. The ultimate goal of GFCS is to: "Enable better management of the risks of climate variability and change at all levels, through development and incorporation of science-based climate information and prediction into planning, policy and practice."

In 2011, the WMO agreed to implement the GFCS, and is working with partnering agencies to develop or strengthen the many contributions to the GFCS, in terms of observations, monitoring, research, services and capacity development, all with a user focus and engagement. As part of this effort, experts from around the world are collaborating to scope requirements for CRM, develop methods for interacting effectively with user communities, and improve applications of climate information for user benefit.

Motivated by the need to ensure the effectiveness of the GFCS, and the associated requirement to improve the practical application of CRM at local levels, in order to reduce climate impacts, build resilience to climate variability and change and contribute to poverty reduction and development, the WMO Task Team on Climate Risk Management (TT-CRM) organized a CRM Symposium in Guayaquil, Ecuador in October 2011. Experts in a wide range of climate and risk disciplines from all the continents and key agencies working with WMO on the GFCS implementation attended. The overarching aim of the Symposium was to help both providers and users of climate information in the development and application of information on climate variability and change, in an operational 'no regrets' sense, for minimizing climate-related risks and maximizing any opportunities that may. As a result of this international meeting, innovative approaches for CRM were discussed, practical examples of best practice were highlighted, and guidance for appropriate processes, tools and techniques to adopt were proposed.

In this chapter, key points and outcomes from the WMO Symposium on CRM are highlighted, including proposing a definition of CRM (see Box 2. 'Climate Risk Management (CRM)'), and recommendations on CRM for WMO Members. The aim is to provide a useful document to all who are interested in establishing or improving CRM processes and systems at the local level (particularly village to country scales). Throughout the chapter, conceptual discussions are complemented with real-life case studies and lessons learnt and shared by experts who are involved in CRM across different sectors and in academia.

Box 2. Climate Risk Management (CRM)

The World Meteorological Organization Task Team on Climate Risk Management (TT-CRM), based on the expert guidance from the Symposium on CRM, proposes a definition of CRM as **a systematic and coordinated process in which climate information is used to reduce the risks associated with climate variability and change, and to take advantage of opportunities, in order to improve the resilience of social, economic and environmental systems**.

2. The process of CRM

Many CRM frameworks, that include different key steps and elements (actors, tools, techniques etc.), have been developed to provide guidance and a degree of consistency for applying risk management to a range of climate-related issues[3,4,5,6]. These frameworks tend to be conceptual and general so that they can be easily applied to a wide range of concerns by both technical and non-technical users, and may vary in their foci and level of detail ranging from international to local, or general to sector specific.

Despite the variety of frameworks and tools for assessing and managing climate risks, utilisation of the output to enable decision making is advancing rather slowly[7]. Continued guidance by risk management experts to begin the CRM process with user and provider engagement and collaboration has often not been considered, and assessments are still typically conducted by scientists and then the information is handed to users in a one-way exchange - the so called top-down approach - that often does not promote effective decision making. It is important to continue to promote collaboration between assessors and stakeholders (providers and users of climate information) at all stages of the CRM process, to enable all parties to understand the steps involved in the knowledge to action pathway, and therefore to facilitate effective and sustainable CRM responses.

The WMO Symposium on CRM discussed various CRM frameworks and steps, which has enabled the TT-CRM to identify the following key steps that are considered essential to consider in a CRM process. These should not be regarded in a linear manner, but rather combined in an iterative or cyclical order:

- User and provider engagement and collaboration
- Climate risk assessment
- Communication and dissemination of climate risk knowledge, information and tools
- Adaptation and capacity development
- Monitoring, evaluation and improvement

Some of the key points for each of these steps are highlighted below, and practical examples, to demonstrate good practice on how they have been applied to real-life CRM projects, are provided.

2.1. User and provider engagement & collaboration

Breuer and colleagues[8] compared the traditional research model, in which researchers develop new technologies or tools that extension agents (professionals trained in skills such as communication and group facilitation, and usually also in technical areas of the sector they serve) deliver to end users, to a pipeline that delivers water. They proposed an alternate loop model, in which the loop encircles end-users, extension agents, and researchers with diverse opportunities for interactions among this co-learning community. The symposium supported this approach, and further recommended that science-based learning communities or communities of practice use as many methods as possible for engagement and collaboration.

In this section the engagement of agricultural stakeholders, particularly through experiences with the Southeast Climate Consortium of the United States of America (USA) (SECC), is emphasized. However, the same principles could be applied successfully for CRM decision makers in other sectors, for example water resource managers, coastal community planners, or wildlife managers.

The work of extension agents has been particularly successfully applied in the agricultural sector. Opportunities to engage end-users, e.g. farmers, in the process of developing a decision support system have included surveys, interviews, sondeos, workshops, focus groups, working groups, presentations and displays at association meetings, and on-line feedback[8,9,10]. An example of good practice in this sector has been the SECC, who developed AgroClimate (see http://agroclimate.org/), an on-line decision support system for extension agents and end users, using all of the engagement opportunities described above[8,11,12] .

Box 3. Steps for engaging farmers and outreach workers

1. Ask what they want.
2. Listen.
3. Give them what you think they asked for.
4. Ask them whether you've given them what they need.
5. Listen.
6. Observe whether they use the information or tools that you have given them.
7. Modify what you have provided.
8. Go back to step 4.

The most important points for engaging farmers and outreach workers as research collaborators identified through the experiences of SECC are straightforward: 1) knowing which questions to ask; and 2) listening to the answers. The various methods that were used for engagement with end users follow the basic iterative steps outlined in Box 3. 'Steps for engaging farmers and outreach workers', but each is unique in terms of the depth and breadth of information that they can evoke.

By providing a range of engagement methods, individuals can self select how they will engage in the community, depending on their level of interest, availability of time, and willingness to commit to an activity. This ability of community members to select the engagement activities in which they will participate applies to nearly all members of the community, including end-users, extension agents, and researchers. The exception to the self selection clause is a core team of three or more individuals who are fully committed to the community. The SECC strives to have a team that includes at least one social scientist (anthropologist or rural sociologist), climate scientist, and agricultural scientist. The ability of this committed team to work together will be the most critical factor in the success of the engagement.

Two of the engagement methods noted above – sondeos and working groups (both powerful and less commonly used in other reported participatory research approaches) – are further described:

2.1.1. Sondeos

A sondeo (Spanish for sounding) is a semi-structured discussion in which a two- to three-person multi-disciplinary team engages one or two people from the target audience in conversation[8]. Most of the sondeos conducted for the SECC have been part of a graduate course in field research methods. The students and their instructors meet to discuss the problem of interest and to agree on a general set of questions to ask. These questions are a guide to conversation, rather than a formal questionnaire. The course members divide into small multi-disciplinary teams to conduct their conversations, typically at the residence or place of work for the people of target audience. For the agricultural community, extension agents have been vital in helping identify people from diverse target audiences, where the targets have included small and large farms, vegetable growers at farmers markets, farmers with and without irrigation, and others.

An important benefit to a conversational approach is that it often elicits key issues that the researcher could not have anticipated, issues that would likely have been missed with an interview or survey that has a list of pre-established questions. In keeping with its conversational nature, the researchers do not take notes during the conversation. Rather, when the conversation is completed, each researcher writes their own synthesis of the discussion. In one day, a single team can usually complete three or four discussions, which typically last about one hour, followed by note writing.

At the end of each day, all teams assemble to discuss their findings and to identify new questions that will guide conversations on the following day. Some questions may be retained throughout the sondeo in order to provide continuity, but the discussions evolve day-by-day as teams engage in conversations, learn, share their learning, and modify their conversation guide questions. After one week of field work, the course identifies a leader to write the sondeo report. Examples of SECC sondeo reports can be downloaded from http://SEClimate.org/pubs.php.

2.1.2. Climate working groups

A working group includes members from science and the broader community who meet regularly, typically 3 or 4 times per year, to engage in dialogue on the new findings from science, information and technology needs of the broader community[9]. The steps for building and nurturing a climate working group are outlined in Table 1. The SECC has successfully established working groups for agricultural and water supply utility communities. Both have about 25 to 30 members and both are highly productive, yet each has distinct features that reflect the differences among the communities.

A key element to the success of a climate working group is commitment, both on the part of the individual members of the groups and the institutions that they represent, if any. For example, most members of the water supply utility climate working group represent a particular water provider, city, or agency. If a member of any particular institution is not able to attend a meeting, the participating institutions identify an alternate who can attend. This policy provides both continuity and assures that the institutions have also committed to the working group.

Phase 1: Exploratory	• Identify key stakeholder groups and individuals • Assess climate-related concerns and priorities • Map network of stakeholder groups • Discuss interest in development of a climate working group • Identify potential participants and assess their interest in working together
Phase 2: Group building and goal setting	• Identify convener and facilitator for an exploratory meeting • Convene stakeholders to exchange information and explore scope for a climate working group • Discuss expectations and establish short- and long-term goals • Define group norms and roles
Phase 3: Implementation	• Develop a work plan and timeline • Identify gaps in expertise and experience • Develop sub-groups as needed to address specific topics
Phase 4: Monitoring and feedback	• Define outcome and process criteria to track progress • Include time for reflection and feedback for iterative group adaptation

Table 1. Phases and activities for the building and nurturing a climate working group. Source: Bartels et al. (2011)[9].

By far, these climate working groups demand the greatest level of commitment from the learning community for any of the engagement methods that have been tested, but it is precisely this commitment that helps them advance science. The climate working groups help researchers build collaborative relationships with different stakeholder groups for on-going learning, both by the scientists and the stakeholders. They link research with real-world decision needs to help improve resource management strategies of stakeholders as well as improving the research and education programs of the science community. Most importantly, climate working groups engage members from diverse stakeholder groups that might not otherwise interact and promote the legitimacy of the science community as a source of information and technology that is relevant to solving the wicked climate problems that society faces.

2.2. Climate risk assessment

As with the definition of climate risk, or CRM, there are many definitions of risk assessment. A common theme across most is the requirement for a process and/or technique that provides information with which to assess the key risk or risks. For example, the Society for Risk Analysis proposes that: "Risk assessment is the process of establishing information regarding acceptable levels of a risk and/or levels of risk for an individual, group, society, or the environment" (see: http://www.sra.org/resources_glossary_p-r.php).

Risk assessments involve analysis techniques, methodologies and tools that have the key quality of assessing uncertainty (a common quality of risk), either quantitatively or qualitatively, and representing this as some measure of likelihood and/or probability. Climate risk assessment is used to help decision makers optimize resources for responding to climate-related disasters and reducing risks and impacts associated with current and future-projected climate variablity and change. It is one of the first stages of CRM, and involves identification and synthesis of hazard and vulnerability information/data that is relevant to the specific climate-related risks identified through the 'User and provider engagement and collaboration' step. One very important consideration in all climate risk assessments is the balance between the quantification of climate hazards (intensity, frequency and/or duration) and the approach to estimate the main elements of vulnerability on the ground i.e. level of exposure, poverty, exclusion, education, organizational capacity, infrastructure among others. Both hazard and vulnerability estimations may be validated using historical information of climate events and changes in socio-economic vulnerabilities and associated impacts. However, such assessment may encounter problems, for example, in some cases quality of data may be poor, data may not be available, skill of forecasts at different scales can be low. As well, even if the climate information is complete and correct, the user may not access it, or may not understand or know how to apply it.

Indicators of climate-related risks (impacts, hazards and vulnerabilities) are often used to focus a risk assessment on the specific areas of interest for the decision maker. Indicators are values that can be monitored (and/or modelled) to assess changes in the state of a system, and are important tools for simplifying complex processes, with potentially multiple drivers and feedbacks, into useful and accessible information. Defining which indicators are appropriate for decision makers, as well as climate monitoring or projection purposes can be a complex process, and many different approaches have been adopted[13,14,15,16]. One of the more common approaches used for indicator-based studies uses the driving force-pressure-state-impact-response (DPSIR), pressure-state-response (PSR) or driving force-state-response (DSR) which organize indicators in the context of a causal chain[17,18,19].

Climate risk assessments typically include statistical analyses of historical climate indicator records and assessment of information on climate-sensitive impacts, together with understanding of the climate mechanisms and the cascade of processes leading to these impacts. Geographical information and mapping may also be used to assess the zones where impacts are recurrent and are associated with human losses and/or infrastructure damages. Temporal changes of impacts and their related climate hazard characteristics are also often a key part of climate risk assessment, and may be directly linked with social, economical or environmental variables which may change exposure and resilience. In an ideal world, there would be millions of meteorological stations contributing to development of an accurate idea of the historical evolution of climate variables. The reality is, however, that there are not enough meteorological stations, the available stations are not evenly distributed in time and space, available data are not always digitized or shared, and there can be problems in some cases in the quality, completeness or homogeneity of the available data. Notwithstanding such issues, individual climate records measured at specific places integrate the history of the complex interactions between land, air, sea, ecosystems, and community in those locales. The final result is expressed in those climate records and consequently the history of this whole interaction process is reflected in time series of the measured values, or of their departures from a chosen reference period.

Analysis derived from climate indices/indexes, such as that provided by the CCl/CLIVAR/JCOMM Expert Team (ET) on Climate Change Detection and Indices (ETCCDI) (see http://www.clivar.org/organization/etccdi/etccdi.php) is a powerful tool that can be applied at local level. In places where high quality climate records (preferably long period, minimal gaps, and homogeneous) are available, the information that can be delivered through such climate analysis is absolutely useful, in conjunction with social and economic and other information for that locale, in DRR, adaptation and CRM processes. Such information is more accurate and in most cases more appropriate than that generated by downscaled models, but in zones where there are no available stations, information generated by downscaling is the next best alternative.

Climate risk assessment may also include a future element, utilising climate change projections and/or 'what-if' scenarios to explore the potential impacts of future scenarios of change. Practical obstacles to using information about future conditions are diverse, ranging from limitations in modeling climate system complexities (e.g. projections having coarse spatial and temporal resolution, limited predictability of some relevant variables, at scales that matter for decision making and forecast skill characterization), to procedural, institutional, and cognitive barriers in receiving or understanding climatic information, and the capacity and willingness of decision-makers to modify actions[20,21]. In addition, functional, structural, and social factors inhibit joint problem identification and collaborative knowledge production between providers and users. These include divergent objectives, needs, scope, and priorities; different institutional settings and standards, as well as differing cultural values, understanding, and mistrust[22,23].

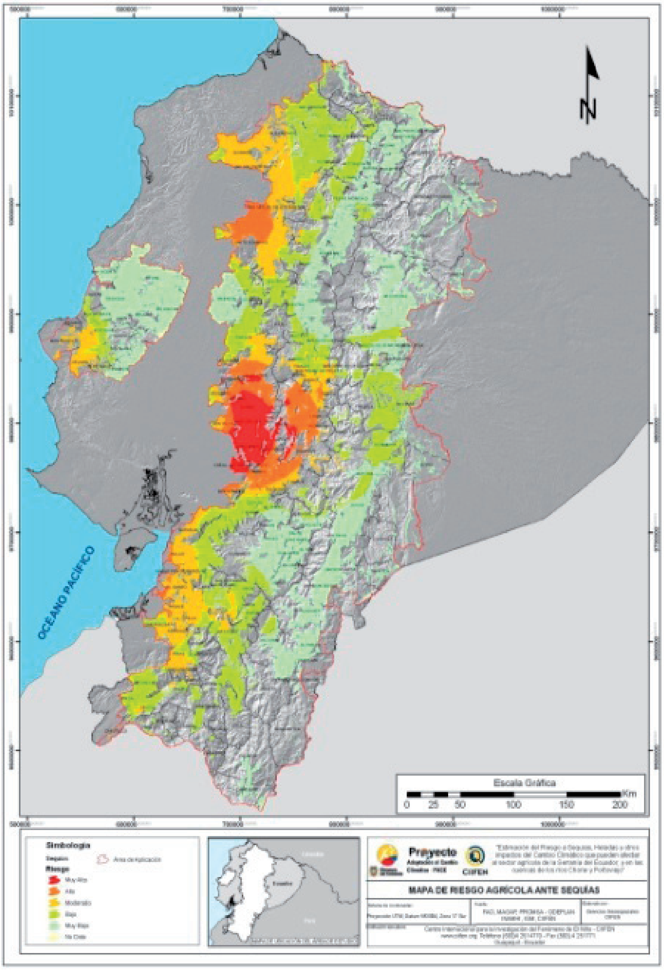


Figure 1. Climate-related risk map for agriculture in the highlands of Ecuador. Color scale: Red: high risk; Light green: low risk; Grey: no data. Source: MAE (2012) www.ambiente.gob.ec

A fundamental part of risk assessment is related to vulnerability. However, it is very difficult to provide a unique formulation or set of indicators for vulnerability, as these will vary across sectors, geographically and in reponse to socio-economic conditions. A typical view of vulnerability considers the combination of several elements: the level of exposure (of an element which must be specified, e.g population, livelihood, infrastructure, etc.), the level of susceptibility which is a degree of how much the natural hazard can affect it, less or divided by the coping capacity of the exposed element which includes all the factors of resilience in the community, livelihood, infrastructure, etc.). While increases in the level of exposure and susceptibility both increase vulnerability, increases in coping capacity reduce it. This approach mixes physical exposure (i.e. the presence (location) of people, livelihoods, environmental services and resources, infrastructure, or economic, social, or cultural assets in places that could be adversely affected by physical events and which, thereby, are subject to potential future harm, loss, or damage) with the social determinants of vulnerability. According to the most recent IPCC-SREX Report (2012)[22], the IPCC describes vulnerability as the propensity or predisposition to be adversely affected. Every location on the planet has its own vulnerability profile and a specific evolution pattern. Historically this vulnerability pattern can be approximated with some social or economical indicators, statistics of disaster or land use multi-temporal comparison. The central focus of climate risk assessment is to understand the relevant climate hazards and their evolution over time, together with the vulnerabilities and how these have evolved, and a likely to change in the future, in a particular area. It is not possible to implement CRM or adaptation actions only with climate scenarios. This information must be complemented with the estimation of current vulnerability and potential future evolution.

No single, consistent approach for conducting risk assessments has emerged, instead a range of different techniques have been used[24]. The choice of a particular technique is influenced by several factors, including: the goal of the assessment, the exposure units to be studied (an exposure unit is defined as the sector, location or activity being assessed), availability of data, choice of models suitable for the projection of future outcomes, and the time frame involved. A major challenge for future climate change assessments is the uncertainty associated with future projections and the propagation of this uncertainty throughout an impact assessment[3]. One approach has been to give a range of uncertainty bounded by low and high scenarios of climate change. However the outcomes of such analyses may be too broad for planning effective adaptation.

An example of a successful climate risk assessment for the agricultural areas in the highlands region of Ecuador was developed in 2011 by the International Research Center on El Niño (CIIFEN), Ecuador. This was requested by the Ministry of Environment as part of the National Plan for Adaptation. For the assessment, agricultural areas were identified based on up-to-date satellite information, and specific field verification. Information and indicators for agriculture aptitude, erosion, hydrological deficit, level of access to water for irrigation, type of soil, were considered, and social and economical indicators were selected. All information was analyzed spatially at parish level and combined to produce a vulnerability map covering the Ecuadorian highland region. This was further combined with historical climate hazard maps of "dry consecutive days" and "high temperature indexes", as reported in the Second National Communication of Ecuador to the UNFCCC, 2009[25]. The resultant map of the climate risks for the agricultural sector in the highlands of Ecuador (Fig. 1) is currently used by national and local authorities to assign priorities, allocate resources and address the key elements involved in the vulnerability of the agriculture sector to cope with the potential climate hazards based in the historical trends.

2.3. Communication and dissemination of climate risk knowledge, information and tools

The term 'risk communication' as used here, refers to intentional efforts on the part of one or more sources (e.g. international agencies, local government, communities) to provide information about hazards and hazard adjustments through a variety of channels among themselves or to different audiences (e.g. the general public, specific at-risk communities), for the purpose of influencing the recipients to apply the information and take appropriate action. It also includes efforts of local communities to characterize and communicate their risk-based experiences. Lindell and Perry (2004)[26] summarized the available research as indicating message effects include pre-decisional processes (reception, attention, and comprehension). Several studies have identified the characteristics of pre-decisional practices that lead to effective communication over the long-term[27,28,29].

Communicating and disseminating risk information can be very challenging. One of the first steps for effective communication is to ensure two-way communication channels, where information providers and users can interact equally and explain misunderstandings. Before starting a CRM process, it is paramount to build and apply "climate information chains"[30], as discussed above. This involves a complex network of institutions involved in the end-to-end process of CRM, i.e. National Meteorological and Hydrological Services (NMHSs), disaster management agencies, national and local authorities, the media, private sector, community representatives, and public and private agencies of strategic sectors such as agriculture, health, water resources. Such a complex network requires diverse means of communication, which has included web-based GIS tools with, *inter alia*, real time information updates, e-mail distribution lists, text alerts and high-frequency radio transmission which is useful for remote locations. A climate information chain should have legitimacy, credibility and be interactive. It is a kind of "living mechanism" that must be kept operational. To get people and institutions engaged in this chain, dialogues, meetings and agreements are also necessary. One example of an operational mechanism for communicating and disseminating risk information would be regional or national climate outlook forums (RCOFs and NCOFs). In these forums, climate information providers and users meet (either face-to-face or virtually), usually on the release of a seasonal climate forecast. The opportunity to share information, discuss issues and build knowledge has proven invaluable in many parts of the world (for further information and references, see http://www.wmo.int/pages/prog/wcp/wcasp/wcasp_home_en.html).

Once climate information chains are set up, the information to be disseminated should consider that climate knowledge should optimally combine scientific knowledge and indigenous knowledge. Both are necessary to ensure the effectiveness of their application. Science is not enough to contribute practical and effective solutions for CRM, but when it is linked with the local culture and experience of the communities, fantastic responses can be obtained. This is exemplified by a young member of a remote community in Ecuador drawing a risk map for his location based on all the experiences, impacts and weaknesses their community has evidenced, but with a better understanding of how the vulnerability was built and the main climate hazards that threaten this community[31]. There is indeed considerable evidence to show that if communities at risk are actively involved in information collection and analyses then they are far more likely to rely on that information than if it is just provided to them from 'outside'[30]. Information is also regarded as credible to local actors if it is collected and reported by individuals recognized by the central bureaucracy and locals as responsible observers with minimal political motive, such as teachers or extension workers[32]. One effective way to consolidate climate information chains is through their usefulness. If such informaiton chains become operational and communities respond effectively during planning, early prepardness and response, users become engaged and empowered by the system because they feel they are part of it.

Examples of risk information generation and diffusion efforts within disaster research and response communities include interpersonal contact with particular researchers, planning and conceptual foresight (as in Red Cross/Red Crescent brochures), outside consultation on the planning process (as per the Federal Emergency Management Agency of the USA (FEMA)), and user-oriented transformation of information and individual and organizational leadership. The characteristics of risk communication messages involve information quality (specificity, consistency, and source certainty) and information reinforcement (number of warnings) that have significant impacts on adoption of adjustments[33,34]. Messaging should also aim to foster 'no-regrets' actions, in which the recipient of the information takes climate-related decisions or action to maximize positive and minimize negative outcomes of climate variability and change.

2.4. Adaptation and capacity development

The IPCC definition of adaptation to climate change is the "adjustment in natural or human systems in response to actual or expected climatic stimuli or their effects, which moderates harm or exploits beneficial opportunities"[35]. From a conceptual point of view adaptation and capacity development are part of the overall CRM approach. In practical implementations of adaptation projects these two compartments often become indistinguishable, for example, Monsoon Forum meetings, which focus on preparedness planning for monsoon season climate variability across Asia (see http://www.rimes.int /societal/monsoon-forum), often also involve training sessions to build capacity across the community. However, there is a strong tendency to try to separate and manage adaptation and capacity development independently of each other, as well as from the other steps involved in CRM. This can cause a negative effect on implementation and makes the institutional information framework which is necessary to get effective responses in human systems more complex.

In accordance with the Global Adaptation Partnership (www.climateadaptation.cc), compelling climate change evidence has prompted the development and implementation of national adaptation plans around the globe, and these have generated specific within-country actions related with vulnerability and risk assessments including long-term climate observations or projections. In some places, these actions overlap with CRM (as well as other risk management) plans of national disaster management or other government agencies. For example, the UK government Department for Environment, Food and Rural Affairs' (DEFRA) Climate Change Risk Assessment was one of the requirements from the UK government's Climate Change Act (for further details see http://www.defra.gov.uk /environment/ climate/government/risk-assessment/).

One of the key issues for adaptation is its local nature. CRM or adaptation processes require ownership by communities, not only their authorities, for successful implementation. Such ownership is strongly connected with the perception that individuals have about the risks, the current impacts over them and how they are able to affect their integrity and progress. To encourage this perception, climate risks must be clearly identified and presented to communities, and this information must be validated with local data, information and local community feedback with experiences of how the population has coped with recent climate impacts. Setting up practices and mechanisms for coping with climate variability in the here-and-now is an important step in preparedness for future climate changes. NMHSs are important partners for providing such local climate and meteorological information to inform adaptation. Given a local perception of climate risk, it is easier to generate adaptation strategies and solutions with the community and negotiated with the authorities, and implement action plans designed by the community. Particularities in local plans are key points from the cultural and social point of view. By adopting a local focus to CRM, adaptation is necessarily a bottom up social construction implemented by local communities.

Past experience in capacity development for CRM suggests that providing a single recipe to conduct capacity development efforts has limitations. Some principles for a more comprehensive and effective approach, based on good practice, are therefore proposed as guidance:

- Consider sustainability of capacity development through continual rotation of technical staff in national and local agencies;
- Design training strategies under a "train the trainers approach";
- Prepare educational material combining scientific and indigeneous knowledge;
- Ensure a robust institutional and stakeholders network to support the capacity development process (with multi-level stakeholder coordination and communication);
- Consider blended training courses, with both face to face activities supported by e-learning systems;
- Implement accountability mechanisms for both capacity development implementers and beneficiaries;
- Ensure an effective and long-term monitoring and evaluation mechanism involving national and local institutions;
- Keep accessible databases of people trained and the impact of the capacity development in their current activities;
- Develop a long-term strategy for education (alongside shorter courses), such as the 5-year Climate Change Adaptation Degree and Master of Science (integrated) programme including agriculture, animal sciences, fisheries, forestry, biodiversity, water resources, health which has been launched by Kerala Agriculrural University, India;

- Co-share capacity development efforts with beneficiaries;
- Encourage allocation of financial support from national, local or private stakeholders to support CRM capacity development;
- Develop the capacity of information providers to deliver improved information, services and products, through, for example, improving the observing networks and climate risk modeling capability.

2.5. Monitoring, evaluation and improvement

Monitoring, evaluation and improvement are important elements in CRM. Monitoring could cover organizational, financial, operational, political, regulatory and other issues. It is required to collect relevant information and data to help quantify the risks, and to assess the success (or not) of adaptation or other interventions. Evaluation is the means by which the accomplishments are compared with the expected goals and what improvements are required to complete the iterative cycle proposed for CRM, within which there is an implicit assumption that progress should be continually sought.

Before implementing any monitoring, evaluation and improvements for CRM, some key indicators should be set up. It is unlikely that one single indicator can be identified to monitor all the risks (hazards and vulnerabilies) necessary to monitor the CRM process. However, to be practical, a few key indicators should be identified with the following attributes - quantitative (e.g: number of victims and cost of impacts), quality controlled based on standards (e.g: ISDR guidelines for design of EWS), and time limited.

The monitoring, evaluation and improvement step enables key oversight bodies (such as a NHMS, a regional health organization or other stakeholder) to track progress against initial conditions, and assess areas where corrective action is required. It can also be an important tool for the pursuit of resources to address gaps and deficiencies in the process, and to facilitate the dialogue between the relevant stakeholders in the risk management process at various levels, which is an essential part of a feedback loop between relevant stakeholders. Through this step the focus is kept on results – on reaching the desired goals – as well as on continual learning. It must be flexible, allowing for and even anticipating new challenges or opportunities, or new methods and understanding in the theory and application of CRM.

3. Role of early warning systems and adaptation planning within the context of disaster risk reduction and CRM

"Neither society nor the environment are static. Consequently, neither is the risk" (Alan Lavell)

It is clear that CRM may encompass a wide range of temporal and spatial scales, depending on the nature of the risks and their socio-economic context. Both climate and disaster risk are considered as some integration of hazards, vulnerabilities and exposures (see Section 1), and EWSs may be integral to the adaptation measures used to manage these risks. An important concept for Disaster Risk Reduction (DRR) taken from ISDR Glossary, 2009[36] concerns "the risk of disaster", which is usually expressed as ‘the probability of life loss or property destruction or damage in a given period of time'. The action of DRR usually refers to the socio-economic objective of reducing that risk. In comparison, CRM is focused more towards the longer-term application of climate information and tools in a multidisciplinary scientific context to address both the positive and negative impacts of climate variability and change on society, infrastructure and life. The concepts of DRR and CRM are therefore complementary, both including a focus on risk management, and there are mutual advantages involved in designing integrated DRR and CRM projects.

If the integration requirements of DRR and CRM are considered according to impact scales (local, regional or global), it is clear that even if some climate phenomena relate to global causes they are materialized through locally-specific contexts, causing damage or losses depending on existing capacities in those local areas. Hence, the need to focus on improving the management capabilities of both DRR and CRM is especially important at local scales. Early Warning Systems (EWS) are often central to DRR and CRM, particularly over relatively short time scales (minutes to weeks), whereas other systems and tools that focus on longer-term adaptation planning tend to be more appropriate for CRM at longer time scales (months to decades). All such measures have the common aim of reducing vulnerability, increasing resilience and improving response capacities of people, economies and ecosystems at risk.

The UN International Strategy for Disaster Reduction (UNISDR) defines early warning as "the provision of timely and effective information, through identified institutions, that allows individuals exposed to a hazard to take action to avoid or reduce their risk and prepare for effective response". Governments often maintain EWS to warn their citizens and themselves about impending hazards, resulting for example, from health, geologic, or, climate and weather-related drivers. Traditional assumptions are that effective functioning of EWS requires only prior knowledge of risks faced by communities and other users of the early warning information. Under a CRM framework EWSs are expanded to other adaptation planning measures, including technical monitoring and warning services for highlighting the risks and their potential impacts, effective strategies for dissemination of understandable warnings to those at risk, and finally, knowledge and preparedness to act[37]. Two additional elements have been introduced, 1) awareness that risks are changing (and which new risks may arise) and, 2) the need for constructing and communicating new knowledge about future conditions that can be understood, trusted and used[38,22]. One goal is to be prepared to use windows of opportunity for engaging and providing leadership, and for legitimizing risk management and successful communities of practice that have arisen during but also between events.

Given the links between near- and long-term climate variability and change, the early warning construct also applies to more extended timescales. For example, WMO 'Climate Watch' systems utilise near real-time and historic climate observations with proactive mechanisms for interacting between users and NMHSs to provide alerts on major climate anomalies and extremes (see http://www.wmo.int/pages/themes/climate/climate_watch.php). Improving the institutional organization of the EWS as well as the associated strategic response to crises are closely linked to developments in understanding of climate vulnerability and governance[39,40,22]. Countries or regions that have developed such systems may also use them to develop and inform strategic adaptation response options to climate changes, thus developing broader institutional flexibility and preparedness, and reducing societal vulnerability.

For most locations early warning is still a linear process based on a "sender-receiver" model of risk communication. In this section, the term "early warning information system" is used to describe the more integrated process of risk assessment, communication and decision support, of which an "early warning" is a central output. An early warning information system involves much more than development and dissemination of a forecast, it is the systematic collection and analysis of relevant information about and coming from areas of impending risk that: (a) Informs the development of strategic responses to anticipate crises and crisis evolution; (b) Provides capabilities for generating problem-specific risk assessments and scenarios, and (c) Effectively communicates options to critical actors for the purposes of decision-making, preparedness and mitigation. Central to the implementation of this more comprehensive vision of "early warning information systems" is a detailed examination of the root causes of the lack of early action[41].

Numerous international and national EWS' exist [42,43,44]. In addition, many early warnings directly and indirectly activate other warning systems in affected sectors and communities, a process that has been referred to as a cascade of early warnings[45]. For the most part, EWS' have been interpreted narrowly as technological instruments for detecting and forecasting impending hazard events and for issuing alerts. This interpretation, however, does not clarify whether information about impending events is actually communicated and used to reduce risks[44,22,46].

An example of good practice with an EWS is provided by the Climate Forecast Application in Bangladesh (CFAB) project. Heavy rainfall episodes in the Ganges-Bhramaputra basin (combined drainage area ~1,662,000 km^2) cause human suffering almost every year. Webster & Hoyos (2004)[46] showed the possibility of using physically based statistical schemes to predict rainfall with lead times of more than 10-days in the monsoon region. Based on this and subsequent research the CFAB project, supported by Office of Foreign Disaster Assistance of the United States Agency for International Development (USAID/OFDA), was launched during the monsoon of 2003 and 2004. Long-lead forecasts for rainfall in the river basin were given using the UK-based European Centre for Medium Range Weather Forecasts (ECMWF), Tropical Rainfall Measuring Mission (TRMM) and other datasets[47]. The Program was a collaboration (see Fig. 2) of the following agencies: Atmospheric and Oceanic Sciences at the University of Colorado, Boulder, Georgia Institute of Technology, ECMWF, Bangladesh Meteorological Department (BMD), Bangladesh Flood Forecast and Warning Centre (FFWC), and Asian Disaster Preparedness Center (ADPC) in Thailand. Coordinated efforts by ADPC, BMD and the Institute for Water Modeling (IWM) resulted in the development of 1-10 day discharge forecasts at major stations of two rivers (Hardinge Bridge, on the Ganges and Bahadurabad, on the Bhramaputra). The FFWC was responsible to produce local-level forecasts in other locations along these rivers. The Center for Environmental and Geographic Information Services (CEGIS)[48] disseminated flood forecasts to communities during the monsoon season, working in close coordination with the Disaster Management Bureau (DMB) and the Department of Agricultural Extension (DAE). With the additional lead times and tailored warnings, community level flood risks were better managed. Communities were able to mobilize in advance (e.g. move livestock to higher ground, secure their fishing nets) in order to protect their livelihood assets. The project ran a second phase from 2006 to 2009 with support from USAID Bangladesh through CARE-Bangladesh. The objective was to transfer technology from the USA to Bangladesh and to build the capacity of national and local institutions for a sustainable end-to-end generation and application of flood forecast products in high-risk locations.

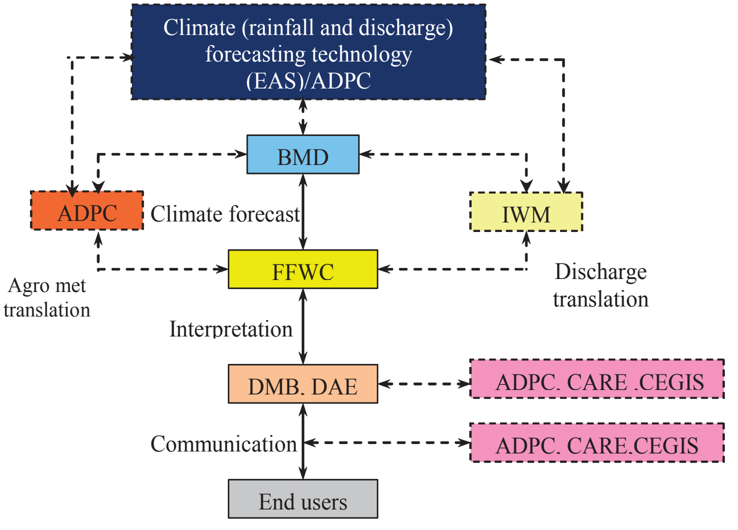


Figure 2. Institutional linkages for 1 to 10-day forecast of rainfall in Climate Forecast Application in Bangladesh (CFAB) project. Solid lines denote forecast/product flow and broken lines indicate coordination between the institutions: Asian Disaster Preparedness Centre (ADPC); University of Colorado; Georgia Tech – Earth Atmospheric Sciences (EAS); Bangladesh Meteorological Department (BMD); Flood Forecasting and Warning Centre (FFWC) of Bangladesh Water Development Board (BWDB), Disaster Management Bureau (DMB); Department of Agricultural Extension (DAE); Center for Environmental and Geographic Information Services (CEGIS); and CARE Bangladesh.

The disaster research and emergency management communities have shown that warnings of impending hazards need to be complemented by information on the risks actually posed by the hazards and likely strategies and pathways to mitigate the damage in the particular context in which they arise. Effective "early warning" thus implies information is introduced into an environment in which much about risk and vulnerability is assumed[49]. Vulnerability analysis provides a contextual basis for early warning by identifying structural, water, energy, and food insecurity attributable to disruption of primary means of access including informal community safety nets[50]. As is long-recognized by the disaster, food and water security communities, and more recently the climate adaptation research communities, successful early warning information systems integrate "input" and "output" indicators. Input indicators include measures of production potential, including rainfall, soil conditions, heat and crop and livestock growth. Output indicators include nutritional indices, behavioral indicators, and signals of economic activity, that deal with the food, water and other supply situations or changes in demand that result from scarcity[51]. The timing and form of climatic information (including forecasts and projections), and access to trusted guidance to help interpret and implement the information and projections in decision-making processes may be more important to individual users than improved reliability and forecast skill.

Experience provided by the U.S. National Integrated Drought Information System (NIDIS) and the United States Agency for International Development (USAID) Famine Early Warning System (FEWSnet) drought early warning information systems developed in the USA has led to the following recommendations for developing EWS:

- Develop a Governance structure.
- Frame the goals and objectives of international and country and intervention strategy from a securities perspective (water, food etc), e.g FEWSnet, NIDIS.
- Strengthen the scientific and monitoring foundations to support early warning.
- Specify of reliable information provided by forecasts, especially for key climate features i.e. ENSO.
- Improve understanding of the modulation and combined impacts of interannual and decadal-scale variations on agricultural and meteorological drought duration and severity.
- Place multiple indicators within a statistically consistent triggering framework-cross-correlation among units for rapid transitions (e.g. climate and vegetation mapping) before critical thresholds are met from onset to severity.
- Develop risk and vulnerability profiles of drought-prone regions and locales including impact of climate change adaptation interventions on food and water availability, access, and use.
- Develop indicators and methodologies to assess the risk to environmental services, value and costs of environmental degradation, and impacts of water and crop subsidies.
- Inventory and map local resource capabilities (infrastructure, personnel, and government/donor/ngo-supported services) available to complement food and water program operations.
- Conduct gaming scenarios with planners and decision makers for selected past and projected events to:
    - Improve understanding on whether and how best to use probabilistic information with scenarios of potential surprise and cumulative risks at each scale.
    - Map decision-making processes and identify policies and practices that impeded or enable the flow of information among information system components.

The NIDIS and FEWSnet experience also provides a good example to demonstrate that successful drought early warning information systems have explicit foci on: (1) integrating social vulnerability indicators with physical variables across timescales, (2) embracing risk communication as an interactive social process and, (3) supporting governance of a collaborative framework for early warning across spatial scales[46]. Forecasts need not be perfect to make early warning useful. For longer-term EWS, it is also important to note that although a trend in the drought-based indicators may serve as a warning, the actual point of transition or threshold (e.g. dune mobilization) to increased severity remains difficult to predict.

Traditional warnings, with justification, remains an important source of climate information in many rural communities. At the community level, farmers in Zimbabwe and Malawi have identified local language radio programs as credible and accessible mechanisms to deliver forecasts if they occur with follow up meetings with extension agents or other intermediaries[52]. Internet based tools, such as Google maps, and graphical tools are already being used for participatory, large-scale information development. However, these tools are inherently limited in communicating the relevant local context and the consequences (positive and negative of information use). For most locations, the governance context in which EWSs are embedded is also key. The links between the community-based approach and the national and global EWSs are weak at present[53]. Improving the complementarity and legitimacy of both approaches is a new challenge to address especially in developing the institutional foundations for global climate early warning information systems envisioned by the Global Framework on Climate Services (see section 1. Introduction).

There is a critical need to approach and support early warning through DRR and Climate Change Adaptation (CCA)[54,22], and the overarching processes involved in CRM. This requires a framework that uses climate change scenarios not above but within risk and vulnerability profiles, thereby capturing the nature of capabilities and decision-making networks. These form the basis for effective EWS design and implementation. The cases above, and other efforts, have demonstrated that social protection and early warning information interventions can provide DRR while helping to meet the goals of adaptation to changes in extreme events. Furthermore, sustainable development prospects are very dependent on the effectiveness of the many networks of EWS'[57]. In these networks, subtle rules of interaction emerge that shape the context in which resource-related decisions are taken, and the rules are negotiated and made[55,56].

To ensure that DRR and CRM are integrated utilising appropriate systems, information and tools, some transversal capacities need to be established between the scientific community studying and analyzing the climate information (at timescales relevant to both DRR and CRM), and the decision-makers who are required to consider the full spectrum of the impacts of climate variability and change. Decision makers across all facets of society also need to be aware of the changes, risks and impacts threatening their societies and find appropriate ways to adapt to and protect these from the most damaging changes. They should also consider climate as a resource, with beneficial aspects that can be exploited, through application of timely and appropriate climate information, tools and products.

4. Case studies demonstrating effective CRM practices

CRM is designed to be a practical process to be implemented on the ground. People, policies, environmental issues, governance, information, cultural aspects among other elements should be organised to communicate in an appropriate manner to deal with extreme weather- and climate-related risks. There is no single CRM solution for a particular situation in any part of the world. However, over time, there are increasing success stories demonstrating good practices for a wide range of CRM situations. The following case studies are examples of such good practice in CRM for different locations and development sectors.

4.1. CASE STUDY 1: Using probabilistic seasonal forecasting to improve farmers' decision in Kaffrine, Senegal (Ousmane Ndiaye, Robert Zougmoré, Jim Hansen, Aida Diongue, El Hadji Seck)

Although agriculture and pastoralism occupy 80 per cent of the population in the Sahel, climate information is not yet widely integrated into farm management decision systems. However, many efforts have been made in the region to produce climate information such as the yearly climate outlook forum preceeding the incoming rainy season[57]. Yet, this hasn't benefitted the user community, particularly the most vulnerable to climate variability and change. This paper documents one ongoing demonstration project in Kaffrine, Senegal, within the peanut growing basin, where rural communities, policy-makers and relevant institutions are testing the use of probabilistic seasonal forecasts for managing climate risk. The process, from training the farming community to evaluating the use of the forecast information, is outlined.

4.1.1. Background

Rainfall in the Sahelian region of West Africa experiences strong variability over time-scales ranging from intra-seasonal (including long dry spells and false onset) to inter-annual and decadal. At the longest time scale, climate change is shifting the desert boundary and altering the landscape. This strong variability has an impact on many sectors, including health, agriculture and water management. The major impacts of climate variability in this region make CRM an imperative for the livelihoods of Sahelian communities. Each time scale of variability requires a specific climate risk plan.

As is the case in most French colonized countries in Africa, ANACIM, the national weather service of Senegal, is in charge of providing meteorological services to the country. ANACIM, in partnership with the CGIAR research program on Climate Change, Agriculture and Food Security (CCAFS), initiated a pilot project in 2011 to test the usefulness of probabilistic seasonal forecast information to peanut farmers in Kaffrine. In addition to ANACIM, the key stakeholders participating in the project include local government technical services, local farmers and NGOs. A big challenge in the training was to go through many key and important steps in achieving good CRM, including producing skillful seasonal forecasts at district level, presentation of probabilistic seasonal forecasts using easy-to-understand terminology, training farmers with the new information for them and translate this into decisions. This study concludes with lesson learned during the initial year of the project.

4.1.2. Building trust

The approach used in the project was to help the various stakeholders identify strategies for successful CRM using seasonal forecasting. The project sought to build the trust of farmers, while working with all relevant local organizations. In West African's culture, trust is the most important requirement in order to be effective in a viable partnership. A common saying states, "the manner to give is better than what is given". In order to build trust with farmers and foster a sustained relationship, a workshop that included 30 farmers from 6 villages around Kaffrine was planned. Since ANACIM does not have the legal mandate to implement any agricultural strategy, partnership with the local agriculture department representative (SDDR) was ensured. The SDDR has the mandate to monitor activities related to the farming system, and also has arbitration authority in case of conflicts over issues such as farm allocation, fertilizer subsidizing, buying harvest products. SDDR was a natural contact point with farmers, since they had already developed a long time partnership and has their trust. It was very important that the project team not appear as a stranger in the system, but work through a known entity. Association was developed with other local technical services including agricultural advisers from the national agency for agricultural and rural advice (ANCAR), which has a presence nationwide at district level and a mandate to advise farmers in term of agricultural strategies. Volunteers from World Vision, a Christian charity that assists children and invests a lot in agriculture in the district of Kaffrine, also participated. Participants included individual farmers, and members of farmer organizations such as JAPPANDO. Women represented about 30 per cent of those participating.

4.1.3. Training

Consistent with the strong oral tradition, time was reserved during the workshop to allow farmers to interact with the experts team. On the first day of the workshop, the floor was given to farmers to describe their experiences with forecasting the weather and climate. The technical experts started with differentiating the concepts of weather (imminent) and climate (longer term), as these concepts are interchangeable in the local language.

4.1.4. Connecting with farmers' indigenous knowledge

A challenge for the CRM approach was to enable farmers to trust and use scientific information. There was a clear need for a common ground, where farmers would use the new scientific seasonal forecasting approach proposed to them without feeling that their indigenous approaches to seasonal forecasting were being rejected. In this culture, the scientists would lose credibility if the farmers were to think the scientists were saying that their elders were wrong and only the scientists were right. The strategy was to listen to them and understand the aspects of their traditional knowledge that might be climate related. The farmers were welcomed as guardians of knowledge passed from generation to generation, and invited as experts to share their indigenous climate knowledge. According to their tradition, being elder means possessing wisdom. The farmers were asked to specify whether each indicator was for immediate weather, or for climate conditions for the upcoming season. Some of the indicators were clearly just coincidental events with no apparent link to the climate, but many were very much related to climate and specifically with the high humidity and high temperature associated with the monsoon system (Table 2). After carefully listening, the scientists acknowledged the farmer's memories and explanations.

Climate variability	Indicators
Onset of the rainy season	When the wind changes direction to fetch the rainfall Apparition of stars shaped as elephant Birds crying as if it calls men to go to field and woman to stay at home Early flowering of many trees species: Néré, dimb, tamarinier, sone Butterflies and libellees are numerous Some persons feel heavy in their body Hot night time
Major rainy event	When wind is shifting direction When dark clouds become white
Good rainy season	When snakes and frogs are more numerous than usual The shooting star direction indicates which zone will receive excess rain this year Net appearance of seven stars in the sky
Good cropping season	When the rain is settled in June the 24th and we start the millet around the 14th of July we can expect good harvest
End of the season	When frogs start chanting When the sky is high When we observed dew in the morning

Table 2. Quotes from farmers on their perceived climate variability indicators

4.1.5. Explaining the basis of seasonal forecasting

Another challenge was how to explain to farmers, in easy-to-understand terms, how climate forecasts work. Many farmers knew about weather forecasts, communicated through the weather bulletin on national TV. But the real challenges were to convince them that the rain could be forcast one to two months ahead, and to help them to understand the probabilistic nature of forecasts at this lead-time. The basis of seasonal forecasting was explained to the farmers by calling upon their intuition. When asked, "When it is hot, why do people go to the beach?" they responded that the sea breeze brings fresher air. They were then asked, "Then why? Isn't it the same sun that heats both land and ocean? Why then does the ocean get cooler in summer?" It was explained that ocean has better memory of the past compared to the continent. Ocean remembers the heat of the past days and weeks. That's why, on a very hot day people go to the beach to benefit from ocean memory of the past weeks. Similarly, when it is cold, the ocean still remembers recent warm days. The ocean's heat memory is the basis for seasonal forecasting. As rain comes from clouds, clouds come from water vapor, and most water vapor from the ocean, they could see how ocean temperature could control rain. The farmers were also informed that satellites are used to monitor ocean temperature throughout the world, and computers quantify the likelihood of rain in Senegal. This very simplistic explanation helped them to make sense of scientific seasonal forecasts, and was sufficient to convince them to trust the forecast during this first contact.

4.1.6. Getting past the technical language barrier

The next challenge was to explain the probabilistic nature of the forecast, which is less intuitive than a deterministic rainfall amount. To start with, the farmers were asked to recollect from their memory the last 5 rainy seasons and rank them from the wettest to the driest. With a pluviograph (Fig. 3), it was explained how rainfall is recorded in millimeters, and what 1mm of rain means. One mm of rain was poured into the soil, then the farmers were asked to compare it with the quantity of rain that they consider sufficient to plant their crops. They indicated that they plant when the soil wetting front is greater than the span of an average man's hand. To help them interpret what a seasonal total means, a discussion was held on how the temporal distribution of rain relates to the seasonal total. It was clear for the farmers that a seasonal forecast gives an idea of the total, but no information about its distribution in time. One farmer explained the difference between a good rainy season and a good cropping season, which was a clear indication that they understood the seasonal forecast output. The farmers were informed also what could potentially be forecast and what could not.

Figure 3. Training farmers to read a rain gauge.

4.1.7. Probability of exceedance graphs

It was decided to express the forecast as a probability of exceedence of rain instead of the probability of occurrence of the three tercile categories that meteorological services in West Africa officially issue in their seasonal forecasts. Farmers understand the notions of uncertainty and probability, but understanding and acting on formal probability formats is challenging. To help them understand the new probability of exceedance format, an exercise of classifying the last 5 years of rainfall that they recollected from memory was conducted (Fig. 4). A chart of 30 years of Kaffrine rainfall data was provided. The farmers could see that it is very likely that they would get at least as much rainfall as the driest year, and very unlikely that they would get more than the wettest year. The middle years represent "normal" conditions. How to identify the 25th, 50th and 75th percentiles of rainfall from the graph was discussed. The idea that a dry forecast would shift the distribution toward the left and wet forecast to the right was introduced. Hypothetical wet and dry forecasts were discussed until the farmers appeared to understand what they meant. As the probability of exceedence is a cornerstone of the training, the farmers were divided into four groups. Two groups were given probability of exceedence forecasts for hypothetical dry years and the other two groups were given hypothetical wet years. The farmers were asked to discuss what they would do differently if this were the actual forecast for the upcoming season. Each group reported back on their forecast and strategies. The whole group was encouraged to comment on these strategies.

Figure 4. Farmers sorting seasonal rainfall.

4.1.8. Communicating the forecast

After the training, the next step was to build a communication strategy to ensure that the information will reach the farmers effectively. A discussion of the best way to communicate the seasonal forecast revealed a number of options. Among the means identified, cell phones appeared first as a cheap technology. This option is accessible, since most of the farmers have a cellular phone, and is consistent with the traditional use of oral communication. Local radio was the other promising means of communicating forecast information. All of the participating farmers listen to the radio, but the listening quality of the radio is very poor when they are on the farm. Some NGOs and farmers association leaders recommended e-mail as a possibility. The administrative authority who was present mentioned the government's network of heads of village. In case of an extreme event, this can be used to reach each village within an hour. The local authority showed his support and promised to help with access to this facility. To avoid conflicts between farmers' organizations, other farmers recommended sending the information through the SDDR, who knows how to contact them.

4.1.9. From theory to practice

A week after the training work, ANACIM sent a group of experts to call a meeting to communicate the actual July-September 2011 seasonal forecast with the farmers. Twenty-two attended. Some key points from the training were revisited: good rainy season versus good cropping season, probability of exceedence interpretation, plausible management response strategies, definition of 1mm. Rain gauges were distributed to some representatives of farmers' organizations who expressed need for this tool, and the meaning of a millimeter of rain was again demonstrated. The forecast was presented with an explanation on how to interpret it (Fig. 5). The forecast in this case was "normal to above-normal." As the year before, 2010, had been exceptionally wet – the highest on record – it was indicated that rainfall this year (2011) would probably be less than 2010. Some explanations about what the seasonal forecast did not say were also offered. Recommendations on any particular management strategies were not made, but rather it was left open to each farmer to decide. Considering that this was a first contact with them, it was preferred to build trust first before offering recommendations.

Figure 5. Training on interpreting the probability of exceedance.

4.1.10. Keeping in touch

Through this project, funds were available to undertake two field visits during the season, and also to call selected farmers from time to time. When the first big rainy event occurred, some farmers were asked if they planned to use the seasonal forecast, whether they spread the word, how the rain was, etc. It was not a forecast but rather a monitoring exercise. It was good to touch base. During the first field trip in selected villages (October 12-13 2011), some farmers made major decisions, such as borrowing money from the bank to invest more in their farm, or hiring workers. Another field visit was made around the end of the cropping season (October 18-22, 2011) to conduct surveys on expected yield.

4.1.11. Evaluation of the seasonal forecast

In January, three months after the rainy season, when farmers have sold their crops and finished their farming work, an evaluation workshop was organized in Kaffrine to assess the use and usefulness of the seasonal forecast strategy. Local extension services were present, as well as farmers' organizations. Fifteen of the farmers who attended the training workshop in June were invited back, along with 13 other farmers who hadn't received information about seasonal forecasting. During the January workshop, participants assessed both 2011 seasonal rainfall and the performance of various crops grown in the district. The participants took the opportunity to discuss in three groups, and interpret the information presented. One group included 12 farmers that had received the forecast and adjusted some decisions in response to the forecast (group I). The next groups included 3 participants who did receive the forecast but didn't make any adjustment to their farming practices (group II), and the last group consisted of 13 farmers who had never received any climate forecast information (group III). They were asked to document actions taken, problems encountered, and recommendations. Group I understood from the workshop that a short cycle crop was suitable because the season was to be less than 2010, but rainfall would be enough. The main problems they listed were: the high spatial variability of the rainfall, the late occurrence of the first rainfall which made it difficult to judge when to start planting, a long dry spell, and early termination of the season. They wanted to know or get: the starting date, finer forecast information in space, a weather bulletin each two weeks, and more training to better understand the forecast. Group II did not use the seasonal forecast because they had already bought their seeds at that time which made it difficult to change any of their farming strategy. Group III, who had never received any climate information, indicated that they had thought 2011 would be like 2010. They missed the opportunity of a long season in 2010, and were prepared to catch up the next year by choosing a long cycle variety, buying fertilizers and hiring wage laborers. The group members concluded that their problem was that they didn't know anything about the course of the rainy season and needed to be part of the group that received seasonal forecast training.

4.1.12. Lesson learned and way forward

The Workshop participants were given the chance to evaluate the whole process – from farmer selection, to organization of the workshop, to training agenda – in order to identify what is needed for improvement. There is a need to improve the communicating system by using already existing channels. World Vision recommended that training more trainers would be the best way. Overall, the farmers appreciated the experience of last year and welcomed more training.

Seasonal climate forecasts could have considerable potential to improve agricultural management and livelihoods for smallholder farmers. But constraints related to legitimacy, salience, access, understanding, capacity to respond and data scarcity have so far limited the widespread use and benefit from seasonal predictions in the Sahel region. The existing constraints reflect inadequate information services, policies or institutional processes in the region. However there is great potential to overcome these constraints. An approach is suggested that packages: i) seasonal and onset forecasts, ii) opportunity for farmers to implement strategies, and iii) insurance tools in case of extreme variable or dry years. Even when the seasonal rainfall or onset matches the forecast, poor farmers wouldn't profit if they don't have access to funds or crop varieties to implement any forecast-based strategy. And it turns out that in Kaffrine, there is often false start of the rainy season, making it imperative to provide farmers with alternatives, for example through index insurance.

As work with farmers in Kaffrine on the forecast continues, research is being conducted and a working group on improving prediction of intra-seasonal variability has been set up. Crop producers and seed bank will be invited into the process, to allow farmers to access suitable varieties for forecast-based strategies. There is some work on index insurance in the region, and it is planned to reach out to involve such groups in this effort. Through this approach it is hoped to gain success, avoid frustration and build long-term partnerships.

4.2. CASE STUDY 2: Climate risk management of plantation crops in the humid tropic region of Kerala, India (G.S.L.H.V. Prasada Rao)

The global economy has adversely been affected to a considerable extent due to weather related disasters which are not uncommon in the recent past. It is true in the case of the Indian economy too. The year 2010 was the warmest year ever recorded, followed by 2009 in India.

Climate model simulations indicate that a marked increase in rainfall and temperature over India could be seen during the current century. The maximum expected increase in rainfall is likely to be 10-30 percent over central India, with temperatures projected to increase between 2 and 3^{o}C by the end of the 21st century.

More frequent storm surges and increased occurrence of cyclones in the post monsoon period, along with increased maximum wind speed are also expected along the coastal belt. As a result, the occurrence of floods and droughts, cold and heat waves and sea level rise may adversely affect the food security to a large extent across India, as seen in 2009, 2002 and 1987. Such impacts also influence plantation crops, which are predominantly grown in the humid tropics like Kerala, in the southwest of India (between latitudes

8°15`N and 12°50`N and longitudes 74°50`E and 77°30`E). The location map of Kerala is given in Fig. 6.

Figure 6. Location map of Kerala. Source: www.mapsofindia.com

4.2.1. Rainfall and thermal regimes of Kerala

The annual rainfall across Kerala is highly variable, averaging about 3000mm, but varying between less than 1000mm to greater than 5500mm.

Seasonally, rainfall is bimodal, due to the influence of both the summer and winter monsoons, with maximum monthly rainfall (>600mm) during the summer monsoon in June and July, and winter monsoon rainfall (200-300 mm) during October. Heavy rainfall during the summer monsoon, followed by a prolonged dry spell is a characteristic feature of the humid tropics, which is particularly prominent in the case of the northern districts, including Kasaragod, where the influence of winter monsoon is negligible (Fig. 7).

Annual average surface air temperature varies between 25 and 30°C, with a seasonal range between around 18°C in winter and 35°C in winter. The altitude across Kerala varies from below mean sea level to above 1500m, and temperature varies significantly with altitude. Accordingly, a sequence of crops is grown across the altitudinal range (see Table 3).

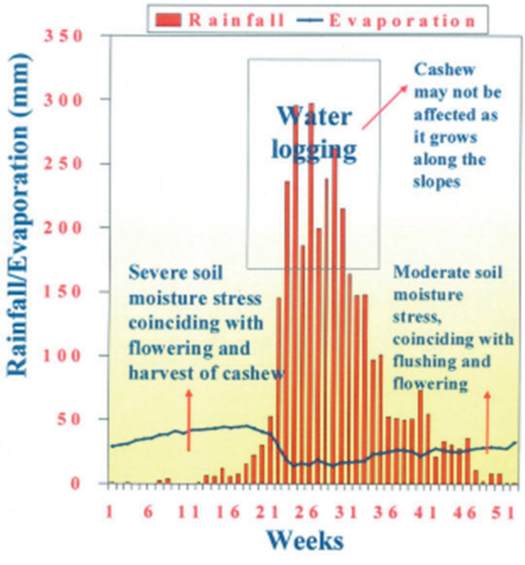


Figure 7. Mean weekly rainfall and pan evaporation at RARS, Pilicode, Kasaragod District, Northern Kerala, India.

Class	Region	Temperature conditions	Altitude (amsl)	Crops
Mega-therms	Low land	High to Moderate temperature throughout the year	0 -10 m	Coconut, arecanut and cashew
Meso-therms	Mid land	Moderate temperature throughout the year, winter temperature relatively low	10 -100 m	Coconut,cocoa arecanut, rubber, cashew and black pepper
Micro-therms I	High land	Moderate to Low temperature throughout the year, winter temperature low	100 -500 m	Rubber, coconut, cashew, arecanut and black pepper
Micro-therms II	High land	Low temperature throughout the year	500-1000 m	Coffee (arabica), rubber, arecanut and black pepper
Micro-therms III	High ranges	Low temperature throughout the year, winter temperature is occasionally goes below 0°C	1000-2500 m	Tea, Coffee (arabica) and Cardamom

Table 3. Altitudinal sequence of crops in Kerala.

Since rainfall is abundant during the monsoon season, surplus water during the first crop season leads to waterlogging which is detrimental to crop growth. In comparison, the second and third crops often suffer from soil moisture stress, and crop failure is a common phenomenon if irrigation is not assured. Erratic rainfall distribution during the monsoon, coupled with failure of the northeast monsoon, may result in drying up of surface reservoirs during summer, which are the major water resources in the region. In recent years, meteorological droughts during the monsoon and summer droughts are not uncommon across Kerala, with the summers of 1983 and 2004 particularly prominent (Fig. 8).

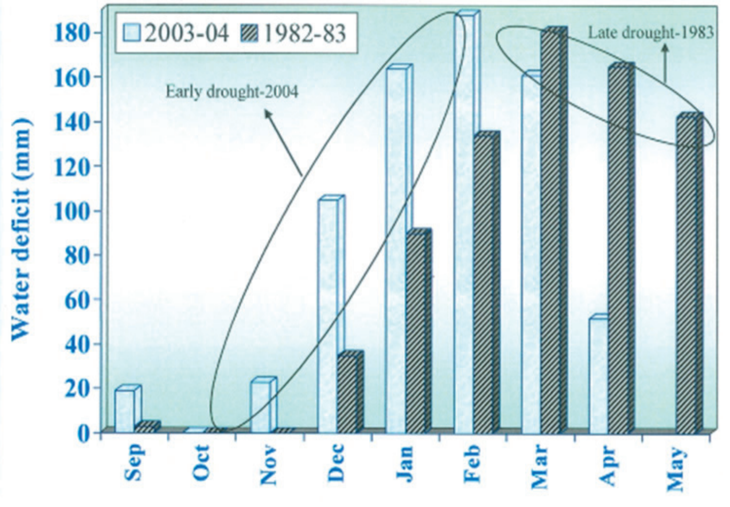


Figure 8. Water deficit in Kerala from Sept to May 1982-83 and 2003-04.

4.2.2. Impact of summer drought on plantations

The prolonged summer droughts, coupled with high temperature and low atmospheric humidity, in Kerala during 1983 and 2004 adversely affected production of many plantation crops, particularly rainfed coconut palms, arecanut, cardamom, coffee and black pepper and as a result the economy of the state was impacted. For example, monthly nut yield declined by up to 50 per cent (depending on management practices) in the year following drought, cardamom yield reduced by 30 per cent in Idukki dristrict and several black pepper gardens were wiped out in Wayanad district. Cocoa yield was also adversely affected due to high temperature in the absence of soil moisture. The impacts of these droughts on the agriculture and economy of Kerala highlighted the need to manage the risks posed by climate variability and change in this region, including other climate hazards such as floods, cold temperatures and heat waves). Various measures are now in place and being developed to pro-actively manage these risks, particularly at local levels. In the next section two of these measures are highlighted: 1) Scarce water resource management specifically through effective management of irrigation; and 2) Weather forewarning and dissemination.

4.2.3. *Climate risk management*

Scarce water resource management – irrigation: Management of irrigation during the summer months under scarce water resources is one of the key tools available for managing the adverse impacts of summer drought on crops. Various methods can be used to assess the irrigation requirements for different crops and time periods throughout a season, e.g. estimate weekly water deficit/surplus by taking the difference between weekly rainfall and open pan evaporation, or through calculations of potential evapotranspiration.

In the RARS, Pilicode location of Kerala, the irrigation requirement for coconut was estimated using the FAO's CROPWAT decision support tool for estimating crop irrigation water requirements based on soil, climate and crop data (see: http://www.fao.org/nr/water/infores_databases_cropwat.html). According to CROPWAT, the monthly average irrigation requirement for coconut in this region varied from 1106 litres/palm/month in December to 1488 litres/palm/month in April. The total irrigation requirement from December to May was estimated as 7807 litres/palm (Table 4). These values have provided guidance to coconut growers in the region on the general amount of irrigation water required to improve yield during average summer months.

Month	ETo (mm)	Water requirement (mm) (ETo x 0.75)	Irrigation requirement (1) ($\pi r^2 h$)
December	3.79	2.84	1106
January	3.95	2.96	1154
February	4.56	3.42	1204
March	5.01	3.76	1464
April	5.26	3.95	1488
May	4.76	3.57	1391
Total			7807

ETo-Reference evapotranspiration; 0.75-Crop coefficient; r-Radius of coconut basin in m^2

Table 4. Estimated irrigation requirements for coconut in the RARS, Pilicode location of Kerala.

More detailed seasonal irrigation advice to coconut growers has been provided by field experiments in which coconut palms were either irrigated at a rate 450 litres/palm/week for differing periods between December and May, or irrigated according to a climatic water balance approach, or not irrigated (Table 5). Results showed that the yield improved in all the irrigated treatments when compared to that of pre-treatments yield or no irrigation, as a result of reduction in the duration of water stress. Irrigation applied as per the climatic water balance approach (T6) showed one of the largest percentage yield increases, indicating that the preferred irrigation treatment for coconut yield is as required during the whole summer.

Treatment	Pre-treatment yield (1976-89) (A)	Post-treatment yield (1991-97) (B)	Difference in yield (A – B)	% increase over pre-treatment yield	Whether significant over pre-treatment yield
T1	103	122	19	18.4	S
T2	103	112	9	8.7	S
T3	94	107	13	13.8	S
T4	87	120	33	37.9	S
T5	92	115	23	25.0	S
T7	91	95	04	4.4	S
T6	60	82	22	36.7	S

T1- Irrigating the palms @450l/palm/week during December and January
T2- Irrigating the palms @ 450l/palm/week from December to February
T3 – Irrigating the palms @450l/palm/week from December to March
T4 – Irrigating the palms @ 450l/palm/week from December to April
T5 – Irrigating the palms @ 450l/palm/week from December to May
T6 – Irrigating the palms as per climatic water balance procedure
(150 l/palm/week in December, 200 l/palm/week in January, 300 l/palm/week in February,
350 l/palm/week in March, 400 l/palm/week in April, 450 l/palm/week in May)

Table 5. Duration of soil moisture stress on coconut yield of WCT.

Agro climatic zonation: Based on climate variables such as precipitation and potential evapotranspiration, the agro climatic zonation can be delineated using the water balance techniques. Such agro-climatic zonation if delineated on crop wise, climatic risks can be mitigated to a considerable extent. In the case of cardamom across the Western Ghats, Zone I & II are superior when compared to that of Zone III, where climate risk is high in terms of high temperature, prolonged dry spells and less length of crop growing season (Fig. 9). Similarly, cashew can be extended from northern districts of Kerala to south of Maharastra along the West Coast and North of Tamil Nadu to Orissa along the East Coast and inland areas away from the Coast. However, tea mosquito bug incidence along the West Coast and cyclones along the East Coast are the constraints for obtaining better cashew production. Similarly, simulation models can very well be used to simulate production potential of various crops in a given watershed area through which the climate risk can be minimized with proper crop management practices. In addition, agroadvisory service based on weather forecasting will go a long way in sustenance of crop production. Another multidisciplinary project launched by the Government of India, that is FASAL, is a classical example to help in GIS based watershed planning in Agriculture as a part of climate risk management.

Weather forewarning and dissemination: A reliable and clearly disseminated weather forecast is a very important tool for forewarning crop managers of potential weather hazards. The India Meteorological Department is constantly working to improve forecast skill and help disseminate the forecast in a suitable form to aid farm level decisions.

To improve dissemination of weather forecasts in a timely manner to agricultural villages, Village Resource Centres that are linked online to an Agro Advisory Service (AAS) have been established (under the ISRO programme) across the Kerala region. AAS' base their advice on the latest weather forecasts and agricultural expertise. The economic impact of a weekly AAS based on medium range weather forecasting has been assessed for different crops and regions (Table 6). This showed that the percentage increase in yield varied from 6.4 – 19 per cent depending upon the crop in the case of AAS farmers compared to the non-AAS farmers. Furthermore, it indicated that seasonal crops need intensive advisory, followed by less intensive for biennials and perennials.

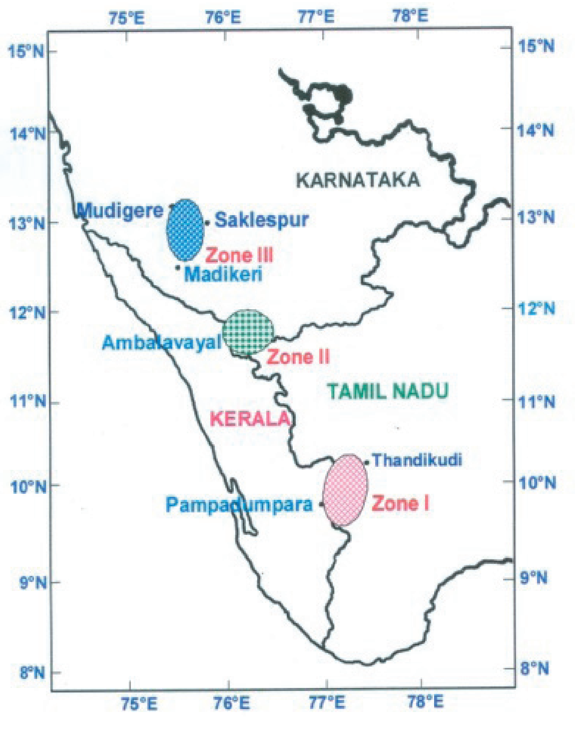


Figure 9. Agroclimatic zones of cardamom across the Western Ghats.

Season	Crop	Yield (t/ha)		% increase over in yield over non-AAS farmers	% increase in net return over non-AAS farmers
Rabi 2003-04	Paddy	3.7	3.0	19.0	30.0
	Banana	31.8	27.9	12.2	11.5
	Coconut*	13460	11025	18.1	13.2
Kharif 2004	Paddy	2.8	2.6	7.1	31.6
	Banana	22.8	20.5	10.1	7.9
Rabi 2004-05	Paddy	3.0	2.7	7.5	34.1
	Banana	31.3	29.3	6.4	11.9
Kharif 2005	Paddy	2.7	2.5	6.5	38.5
	Banana	25.3	22.4	11.2	7.1
Rabi 2005-06	Paddy	3.3	2.9	13.6	36.0
	Banana	27.8	24.5	11.8	12.3
Kharif 2006	Paddy	2.8	2.5	9.0	29.4
	Banana	24.5	21.6	11.9	10.0
Rabi 2006-07	Paddy	3.2	2.9	9.4	34.7

* nuts/ha in the case of coconut

Table 6. Impact of AAS on crop yields from 2003-04 to 2006-07.

4.2.4. Lessons learned

Adaptation strategies and awareness raising are particularly important for managing the risks posed by climate variability and change, not only on crops but also across all sectors that are sensitive to weather and climate. Various agroclimatic techniques have been used in the Kerala region of India to effectively manage some of the risks posed by climate to crop productivity. Expansion and further development of such techniques will be vital for the continued sustenance of agricultural production in humid tropical regions and particularly monsoonal regions. As pointed out by Prof. M.S. Swaminathan "India's strength lies in its ability to manage monsoons" instead of saying "Indian agriculture is a gamble of the monsoon".

4.3. CASE STUDY 3: Climate risk management through structural adjustment and regional relocation: A case of rice industry in Australia (S. Mushtaq, G. Cockfield, N. White, and G. Jakeman)

Climate change poses significant challenges to the Australian agricultural sector due to likely increased climate variability and increased frequency of extreme events. Climate change projections suggest that the southern part of Australia will generally become drier, while there is a likelihood of increased rainfall and the frequency and intensity of extreme events in parts of the north[58]. The possibility of climate change leading to less rainfall in southern mainland Australia, and as a result on-going water policy reforms, has triggered robust CRM strategies by agriculture sector, particularly in the rice industry. The success of any CRM strategy depends on risk management systems that reflect a more detailed understanding of the complexity inherent within human-environment interactions with more reliable future climate information and associated risks. This case study evaluates climate risks strategies employed by the rice industry in Australia.

4.3.1. Climate risk management in rice industry: Challenges and opportunities

The Australian rice industry has a relatively small number of producers, mostly within the Riverina (southern New South Wales), generating considerable export income. The rice industry and rice growers have adopted a risk-averse approach. CRM in the rice industry is based on a systematic process of managing climate and water availability risks to take advantage of opportunities to improve financial, economic, social and environmental systems. The rice processor has a global supply-chain that ensures continuous rice supplies. During years of low water availability, growers trade water and shift to low water intensive or dry land farming. This has resulted in highly variable domestic rice production. Water will probably become more expensive, less available and allocations will be less secure. The production capacity of the domestic rice industry will be significantly influenced by droughts and environmental water buy-backs. One strategy for Australian growers is to increase production in areas like the Burdekin (north Queensland) that have a sustainable water supply.

4.3.2. Evaluating structural adjustment as a CRM in Rice Farming System

Rice farmers are continually faced with pressures to adjust to changing environmental, climatic and economic conditions. Structural adjustments reflect the decisions by rice growers to adjust the size and farming operations to manage climate, environmental and economic risks[59]. The following sections provide empirical evidence of structural change in rice farming.

Farm sizes, irrigated area and rice area*:* Rice farmers have greater flexibility in farm adjustment and structural change than dryland farmers. This allows them to reduce rice area and maintain farm income from dryland crops. It is hypothesised that increased water scarcity in the Riverina has resulted in an increase in farm size while total rice and irrigated area have reduced. Fig. 10 shows that water availability has a significant impact on rice production and irrigated area and that the total operated farm area is increasing significantly. The increase in farm area can be attributed to the decreasing number of farms and temporary and permanent water trading.

Crop shifting: An assessment of the Riverina (Fig. 11) indicates that farmers are continuously adapting to climate variability and climate change by changing crop mixes and farm restructuring. Rice area per farm is generally declining and being replaced by winter dryland wheat. Some farmers have adjusted their farming operation by shifting from rice to wheat along with a larger area of dryland wheat. The reduction in rice and the increase in wheat area will have an industry-wide impact, e.g. rice mills and storage depots were closed as a result of the lower level of rice supply during 2007-08.

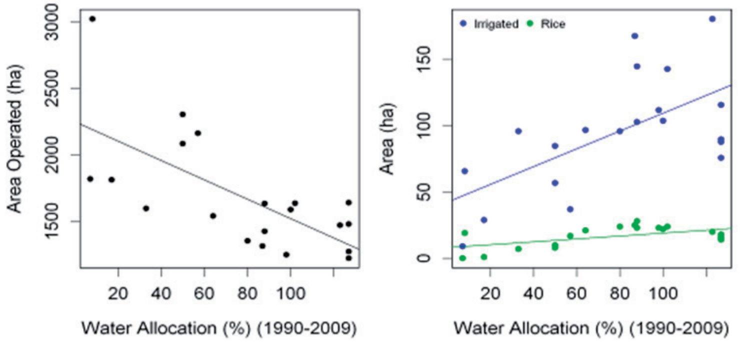


Figure 10. Area operated per farm, p=0.001, (left); rice area, p=0.003, and total irrigated area, p=0.01, (right) as a function of water availability and in Riverina, NSW, Australia; Source: ABARE Farm Survey.

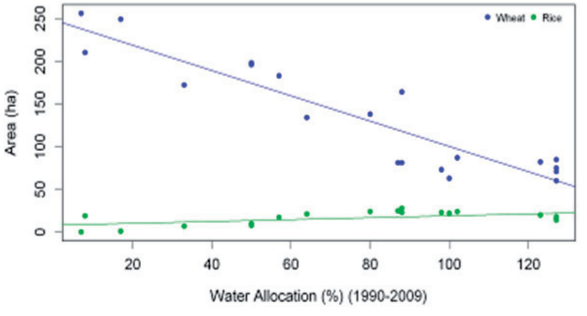


Figure 11. Wheat, $p<0.001$, and rice, p=0.01, production by area per farm as a function of water allocation in Riverina from 1992-2009, NSW, Australia; Source: ABARE Farm Survey.

Water trading: Water movement to more efficient and higher value commodities results in a consolidation of farms without showing evidence of a corporate takeover of the industry[60] showed that water markets facilitate the process of farm adjustment and structural change within the irrigation industry. To maintain a liveable income during drought periods some farmers adjust their operations by temporarily trading water to other growers to take advantage of higher water prices. Fig. 12 shows the relationship of water trading to rice area in the Coleambally Irrigation Area (CIA), Riverina (NSW). However, over the last 5 years the CIA is trading-out water to satisfy the demand of high value crops such as rice. In some instances rice farmers have had to purchase temporary groundwater in order to satisfy forward contracts[61].

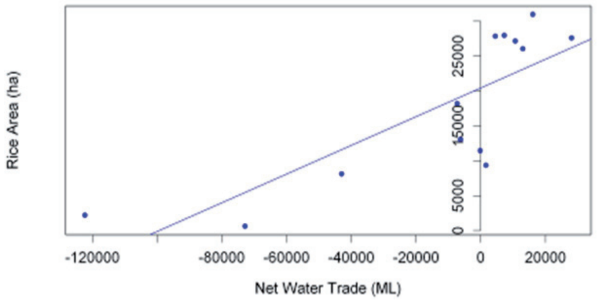


Figure 12. Relationship of net water trading and rice area in the Coleambally Irrigation Area, Riverina, NSW, Australia, p<0.001; Source: Source: ABARE Farm Survey; Coleambally Irrigation, 2009 [62] .

Financial impact: The reductions in available water have significantly influenced farm business profit and overall family income (Fig. 13) and income is sustained through off-farm activities. During 2007-2008 (<10 per cent water allocation) overall average family income per farm and farm business profit was –$27,893 and –$109,536, respectively. Clearly, this is not sustainable in the long-run. Under the anticipated climate change considerable adjustment in terms of cropping pattern or off-farm activities will be required to sustain a reasonable family income. Alternatively, relocation of farms or some of the production could be an option.

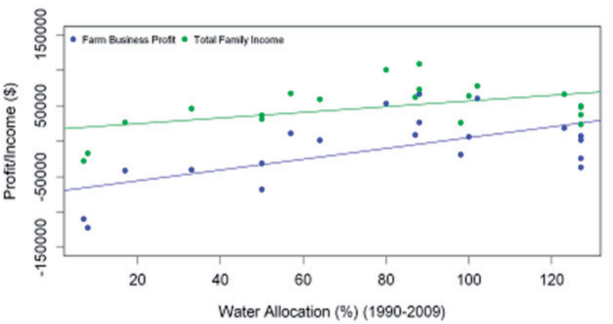


Figure 13. Relationship of water availability and farm business profit, p=0.003,and total family income, p=0.03, Riverina, NSW, Australia; Source: ABARE Farm Survey; Where: Farm business profit ($): Farm business profit equals farm cash income plus build-up in trading stocks, less depreciation expense, less the imputed value of the owner manager, partner(s) and family labour. Total family income ($): Family share of farm cash income less family share of depreciation plus all off- farm income of owner manager and spouse.

4.3.3. CRM Potential from regional relocation

A potential CRM strategy under climate change is to relocate rice production to regions with plenty of water such as the Burdekin. The CGE model was used to estimate the regional economic impact of such strategy. The model compared the net impact of shifting rice production from Riverina to Burdekin on fallow sugarcane land, assuming with no competition and displacement of sugarcane land, under 2030 and 2070 future time periods.

4.3.4. Economic impact regional relocation CRM strategy

The macroeconomic impact of relocating rice production from the Riverina to the Burdekin is presented in terms of changes in real output and real income. Relocation also affects a range of other variables (notably employment) but these are not presented here.

Table 7 summarises the projected changes in real output and real income for each region. The loss of water and consequent switching from rice to wheat is projected to reduce the real economic output and income of the Riverina. Using a 4 per cent real discount rate a cumulative decrease of –$915 million over the 59 years to 2069-70 has been estimated. The decrease in 2069-70 represents an average decrease in real economic output of around $550 per person projected to be living in the Riverina at this time (295,000 people).

As a result of rice relocation on fallow sugarcane land, real economic output increases in Burdekin. A cumulative increase of total of $759 million is projected over the 59 years to 2069-70. The increase in 2069-70 represents an increase in real economic output of around $7,000 per person projected to be living in the Burdekin at this time (18,500 people). The net movement of labour is primarily between the Riverina and Burdekin with minimal net movement of labour to/from the Rest of Australia. Consequently it is projected that there will be minimal impacts (a cumulative total of –$211 million) on the Rest of Australia.

	Real economic output ($m)			**Real income ($m)**		
	2029-30	**2069-70**	**NPV (2010-11 to 2069-70)**	**2029-30**	**2069-70**	**NPV (2010-11 to 2069-70)**
	2010-11	**2010-11**	**2010-11**	**2010-11**	**2010-11**	**2010-11**
Southern Rice	–45	–139	–915	–72	–161	–1,298
Burdekin LGA	35	131	759	58	178	1,149
Rest of Australia	–1	–21	–54	6	–26	9
Total Australia	–11	–29	–211	–8	–9	–140

Table 7. Cumulative change in real economic output and real income under scenario 1, relative to reference case for 2010-11.

4.3.5. Lessons learned and implications for CRM

With the expected reduction in water allocation, losses in rice production cannot be wholly offset by productivity gains given current production techniques and increasing temperatures and rainfall variability. The reduction in output will reduce net exports and have some impact on Gross Domestic Product (GDP), especially because of the extensive value-adding that occurs in Australia. Relocation to Burdekin is one potential risk management option, but limited agronomic knowledge and uncertainty associated with the future climate and associated financial risk pose barriers to relocation. The displacement of an existing intensively-produced crop, such as sugar would result in a much larger net national loss, also meaning that there would be a net reduction in regional income and outputs. It is concluded that there is unlikely to be a rapid increase in rice production in the north without more reliable future climate assessment to build confidence for making informed relocation decisions, infrastructure support, and R&D and extension support to enhance rice productivity and better communication.

4.4. CASE STUDY 4: Climate risk management and health: A call for user friendly climate information (Alexander von Hildebrand)

Climate change poses significant threats to climate-sensitive health outcomes, for example, through increasing the risk of malnutrition due to reduced access to food, waterborne diseases resulting from flood- or drought-related contamination of food and water, and physical and psychological from trauma following more frequent and hasher extreme events.

The appearance of infectious diseases in new geographical areas in response to warmer temperatures, increases in precipitation, and/or other climatologically-related changes, will increase the burden of malaria and dengue, and the *"combination of increasing vulnerability and risk of weather-related hazards are expected to result in more extreme events and disasters"*[63]. Children, the elderly, the poor – and amongst them, women – are expected to suffer most.

To assess the risks to human health posed by climate change and take appropriate actions to reduce their impacts, national health authorities need to know the current and potential future burden of climate-sensitive health outcomes, in order to adapt to the resulting demand for more and climate resilient health services. For this purpose, WHO has developed a Vulnerability and Adaptation Assessment Tool[63]. In order to implement it, user friendly climate and meteorological information is vital. This section will therefore discuss the importance and the urgent need for the availability of science-based climate and meteorological information as a pre-condition for managing climate risks to protect health from climate change.

4.4.1. Role of CRM in health

While the effects of climate on health are becoming better known, more needs to be done to achieve stronger engagement of the health sector and health professionals in climate-change action[66]. One key issue is managing uncertainty in climate and health sensitivity information, which poses significant problems for the health community in decision making processes. "*Climate and meteorological information are a major component of climate adaptation. Tools and knowledge systems which clarify and reduce uncertainty about the climate sensitivity of diseases, will be essential inputs to disease control policies such as malaria elimination, as well as preventing health risks from extreme weather events and climate variability.*"[67]. Therefore, to adapt well to predicted changes of the climate system, "*climate and meteorological information must be taken into greater consideration in health science, practice, and policymaking*"[66]. To this end, the World Health Organisation (WHO) works with partners and collaborating Centres to develop tools, information resources, and dialogs which facilitate climate informed management of health risks.

4.4.2. CRM in the World Health Organization

"Each year, about 3.5 million people die from malnutrition, 2.2 million from diarrhoea, 800 000 from causes attributable to urban air pollution, and 60 000 in climate-related disasters, mostly in low resource settings and also frequently in humanitarian emergency situations. Climate change brings new challenges and costs to the control of infectious diseases as some are highly sensitive to temperature and rainfall, including cholera and the diarrhoeal diseases, as well as vector borne diseases including malaria, dengue and schistosomiasis. Climate change threatens to reverse the progress that the global public health community has been making against many diseases, and increase the challenges for the humanitarian community to respond to natural, biological and social emergencies."[64].

It is clear that climate factors play an important role in the definition of some human diseases. For other diseases where climate is only considered as one of many determinants, WHO have stated that it is also important to understand the various causal pathways from climate change through to health outcomes, in order to identify opportunities to address the environmental determinants of poor health outcomes.

WHO promotes *"measures to reduce the health impacts from climate risks and associated climate change, such as strengthening public health systems based on partnerships with multi-sectoral actors, enhancing capacity of health systems to reduce risks and respond to public health emergencies, protecting hospitals and other health infrastructure from climate risks and effects of climate change, strengthening surveillance and control of infectious disease against climate risk, improving the use of early warning systems by the health secto and building public health interventions at local level to increase community resilience."*[65]. Climate information is needed and should be available in ways that users in each country can understand, especially at the local level. This would facilitate the development of, for example, *"health action plans to enhance early warning and effective response over a range of time scales: from hours or days (for flood or heat wave warnings), to weeks (for seasonal epidemics of vector-borne disease), to months (seasonal forecasts of precipitation anomalies allowing planning for flooding or drought), to years (for drought and associated food insecurity)."*[63].

4.4.3. The WHO's Vulnerability and Adaptation Assessment Tool

The Vulnerability and Adaptation Assessment Tool was developed by WHO to help manage climate risks to health (Fig. 14). It departs from gathering information on the extent and magnitude of current and future importance of climate dependent health outcomes, in order to identify policies and programmes that can prevent or reduce the severity of future health impacts.

A basic premise for the effective definition and, more importantly, implementation of a CRM process is to ensure the involvement and empowerment of the various stakeholders who will be responsible for implementing and assessing the results of the various actions to be undertaken is to establish an iterative process for monitoring and managing the health risks of climate change. Furthermore, to establish plans for communicating the CRM process *"The credibility and legitimacy of the assessment results will be increased if stakeholders and the intended end-users have been informed of and included in discussions throughout."*[63].

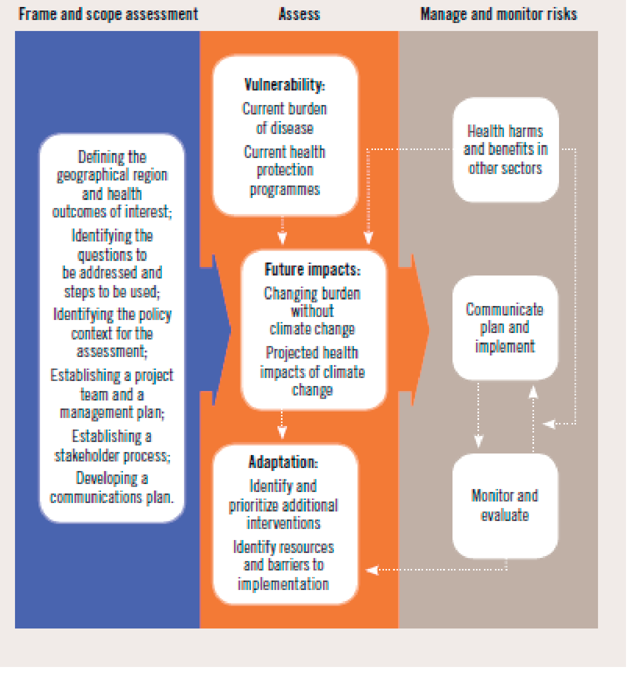


Figure 14. *WHO Vulnerability and adaptation assessment tool*

The three steps conducted in a particular assessment using the Vulnerability and Adaptation Assessment Tool will start with a description of the current burden of climate-sensitive health outcomes and of the most vulnerable populations and region. It is important at this step to address the question: What factors other than weather and climate determine vulnerability of populations and health systems? The second step involves description of the current capacity of the health sector and other sectors to address these risks to climate-sensitive health outcomes. The analysis of results from steps 1 and 2 will demonstrate gaps in the existing health system response. The third step tries to define how the burden of climate-sensitive health outcomes is likely to change over the coming decades, in order to asses the climate change vulnerabilities and their key drivers. This is done firstly irrespective of climate change, and then secondly taking into account the likely health impacts of climate change, including the most vulnerable populations and regions, over the next decades and in the longer term. The analysis of results will help determine how well the health system is, or is not, prepared for example to changes in demand due to changes in the geographical distribution, and incidence or timing of climate-sensitive health outcomes. Gaps in health system response identified earlier, in steps 1 and 2, may be expanded upon during this step.

The information provided from implementation of the Vulnerability and Adaptation Assessment Tool enables the appropriate health experts to define the nature of additional public health policies and programmes that will likely be needed for effective health management, in order to address possible additional burden of adverse health outcomes due to climate change, and to define what policies and programmes are needed in other sectors to protect health.

Throughout the assessments, it is important to take into account that "*Future vulnerabilities may be different from current vulnerabilities because of changes in public health and health-care policies, governance and institutions, socioeconomic development, availability of human and financial resources, and other factors. Impacts can change with both changing vulnerabilities and environmental changes. Public health policies, programmes and interventions to address vulnerabilities and impacts will need to be revisited regularly to ensure continuing effectiveness in a changing climate*"[63].

4.4.4. The need to improve integration of baseline health and climate data for CRM

To implement well the Vulnerability and Adaptation Assessment Tool climate and weather data resources are required that are appropriate for health sector applications. This information is key to enable adequate answers to be established for the following questions:

- Which regions and populations in a country are the most vulnerable to climate and climate change?
- What are the health risks posed by climate change over the next decades and the longer term?
- How well is the health system prepared for changes in demand due to changes in the geographical distribution, incidence or timing of climate-sensitive health outcomes?

To enable the analysis of relationships between current and past weather/climate conditions and health outcomes, relevant data is required both on health and climate. To date there have been only a few studies which have combined the most appropriate health and weather/climate data available, for example, from ministries of health and national meteorological and hydrological services, respectively. Improved integration of these data and the expertise that their host organizations provide would significantly improve the analyses necessary for CRM.

The Vulnerability and Adaptation Assessment Tool invites the use of spatial mapping to describe the geographical distribution of current or projected future vulnerabilities and hazards. "*A geographical perspective and the use of geographical information systems (GIS) offer opportunities to show current distributions of, for example, vulnerable populations and the spatial relationship to disease vectors, river basins prone to flooding, health facilities, and other important variables of interest to public health officials*"[63].

It is important to keep in mind the following caution when models are used to project the health risks of climate change. "*Modeling can be a complex undertaking requiring highly technical expertise and specific data inputs that take time and effort to acquire. The capacity to design and run models to project health impacts can be developed through training courses and other mechanisms. A goal of the assessment could be to build research capacity and increase the availability of models to project health impacts in future studies*"[63].

It is all about what happens locally. "*Risk management works best when tailored to local circumstances. Combining local knowledge with additional scientific and technical expertise helps communities reduce their risk and adapt to climate change (robust evidence, high agreement)*"[63].

The health impacts that may occur in a particular location will depend on the actual climate and climate changes experienced and the vulnerability of the community and region. Qualitative data may allow changes to be assessed over short time periods, but, it is clear that "*Models are generally used to quantitatively estimate how the health risks of climate change could increase or decrease over time, particularly for longer time periods*"[70]. Indeed, "*models can explore the range of potential impacts of a changing climate in the context of other drivers of population health to better understand where, when and in what population groups' negative health outcomes could occur*."[63].

For decision-makers, it is important to have certainty that their decisions are "climate-proof". The availability and use of locally specific *c*limate and meteorological information relevant to health outcomes is vital for these decisions[63].

4.4.5. Examples of good practice in CRM for health

WHO developed a Technical Document on Vulnerability and Adaptation Assessment Tool with input from over 20 countries that designed to provide basic and flexible guidance on conducting national or sub-national vulnerability and adaptation assessments. The document provides examples and references for users[63]. Some examples of practice provided in this document that would benefit from better access to climate information are the following:

Exercise to plot climate-sensitive diseases in geographically defined populations:

Motivated by concerns about health vulnerabilities related to climate change, a joint WHO/WMO/UNEP/UNDP workshop was conducted in the Hindu Kush–Himalaya regions[66]. Only crude estimates of the current burden of climate-sensitive diseases were available because of the lack of health surveillance data at the local level. Therefore, a qualitative assessment was conducted as a first step to generate this information. Expert judgment was used to determine the extent to which climate-sensitive diseases could be a concern in populations in mountainous and non-mountainous regions of six countries (Table 8).

Country	Afghanistan	Bangladesh	Bhutan	China	Nepal	India
Heatwaves	M-P	P	-	P	P	P
Flood deaths/morbidity	P	P	P	P	P	P
Glacial lake floods	M-P	-	M-P	M-P	M-P	M-P
Flash floods	M-P	P	M-P	M-P	M-P	M-P
Riverine floods	P	P	-	P	P	P
Vector-borne disease	P	P	P	P	P	P
Malaria	P	P	P	P	M-P	P
Japanese encephalitis	-	P	-	P	P	P
Kala-azar	P	-	-	-	P	P
Dengue – P P P – P						
Waterborne diseases	M-P	M-P	M-P	M-P	M-P	M-P
Water scarcity, quality	M-P	P	P	M-P	M-P	M-P
Drought-related food insecurity	M-P	P	–	M-P	–	M-P

M-P health determinant or outcome occurs in mountainous and non-mountainous (i.e. plains) areas;
P health determinant or outcome occurs only in non-mountainous areas;
– health determinant or outcome is not present in the country (WHO/SEARO, 2006).
Source: Kristie L. Ebi, Rosalie Woodruff, Alexander von Hildebrand and Carlos Corvalan (2007). Climate change-related health impacts in the Hindu Kush-Himalayas. Ecohealth, 4:264–270[67]

Table 8. Current climate-related health determinants and outcomes in the Hindu Kush–Himalaya regions

Qualitative estimates of future health impacts of climate change using expert judgment:

During the assessment of health risks and responses in the first Portuguese national assessment, a qualitative assessment was conducted of the possible impacts of climate change on vector-borne diseases, including malaria, West Nile virus, schistosomiasis, Mediterranean spotted fever and leishmaniasis; the latter two are endemic to Portugal.

Although human cases of vector-borne diseases have generally decreased over recent decades, many competent vectors are still present in Portugal. Disease transmission risk was categorized qualitatively based on vector distribution and abundance and pathogen prevalence. Four brief storylines of plausible future conditions were constructed based on current climate and projected climate change, and assuming either the current distribution and prevalence of vectors and parasites, or the introduction of focal populations of parasite infected vectors. These storylines were discussed with experts to estimate transmission risk levels. For Mediterranean spotted fever, the risk of transmission was high under all storylines, suggesting that climate change is likely to have a limited impact. For the other diseases, the risk level varied across the storylines. For example, the risk of leishmaniasis varied from medium under current climate to high under both climate change storylines. The risk of schistosomiasis varied from very low (current climate and current vector distributions) to medium (climate change and focal introduction).[68]

Climate and health observatory: Innovations in data sharing, communications and partnership building in Brazil, by Christovam Barcellos, FIOCRUZ Brazil:

Given the complexity of processes that drive climate change impacts on human health, it is necessary to gather data from different institutions in order to understand monitor and project these outcomes. These data include not only climatic and human health variables but also trends in socio-demographic and environmental factors and institutional capacity. The experience of the Brazilian Climate and Health Observatory demonstrates how to bring multiple institutions and stakeholders together to support actions to decrease human health vulnerability to climate change. The observatory has the following functions: gathering available data on climate, environment, society and health; conducting situation analyses and identifying trends and patterns related to climate change impacts on health (e.g. semi-qualitative graphs and maps); providing information to national alert systems and for monitoring health emergencies associated with extreme weather events; supporting research and development on climate and environmental changes and associated health impacts; promoting the active participation of civil society and citizens on issues related to climate change, environmental degradation and health impacts (e.g. news reports, commentaries, photographs).

The observatory project is supported by the Brazilian Ministry of Health and PAHO and is coordinated by the Oswaldo Cruz Foundation. Through workshops, participants developed institutional agreements for sharing data and identified specific data formats, timescales and spatial resolution to be used at the observatory.

Climate change and health impacts to be addressed first include direct impacts from heatwaves, floods and droughts; the expansion of vector-borne diseases; the vulnerability of water supply and sanitation systems, and the increasing risk of water-related diseases; and the interaction between climate change and impacts on air pollutants that increase the risks of respiratory diseases.

City of Quito, Ecuador Climate change mitigation and adaptation plan:

In 2012, the Municipal Council of the city of Quito, Ecuador developed a climate change mitigation and adaptation plan. It aims at reducing GHG emissions by 15 per cent relative to projected growth, and social environmental and economic climate vulnerability by 20 per cent. The plan will create innovative mechanisms for reducing the carbon footprint of the private sector. The municipal plan strengthens the generation and management of climate information and knowledge in close collaboration with the national climate institution, National Institute of Meteorology e Hydrology (INAHMI). Source: http://www.quitoambiente.gob.ec/home/noticia.php?idNoticia=108

Capacity building: Climate Information for Public Health:

In collaboration with the Pan-American Health Organization (PAHO), the International Research Institute for Climate and Society, in partnership with the Center for International Earth Science Information Network (CIESIN) and the Mailman School of Public Health initiated a two-week course on Climate information for Public Health in 2008[69]. It is held annually and "*provides a balance of concepts and methods from the health and climate communities using an approach deeply oriented toward methodology, gathering and using evidence for decision-making in order for the participants to get in-depth knowledge and skills in decision-making for health-care planning of climate sensitive diseases*"[70].

The course helps participants to recognize the role climate plays in driving the infectious disease burden and public health outcomes, understand management and data integration as an opportunity to improve the decision making process in Public Health and realize the benefits and limitations of different climate and environmental data sources including remotely sensed data, meteorological data and climate predictions.

4.4.6. Lessons learned and the way forward

As pointed out by the United Nations Task Team on Social Dimensions of Climate Change,

"Global and regional one-size-fits-all climate analysis may not reflect the reality of a particular community or country and can under- or over-emphasize risks relevant to certain communities. The outcomes of downscaling should be incorporated where relevant and feasible, and combined with complementary mappings that may include social impact assessments and vulnerability maps, in order to identify social climate-induced hotspots (places where particularly severe problems may need to be addressed) and their intersection with other kinds of vulnerabilities such as lack of access to preventive and curative health services, that can reduce health vulnerability to climate change"[71].

Health sectors in countries need to possess tools to conduct climate change vulnerability and adaptation assessments. "*The goal remains to better understand how climate variability and climate change can and do affect health risks today and in the future, in order to better inform policies and programmes that can protect public health*"[70]. However, once there is motivation for action, *"decision makers need to know the magnitude of potential risks and identify a range of options (including their feasibility, benefits, acceptability, effectiveness and costs); the availability of resources and their distribution across the population; and the structure of critical institutions, including the allocation of* decision-making authority"[72].

The call for the production, availability, delivery and application of locally specific, science-based climate and meteorological information is a fundamental requirement for improving the application of CRM to address health risks. Integration of this information with appropriate health information and data will provide an opportunity, but also a challenge, to health authorities to demonstrate leadership within and outside the sector on mitigation and adaptation to climate change in order to protect health.

5. Conclusions and recommendations

From global to local levels, public and private sector institutions are seeking the tools and knowledge for CRM (including capacity building and adaptation). People require climate information over wide ranges of time and space scales for planning and operational purposes. It is imperative, therefore, to ensure that they have the highest quality and widest possible range of products, information (including about uncertainties), and guidance on how the information can be used to provide optimal results and ensure appropriate decisions are made.

Effective CRM must be founded on scientifically sound risk assessment techniques that develop understanding, and where possible, quantification of the risks associated with natural hazards, socio-economic vulnerabilities and their impacts. Climate risk assessment, in turn, requires quality assured historical, real-time and future-projected data on climate-related hazards and socio-economic vulnerabilities, with reliable regional detail. Changing patterns of climate hazards make those data needs even more imperative. Understanding the challenges posed by climate change to longer-term strategic planning and investments (e.g. infrastructure planning and retrofitting based on building codes as a 100 year flood may become a 30 year flood), would help providers frame future products and services, and users build resilience to future as well as current climate. To be effective, processes adopted for CRM must be perceived by individuals, communities and governments as providing real possibilities of improving their living standards through awareness, adaptation, prevention and increased resilience to climate impacts. If authorities set up sustainable, coherent and participative CRM plans that are 'owned' at local level, these will provide strong foundations for the development of successful adaptation strategies.

5.1. Common constraints to Climate Risk Management

A number of constraints limiting the use of climate information for decision-making by sectors have been discussed throughout this chapter. Some of the most common constraints to be addressed before implementation of sustainable CRM systems are:

- Limited national capabilities for climate data processing, analysis, modeling and the generation of information and forecast services including sectoral applications in strategic sectors such as agriculture, health, water resources and others.
- Lack of capacity to communicate information between NMHS and national agencies and authorities.
- Limited capability of governmental institutions and sectoral institutions to identify their climate information requirements.
- Limited coordination between local institutions and agencies.
- Weak or non-existent planning of the investment budget for actions aimed at prevention of and preparedness for national disasters at the level of national and subnational governments.

- Weak or non-existent accountability system for encouraging and mandating use of available scientific information for appropriate risk reduction measures.
- Limited involvement of the private sector to engage with stakeholders in the development of risk management measures.
- Limited knowledge of climate and limited training in application and use of climate information and products by users in many sectors;
- Problems in identifying threats to carry out works on risk reduction because of limited use of the necessary tools for assessment.
- Limited synergy among national institutions to share climate and climate-relevant information at national level.
- The absence of an efficient communication system on extreme climate events to disseminate information such as warnings and advisories.
- Overly technical language in some climate information and products, making them difficult to understand by lay people.
- The lack of a culture of prevention of damage and proper maintenance, supported by finances, for the vulnerable physical and social infrastructure.

Even though climate scientific knowledge and probability modelling have advanced significantly over the last few past decades, especially with respect to the understanding of climate variability and change, the level of uncertainty inherent in probabilistic climate products has tended to make their communication by scientists for integration and understanding by users of the information more challenging. While the scientific community regards uncertainty as an inherent property of the climate system, which can be assessed through probability analysis, decision makers may consider uncertainty as a barrier to decision making, especially when other socio-economic and political variables also need to be considered. The result is a complex and confused integration situation among actors which requires careful communication to ensure the complex information is understood by all and appropriate decisions are made. To overcome this limitation, it is recommended that the climate scientific and stakeholder communities create an agenda sustained by the transfer of requirements, knowledge, tools and instruments allowing the formation of a community of practitioners with analytical skills and sharing basic agreements for action. To improve disaster and climate risk governance, it is proposed that decision-makers should also assume responsibilities concerning the understanding of risks within their sectors and regions, and consider the need to integrate DRR management and CRM using the wide range of currently available instruments and development mechanisms.

The role of the NMHS is decisive within the national efforts to cope with the opportunities and impacts of climate variability and change and to encourage effective CRM. The activities of NMHS regarding climate observations, data management, analysis maintaining a 'climate watch' and forecasting are strictly necessary to estimate the "near climate change", estimate indexes and provide local trends to be applied for risk management and development of adaptation plans. The nature of climate risks, including climate change implies accurate monitoring efforts under a rigorous methodology and standards that only NMHSs can provide in a sustainable way. The current efforts of NMHSs in weather and climate forecasts, the analysis of extremes and other prediction services are valuable for CRM and consequently a solid basis to work on climate change estimations at local level where global projections are not necessary applicable. To encourage communities to engage with such climate services and to enhance the relationship with the vulnerable local populations, a number of recommendations, that will underpin a "new business model" for conducting effective CRM at local levels, are proposed for the operations of the NMHSs of the WMO.

5.2. Recommendations for NMHSs for improving CRM at local level

For improving CRM at local levels, it is recommended that NMHSs:

1. Focus on users, and enhance collaborations with CRM communities in order to assess their needs and address these through provision of high quality and opportune climate services, including through organizing or participating in both face-to-face and on-line Regional and National Climate Outlook Forums which offer direct interaction with users including sectors and the media, and participate in capacity development of users through training in climate matters, adaptation and in the use of climate products.
2. Tailor climate products for CRM end users, including planning departments, local authorities, government agencies involved in environment and risk management, based on their requirements.
3. Monitor the climate and its evolution, conduct Climate Watch programmes, develop regular information for users on the past and current states of the climate, and couple these with reliable, user-friendly predictions for upcoming seasons, as part of a climate services culture.
4. Build and sustain observing networks to provide the data needed for a range of climate services for CRM; conduct data rescue exercises to enhance digital climate databases.
5. Promote training and development of their meteorological staff in diverse aspects of climatology and climate services and in CRM.
6. Enhance climate research, development of climate indexes and other analysis, within their operational activities.
7. Combine climate products with other geospatial information related to vulnerability derived from other institutions for development of more decision-ready, actionable products (this will require strategic alliances, in "win–win" relationships with co-benefits.
8. Enhance the liaison with local communities, communitarian networks and local media for efficient dissemination of tailored products applicable for CRM.
9. Seek and act on user's feedback for product evolution and improvement.

5.3. Moving forward with GFCS

As previously noted, the WMO, along with partnering organizations and with the support of Member countries, has embarked on a new era for climate services, with the decision to implement the Global Framework for Climate Services (GFCS). This decision was reached in the full understanding that GFCS success will require strengthening of observations and monitoring, research, modeling and prediction, the Climate Services Information System (the operational 'engine' of the GFCS which includes the NMHSs of WMO's 189 Member countries), and much improved interaction with users. A great deal of capacity development will be required to ensure the capability of climate providers to deliver quality information and the ability of the users to take up and apply the information. Indeed, in many places this is already underway, such as the provision of climate services in developing countries via the Regional Climate Outlook Forum supported by the WMO Climate Information and Prediction Services (CLIPS) project (Semazzi, 2011)[73].

The curent effort of the CCl Task Team on CRM, which has provided the motivation for this chapter, marks the beginning of a new collaboration on CRM. As the GFCS is implemented and improves over time, much will be learned, and the concept of CRM, *inter alia*, will be tested and evolve, particularly with the cooperation of the organizations affiliating themselves with the GFCS. We hope this chapter provides techniques, case studies, good practices and guidelines that will be useful for this endeavour.

Author details

R. Martínez*
International Research Center on El Niño (CIIFEN). Guayaquil, Ecuador

D. Hemming*
Met Office Hadley Centre, UK Meteorologcial Office, Fitzroy Road, Exeter, UK

L. Malone*
World Meteorological Organization (WMO), Geneva, Switzerland

N. Bermudez
United Nations Development Programme (UNDP), Quito, Ecuador

G. Cockfield and S. Mushtaq
Australian Centre for Sustainable Catchment, University of Southern Queensland, Toowoomba QLD, Australia

A. Diongue
Agence Nationale de l'Aviation Civile et de la Météorologie (ANACIM), Senegal

J. Hansen
International Research Institute for Climate and Society (IRI), New York, USA

A. Hildebrand
Panamerican Health Organization (PAHO), Quito, Ecuador

K. Ingram
University of Florida, Gainsville, FL, USA

G. Jakeman

---

* Leading Authors

ACIL Tasman Pty Ltd. Canberra, Australia

M. Kadi
African Centre of Meteorological Application for Development (ACMAD), Niamey, Niger

G. R. McGregor
School of Environment, University of Auckland, New Zealand

P. Rao
Academy of Climate Change Education & Research, Kerala Agricultural University, Vellanikkara, Thrissur, Kerala, India

R. Pulwarty
National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, USA

O. Ndiaye
Service Départemental Du Développement Rural, Kaffrine, Sénégal

G. Srinivasan
Regional Integrated Multi-hazard Early warning System (RIMES), Bangkok, Thailand

Eh. Seck and N. White
AgriScience Qld, Department of Agriculture, Fisheries and Forestry, Toowoomba QLD, Australia

N. White
Dept. of Biological and Physical Sciences, University of Southern Queensland, Toowoomba QLD, Australia

R. Zougmoré
Climate Change Agriculture and Food Security (CCAFS), West Africa, Bamako, Mali

6. References

[1] Ritchey T. Wicked Problems: Structuring Social Messes with Morphological Analysis. Swedish Morphological Society; 2011.
[2] World Meteorological Organization. Climate knowledge for Action: A Global Framework for Climate Services – Empowering the most vulnerable. WMO-No. 1065. 2010.
[3] Jones R. An Environmental Risk Assessment/Management Framework for Climate Change Impact Assessments. Natural Hazards. 2001; 23: 197-230.
[4] Taubenböck H, Post J, Roth A, Zosseder K, Strunz G, Dech S. A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing. Nat. Hazards Earth Syst. Sci. 2008; 8:409-420.
[5] Wilby R, Troni J, Biot Y, Tedd L, Hewitson BC, Smith DM, Sutton RT. A review of climate risk information for adaptation and development planning. International Journal of Climatology. 2009; 29:1193-1215.
[6] Hemming D, Boorman P, Hartley A, Kay G, McInnes R, McNeall D. Identification of dangerous thresholds in regional climate and other biophysical quantities related to key impacts. Report (Deliverable # D2.4.3) for the UK Department for Energy and Climate

Change and Department for Environment, Food and Rural Affairs under the Hadley Centre Climate Programme. 2011; GA01101 2010-2012.

[7] Johnson L, Weaver C. [Internet]. 2009. Available from: http://www.utexas.edu/lbj/directory/faculty/catherine-weaver.

[8] Breuer NE, Fraisse CW, Hildebrand PE. Molding the pipeline into a loop: the participatory process of developing AgroClimate, a decision support system for climate risk reduction in agriculture. Journal of Service Climatology.2009; 1: 1-12.

[9] Bartels W, Arnold J, Breuer NE, Furman CA, Staal L, Irani TA, Jones JW. Supporting dialog and learning among stakeholders through climate working groups. Southeast Climate Consortium Technical Report Series. SECC Technical Report. 2011; 12-001:00-00.

[10] Bartels W, Furman CA, Royce F, Ortiz B, Zierden D, Fraisse CW. Developing a learning community: Lessons from a climate working group for agriculture in the southeast USA. Southeast Climate Consortium Technical Report; 2012. Series: SECC Technical Report 12-002:00-00.

[11] Crane T, Roncoli C, Breuer N, Broad K, Paz J, Fraisse C, Ingram KT, Zierden D, Hoogenboom G, and O'Brien JJ. Collaborative approaches to the development of climate-based decision support systems. Proceedings of the 88th AMS Annual Meeting, New Orleans, LA, January 20-24, 2008.

[12] Fraisse CW, Breuer N, Bellow JG, Cabrera V, Hatch U, Hoogenboom G, Ingram KT, Jones JW, O'Brien JJ, Paz J , Zierden D. AgClimate: A climate forecast information system for agricultural risk management in the southeastern USA. Computers and Electronics in Agriculture. 2006; 53(1):13-27.

[13] Hammond A, Adriaanse A, Rodenburg E, Bryant D, Wood R. Environmental indicators: A Systematic Approach to Measuring and Reporting on Environmental Policy Performance in the Context of Sustainable Development. World Resources Institute. 1995.

[14] Hák, T., Moldan, B. & Dahl, A.L. (eds). (2007). Sustainability Indicators. A scientific assessment. SCOPE 67, Island Press. 2007.

[15] OECD. Handbook on Constructing Composite Indicators: Methodology and User Guide OECD, European Commission, Joint Research Centre. OECD Publishing. 2010.

[16] EEA-JRC-WHO. Impacts of Europe's changing climate. — 2008 indicator-based assessment. Joint EEA-JRC-WHO report. EEA Report No 4/2008.

[17] Smeets E, Weterings R. Environmental indicators: Typology and overview. European Environment. Agency, Technical report No. 25, Copenhagen, 1990; 19.

[18] Wascher D, Perez-Soba M. (eds.) Learning from European Transfrontier Landscapes – a project in support of the European Landscape Convention. Alterra report 964, LANDSCAPE EUROPE, Wageningen. 2004.

[19] Rochet M, Rice J. Do explicit criteria help in selecting indicators for ecosystem-based fisheries management? 2005.

[20] Marx S, Weber E. Decision making under climate uncertainty: The power of understanding judgment and decision processes. In: T. Dietz & D.C. Bidwell (Eds.), *Climate Change in the Great Lakes.* MSU Press. 2012; Chapter 5: 99-128.

[21] IPCC. Special Report on Managing the Risks of Extreme Events (SREX) and Disasters to Advance Climate Change Adaptation. Cambridge Press. 2012.

[22] Pulwarty R. Climate and water in the West: Science, information and decision-making. Water Resources. 2003; 124: 4-12.

[23] Weichselgartner J. Disaster mitigation: the concept of vulnerability revisited. Disaster Prevention and Management. 2001; 10(2): 85-94.
[24] CIIFEN. Manual de buenas prácticas para adaptación al cambio climático en zonas costeras. Una guía para la formulación, diseño e implementación de medidas de adaptación al cambio climático en zonas costeras. Programa Costas y Bosques Sostenibles. 2011; 1 1:63
[25] Ministerio del Ambiente del Ecuador. Second National Communication to the UNFCCC. 2010.
[26] Lindell M, Perry R. Communicating Environmental Risk in Multiethnic Communities. Thousand Oaks, CA: Sage Publications. 2004.
[27] Fischhoff B. Transaction analysis: A framework and an application to insurance decisions. Journal of Risk and Uncertainty 1993; 7: 53-69.
[28] Cutter S, Mitchell T. Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. Annals of American Geographers. 2000; 90(4): 713-737.
[29] Pulwarty R. Communicating Agrometeorological Information, including Forecasts, for Agricultural decisions. WMO Guide to Agricultural Meteorological Practices (GAMP) 27 WMO Geneva. 2007.
[30] Goddard L, Aitchellouche Y, Baethgen W, Dettinger M, Graham R, Hayman P, Kadi M, Martinez R, Meinke H. Providing Seasonal-to-Interannual Climate Information for Risk Management and Decision-making. Procedia Environmental Sciences World Climate Conference - 3. 2010; 1: 81-101.
[31] CIIFEN. Implementación de un Sistema de Información multimodal de eventos climáticos extremos en el Ecuador. Quinto Plan de Acción DIPECHO Sudamérica del Departamento de Ayuda humanitaria de la Comisión Europea. 2008; 1: 1-48.
[32] Campbell J. Modeling the performance prediction problem in industrial and organizational psychology. In: M. D. Dunnette & L. M. Hough (Eds.), Handbook of Industrial and Organizational Psychology. Palo Alto, CA: Consulting Psychologists Press, Inc. 1990; 1: 687-732.
[33] Mileti, D.S. and C.Fitzpatrick. The Great Earthquake Experiment: Risk Communication and Public Action. Boulder, CO: Westview Press. 1990.
[34] Dennis M, O'Brien P. Warnings During Disaster: Normalizing Communicated Risk. Social Problems. 1992; 39(1): 40-57.
[35] Intergovernmental Panel on Climate Change. Climate Change 2007 - Impacts, Adaptation and Vulnerability. Cambridge University Press, Cambridge . 2007.
[36] UNISDR Terminology. [Internet], 2009. Available from: http://www.unisdr.org/files/7817_UNISDRTerminologySpanish.pdf
[37] Traore A, Rogers D. Towards a multi-hazard early warning and response system in West Africa. Bonn, Germany: EWC-III. 2006.
[38] IFRC. World Disasters Report 2009: Early warning, early action. International Federation of the Red Cross. 2009; 68-93.
[39] Wisner B, Blaikie P, Cannon T, Davis I. At risk. natural hazards, people's vulnerability and disasters. New York: Routledge. 2004.
[40] Hewitt K. Interpretation of Calamity: From the Viewpoint of Human Ecology. Boston, Allen.1983.

[41] Verver M. The East African food crisis: Did regional early warning systems function? The Journal of Nutrition. 2011 n.111.150342.
[42] UNISDR. Hyogo Framework for Action 2005-2015: Building the resilience of nations and communities to disasters. 2007.
[43] UNEP. UNEP's activities in 2006. Year of Publication: 2007. ISBN No: 978-92-807-2801-9.
[44] Pulwarty R, Verdin J. Crafting early warning information systems: the case of drought. In: Measuring Vulnerability to Natural Hazards: Towards Disaster Resilient Societies. 2012 (Birkmann, J, ed. (United Nations University Press, New York, forthcoming).
[45] Glantz M. Usable Science 9: El Niño Early Warning for Sustainable Development in the Pacific Rim and Islands. Report of workshop held 13–16 September 2004 in the Galapagos Islands, Ecuador. Boulder, CO: ISSE/NCAR. 2005.
[46] Webster P, Hopson T, Hoyos C, Subbiah A, Chang R, Grossman R. A three-tier overlapping prediction scheme: Tools for strategic and tactical decisions in the developing world. Predictability of Weather and Climate, Ed. T. N. Palmer and R. Hagerdorn, Cambridge University Press. 2006; 718: 645-673.
[47] Hopson T. Operational Flood-Forecasting for Bangladesh, University of Colorado Ph.D. Dissertation, Department of Astrophysical, Planetary, and Atmospheric Science. 2005.
[48] Subbiah A, Fakhruddin E. CFAB: Climate Forecast Applications for Communities. Asian Disaster Preparedness Centre (ADPC). Report, January, 2012; pp. 23.
[49] Clark M, Olson J. Nursing within a Faith Community: Promoting Health in Times of Transition. Thousand Oaks, CA: Sage Publications, Inc. 2000.
[50] Misselhorn A. What drives food insecurity in southern Africa? A meta-analysis of household economy studies. Global Environmental Change. 2005; 15: 33-43.
[51] Morgan R. The development and applications of a drought early warning system in Botswana. Disasters. 2005; 9: 44-50.
[52] Pulwarty R. Communicating Agrometeorological Information, including Forecasts, for agricultural decisions. WMO Guide to Agricultural Meteorological Practices (GAMP). 2007; 27.
[53] Birkmann J, Chang Seng S, Suarez D. Adaptive Disaster Risk Reduction, Enhancing Methods and Tools of Disaster Risk Reduction in the Light of Climate Change, DKKV Publication. 2011; 43.
[54] UNISDR. Global Assessment Report on Disaster Risk Reduction. UN International Strategy for Disaster Reduction. 2011.
[55] Bohle, H., Keck, M, and P. Sadakpolrak 2009: Adaptive Food Governance. In: IHDP Update 3, Magazine of the International Human Dimensions Programme on Global Environmental Change: 543-58.
[56] Sakdapolrak P. Water Related Health Risks, Social Vulnerability and Pierre Bourdieu. In: UNU-EHS; 2007 (6): 50-60.
[57] WMO, ACMAD. Climate Forecast in Africa. WMO/TDNo. 927, October 1998.
[58] IPCC. Working Group II Impacts, Adaptation and Vulnerability-Chapter 17 Adaptation Practices; 2007 and Technical Report: Climate Change and Water; 2008.
[59] Mc Coll J, Young C. Managing Change: Australian structural adjustment lessons for water. Policy and Economic Research Unit CSIRO Land and Water; 2005 (Report No. 16/05).
[60] Bjornlund, H., McKay, J., 1999. Do Permanent Water Markets Facilitate Farm Adjustment and Structural Change Within Irrigation Communities? Rural Society.

[61] Sanders F, Burt M. Subcritical Contraction for Improved Open-Channel Flow Measurement Accuracy with an Upward-Looking ADVM (with Charles M. Burt and Brett F. Sanders), Journal of Irrigation and Drainage Engineering; 2010.
[62] Coleambally Irrigation [Internet]. Australia 2009. Available from: http://new.colyirr.com.au/
[63] Health issues gain traction at UN Climate Conference [Internet] WHO; 2011. Available from:http://www.who.int/globalchange/mediacentre/events/2011/durban_conference_update/en/index.htm
[64] Protecting health from climate change - World Health Day 2008 [Internet] WHO; 2008. Available from: http://www.who.int/world-health-day/toolkit/report_web.pdf.
[65] Disaster Risk Management for Health Fact Sheet [Internet] WHO; 2011. Available from: http://www.who.int/hac/events/drm_fact_climate_risk_management.pdf
[66] Confalonieri U, Menne B, Akhtar R, Ebi K, Hauengue M, Kovats R, Revich B, Woodward A. Human health. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds.,Cambridge University Press, Cambridge, UK; 2007: 391-431.
[67] Ebi K, Woodruff R, Hildebrand A, Corvalan C. Climate change-related health impacts in the Hindu Kush-Himalayas. Ecohealth;2007(4): 264–270.
[68] Casimiro E, Calheiros J, Duarte F, Kovats, S. National assessment of human health effects of climate change in Portugal: Approach and key findings. Environmental Health Perspectives; 2006 (114): 1950–1956.
[69] Climate Risk Management Overview [Internet] WMO. Available from: http://www.wmo.int/pages/themes/climate/risk_management_overview.php
[70] IRI Resources tools [Internet]. Available from: http://portal.iri.columbia.edu/portal/server.pt?open=512&objID=1094&mode=2
[71] The Social Dimensions of Climate Change, Discussion draft [Internet] 2011. Available from: http://www.who.int/globalchange/mediacentre/events/2011/social-dimensions-of-climate-change.pdf
[72] IPCC. Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation; 2010.
[73] Samazzi F.H.M. Framework for climate services in developing countries. Climate Research; 2011 (47): 145-150.

Risk Management at the Latin American Observatory

Á.G. Muñoz, D. Ruiz, P. Ramírez, G. León,
J. Quintana, A. Bonilla, W. Torres,
M. Pastén and O. Sánchez

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/2568

1. Introduction

The *Observatorio Latinoamericano de Eventos Extraordinarios* (OLE[2]), or *Latin American Observatory*, is a regional collaborative network that, in addition to the existing infrastructure in each country, aims to ultimately increase the efficiency of the decision-making processes, especially in terms of getting more accurate environmental information and exchanging experiences on data, methodologies and scientific products, all of which are done with a standardized methodology and a Web-sharing service.

Present partners of this collaboration are national weather services, universities and research institutes of Central America, the Andean countries and from Southeast South America (see Figure 1). Nowadays, the Centro de Modelado Científico (CMC) of Universidad del Zulia, the founder of the initiative, regionally coordinates the OLE[2]. The coordinating role involves, among other tasks, the suggestion of (a) methodologies for the provision of scientifically based tools and products, (b) mechanisms for the successful use of the environmental information in the policy making process at the different levels (e.g. national and province governments, private sector), (c) facilitation of continuous and effective communication of the different partners and (d) the interchange of experiences and products (by means of the Observatory's products web interface[1], wiki page[2], forum[3] and email list[4]). For more information see references [1] and [2].

[1] http://ole2.org
[2] http://mediawiki.cmc.org.ve
[3] http://laft.cmc.luz.edu.ve:8080/foro_ole2
[4] ole2@cmc.org.ve

The goal of the OLE² is to monitor and forecast key environmental variables and develop accurate products based on different scientific tools in order to help decision makers improve risk management and set up efficient early warning systems. The Observatory provides several model outputs for meteorological, seasonal, and hydrological forecasts, 5-day high-resolution oceanographic prediction for the eastern Pacific, droughts, fire and flood indices, ecosystem dynamics (like duckweed/algae occurrence in Lake Maracaibo), and climate and health applications (e.g., regarding malaria), among others.

The Observatory is organized in interconnected working groups (WGs) or axes: namely, meteorology; climatology; hydrology; air quality; climate and health; and climate change and variability. Each WG has its own methodologies and products, but all of them exchange information and results with other axes and countries if necessary. It is also regionally divided in three branches: the Observatorio Andino, Observatorio Centroamericano and Observatorio del Sudeste de Sudamérica (Andean, Central American and the South Eastern South America Observatories). They operate using the infrastructure already in existence in the different regions by consolidating the collaboration networks through the recognition of common needs and the provision of interdependent solutions. The final products, be it a seasonal forecast or an early warning system for the last-mile user, is built between different partners by sharing their experiences and expertise on different fields. This has created a strong identification between the participating institutions and the Observatory's products. This fact and the continuous need of collaboration to provide regionally successful decision-making tools on time, have proven a natural way to guarantee the sustainability of the initiative, originated in 2007 in Venezuela as a national observatory.

Advances and difficulties are shared and discussed through e-mail lists and videoconferences. The purpose of the videoconferences is to discuss the ongoing projects, products, and new methodologies, and to provide special assistance on models and several other technical issues.

The Observatory has also provided a Grid Technology Infrastructure since 2010, known as AndesGrid [3] to help institutions with less computational power to increase their capabilities using other institutions' available computer resources. The initiative enables the partners to share not only model outputs and, but also their local observations (e.g. rain gauge and temperatures) in "real time". For additional details, see [3].

Of the OLE²'s main achievements in this period, the most important is the enhancement of the institutions' human resources by means of continuous technical support for forecast, modeling, verification, risk assessment and management and many other issues, all of which is done in their native language. The project has also succeeded in standardizing forecasts, data formats, and methodologies, providing common models, tools, and procedures that are used on a daily basis in all the involved institutions. A key tool has been the wiki, which contains all of the needed steps for each task, which by itself is another of the most important elements for ensuring the long-term continuity of the OLE².

Figure 1. The Observatorio Latinoamericano's partners (as January 2012) are mainly national weather services, universities and research and development institutes. (The red line links the countries participating in the AndesGrid iniciative [3]. See main text).

Among the different efforts, risk assessment and management are some of the most important tasks targeted by the Latin American Observatory partners. Even when a general framework is used by the OLE² on this field, the different institutions develop tailored applications to take advantage of the local experience and mechanisms already in existence in their countries. However, the general framework closely follows the ideas discussed by Mora [4].

Risk management is from a wide perspective, a policy system that, among its components, permits stakeholders in several levels, to identify, analyze and quantify the probability of human, material and ecosystem losses, as well as to provide metrics and technical criteria for prevention, mitigation, reduction or transfer of risk [4]. A key idea to be still implemented in most parts of the world is that risk management needs to be understood and addressed as an investment, not as a cost. [4] It is clear that nowadays the establishment of efficient risk reduction plans does not possess a real priority on the agenda of decision-makers and stakeholders, except *a posteriori*. This is especially true in most parts of Latin America, as is exposed in some case studies included in this chapter. Moreover, in general, there are no consolidated cultures incorporating risk management throughout decision-making processes for public and private investment and planning (see [4] and references therein). Also, private sector involvement in this process presents several unique challenges to be overcome [5].

Since 1975 in Latin American and the Caribbean (LAC), the costs of losses attributable to major natural hazards have been estimated at around US$300 billion, with more than 280,000 human deaths and affecting more than 160 million people [6-8]. Table 1 references the major natural disasters in LAC in the last three decades.

Year	Hazard/Event	Countries
1983	Floods	Argentina, Bolivia, Brazil, Peru
1983	Earthquake	Chile, Colombia
1985	Earthquake	Mexico
1998	Floods, landslides associated to hurricanes George and Mitch	Central America and the Caribbean
1999	Floods and landslides	Venezuela (Vargas State)
2005	Most active hurricane season recorded in history [8]	Central America and the Caribbean
2009	Drought (strong El Niño event)	Venezuela, Colombia and Ecuador
2010	Earthquake	Chile
2010, 2011	Floods (strong La Niña event)	Panama, Colombia and Venezuela

Table 1. Major natural disasters in LAC between 1983 and 2011 [6-8].

Consequently, it has been an urgent need for the Latin American Observatory's partners to address the establishment of effective policies for risk assessment and management in their respective countries. To this end, an implementation plan was suggested by CMC and this chapter presents briefly the adopted definitions and general ideas (sections 2-4), as well as case studies for different managerial sectors in Latin America (section 5) in order to discuss in some detail, present risk assessment/management policies and methodologies. Finally, we discuss the benefits of sharing experiences among different countries to address common problems in geographically complex regions, as well as how to achieve success in countries where the scarcity of economic and trained human resources imposes serious limitations on the effectiveness of risk management systems.

2. Risk definition and management

In this and the next section the main definitions related to risk management, hazards and vulnerabilities are discussed in order to provide the foundations chosen by the Observatory in these matters. We don't pretend to address all the details in this complex field, but to outline the main ideas for future reference.

Different agents play a role in the occurrence of natural hazards, which derive from the damaging potential or a combination of them. These agents are frequently [4,6] classified in terms of their origin, namely

- hydro-meteorology: both global and local processes (hurricanes, floods/droughts, weather extreme events)
- internal geodynamics (seismicity and volcanism)
- external geodynamics (landslides, intensive erosion, torrential debris flows).

Following Mora [4], natural hazards (*H*) can be formally defined in terms of the probability that an event becomes so intense (*a*) within time and space frames that it produces significant damage (*d*). The intensity and damage definitions depend on the region and timescale under consideration. Vulnerability (*V*) can be associated with the probability that, according to the intensity of the natural event, damage might occur as a function of the degrees of exposure and fragility of the elements involved. Risk (*R*) therefore is the combined probability (convolution, *) that a hazard might cause significant damage, according to the following relationship:

$$\int_{h,d} p(R)da = \int_{h} p(H)da * \int_{d} p(V)da \qquad (1)$$

The probabilities are not, in general, constant, so the spatial and temporal variability must always be carefully considered. Moreover, *V* and *H* are not independent.

Risk management involves not only the quantification of the probability of loss via equation (1) and the derived secondary effects, but also the generation and execution of policies for [4]

- identification, analysis and quantification of risk
- formulation and application of measures of prevention, reduction and mitigation of causes
- financial, social and environmental protection
- execution of related protocols (e.g. preparedness, response, rehabilitation, reconstruction).

Efforts to foster these processes may include the establishment of adequate policies and protocols to incorporate stakeholders into risk management, improvement of the climate/geodynamic information available, identification of the different sources of vulnerability and their evolution in time, definition of “acceptable” risk levels and the establishment of early warning and monitor systems, and the adequate ways to communicate their products to the end-mile users.

Risk management policies foster and integrate strategically balanced processes capable of providing decision-making tools in a synergistic, decentralized and participative way. In the OLE[2] this is a central philosophy, which also enables the treatment of risk in trans-boundary zones (e.g. the Observatory enabled the continuous exchange of climate services between the National Weather Services and decision makers during the 2010 floods that affected Panama, Colombia and Venezuela in similar ways).

An important aspect of risk management is related to operative platforms that enable decision makers and users to interact and to obtain information useful to assess vulnerabilities, the impact of natural hazards and the risk probabilities for a certain region. This infrastructure is aimed at providing this kind of information *ex ante*, but it can also be used during or after an emergency. It usually includes

- macro- and micro-zoning of hazards and vulnerability and its probable evolution in time (see section 4)
- damage assessment
- identification of priority areas of intervention
- protocols for vulnerability reduction (hazards are not, in general, reducible)
- instruments for financial strategy and budgetary allocations
- identification of "risk bearers" and their responsibilities
- administrative and legal information (e.g. design, operation, lease, transfer and reclamation contracts)
- information on past events
- areas of further investments to reduce future disasters.

The implementation of these platforms in LAC is more advanced in a few countries (see section 5.1 for a Central American example). The Observatory's partners are taking advantage of the experience developed by other members of the regional initiative to improve their own risk management infrastructure.

Finally, but no less importantly, the Observatory is now committed to following the International Research Institute on Climate and Society's (IRI) Four Pillars for Climate Risk Management [31], namely

1. **Identify vulnerabilities and potential opportunities** due to climate variability/change for a given water, agriculture, or health system.
2. **Quantify uncertainties in "climate information"** in order to reduce uncertainties in using that information.
3. **Identify technologies and practices** that optimize results in normal or favorable years as well as technologies and practices that reduce vulnerabilities.
4. **Identify interventions, institutional arrangements and best practices** that reduce exposure to climate vulnerabilities and enable the opportunistic exploitation of favorable climate conditions.

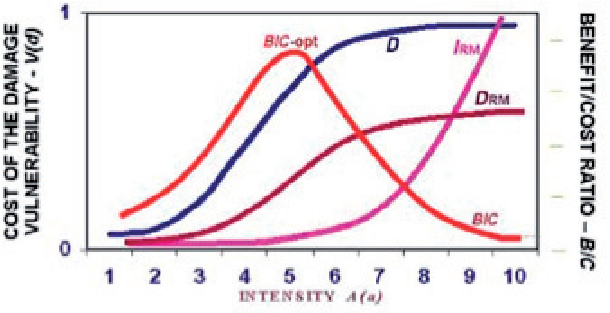


Figure 2. Benefit/cost (B/C) ratio defined as a function of damage expected without risk management (D), reduced by applying risk management (D_{RM}) and the investment involved (I_{RM}). (After Mora [4])

3. Benefits of risk management

It is possible to use risk management as a tool for providing an idea of the economic value of potential damages caused by hazards and the actions undertaken to reduce vulnerabilities. Figure 2 sketches an accumulating damage curve *D* without risk management. If risk management policies are used, this same curve is reduced to a more optimistic distribution, D_{RM}. The reduction of damage due to risk management is then

$$RD = D - D_{RM} \tag{2}$$

The associated investment (I_{RM}) for such a reduction depends on both the hazard and vulnerability involved (which, again, are not independent). On the other hand, the Benefit/Cost (*B/C*) ratio allows for the finding of optimal levels of investment. It's possible to show [4] that the net benefit of risk management is

$$B_{RM} = (D - D_{RM}) - I_{RM} \tag{3}$$

and the Benefit/Cost ratio can be defined as

$$B / C = \frac{D - D_{RM}}{I_{RM}} - 1 \tag{4}$$

Having real distributions of the different quantities it might be possible to select optimal points (e.g. the saddle point named B/C_{OPT} in Figure 2) to limit profitable investments. See [4] and [6] for details.

4. A Multi-scale approach for risk management

The traditional approach to computing risk probabilities using equation (1) involves the consideration of static or, in the best-case scenario, long-term mean vulnerability maps of the associated hazards. Nonetheless, the probability density function associated with hazards in equation (1) is frequently computed from time series that may contain useful information at different time scales. Given the fact that a decision maker might be interested in a certain time scale (e.g. on the order of its own managerial period), a temporal decomposition of the original time series used to identify the hazard probability may be extremely helpful.

For example, Figure 3 presents such time-scale decomposition for Southeast South America's precipitation. The lower panels sketch the non-linear long-term trend (a proxy for climate change), and decadal and interannual variability time series [9]. The rationale is that in considering the interaction of the three climate signals (the explained variance being a way to ponder the specific weight of each one of them) it is possible to provide better climate services to decision-makers, that are more adequate for their scale of interest, e.g. next 10 years.

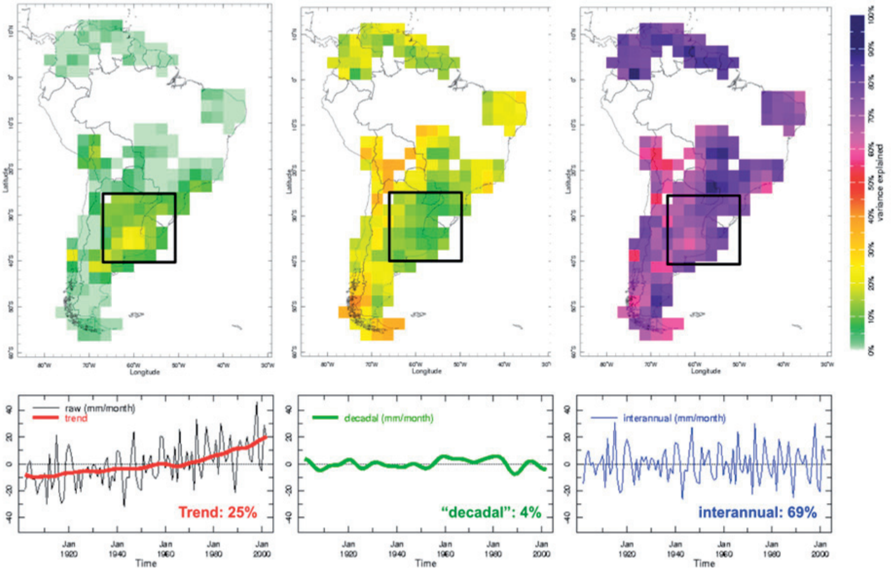


Figure 3. Decomposition for Southeast South America's precipitation (boxed region) for three different time scales: long-term trend (left panel), decadal variability (central panel) and interannual variability (right panel), with their respective explained variance (colors in the map and values in the bottom corner of each time series). Grid boxes in white are not statistically significant at a p-level=0.95. (After Goddard *et al.* [9])

As mentioned before, hazards and vulnerabilities are not independent. The latter encompasses exposure to hazards, sensitivity to these hazards and adaptive capacity. All of them, in fact, evolve in time. Therefore, a similar multi-scale time decomposition must also be done to correctly assess the related vulnerabilities. This provides a more realistic risk probability distribution, and thus better assessments and policies can be developed.

5. Cases of study

In this section a few state-of-the-art examples of risk assessment and management in several countries in Latin America are briefly described. Signatory and participant institutions of the Latin American Observatory operate over these structures, providing tailored methodologies based on what has been explained in previous sections.

5.1. Seasonal Climate outlooks applications for food security decision-making in Central America

In Central America, seasonal climate outlooks turn into risk scenarios used by food-related sectors to help support their decisions and prevent food insecurity. This is a coordinated effort with specialized entities of the Central American Integration System.

The Central American Climate Outlook Forum (CA-COF), coordinated by the Comité Regional de Recursos Hidráulicos (CRRH-SICA)[5] (Regional Water Resources Committee), has consolidated a process to issue three seasonal outlooks per year, bringing together the capacities of all seven weather services in the region. The CA-COF takes advantage of international and global sources of information, which it analyzes along with its own sources and historical data to produce Climate Outlooks for the Central American region. Over the last decade the Forum has issued 38 regional climate outlooks.

To facilitate the use of climate risk information for decision-making with the aim of reducing food insecurity risks, CRRH-SICA has fostered a mechanism over the past few years that turns the Outlook into risk scenarios for those sectors related to nutrition and food security, particularly agriculture, fisheries, potable water and public health, as well as cross-cutting areas like risk management and emergency response. These scenarios are used when deciding preventive measures to mitigate the impact of climate variability on food security.

In a joint effort between the specialized agencies of the Central American Integration System (SICA) and the Regional Food Security Program (PRESANCA II) funded by the European Union, immediately after an Outlook is issued, a group of regional experts is convened to enhance the Outlook with information from the different sectors mentioned above, turning the product into sectoral climate risk scenarios that could guide early warnings of actual and potential hazards to food security.

The working group, composed of CA-COF members and the experts, uses these scenarios to identify sector-specific preventive measures. The food security risk scenarios and the suggested measures are circulated among government entities and other organizations involved in addressing food and nutritional security through their own networks to ensure all this information reaches the most appropriate decision makers.

---

[5] CRRH is the technical Secretariat of the Central America Integration System, responsible for the coordination of activities related to weather forecast, climate, water resources and climate change assessments with Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua and Panamá.

This same method is used at the national level to produce specific recommendations for diverse sectors in each individual country. Those are then distributed among national authorities and stakeholders in private sector with the seasonal outlook map and lists of suggested prevention measures. An example of the matrix of climate risk for agriculture is presented in Table 2.

As part of the same, routine-specific measures are identified at a national level by similar working groups of climatologists and national experts, and suggested to authorities.

A process is now underway to evaluate the use of outlooks and climate risk information directly with the beneficiaries, both for the private and the public sectors, to strengthen decision-making efforts.

Climate Risk /potential damages for Main Crops. Quarter December 2011 -March 2012							
Crop				Country			
	Belice	Costa Rica	El Salvador	Guatemala	Honduras	Nicaragua	Panama
Maize	Plantations in coastal areas likely to be affected by above-normal preci-pitation	No risk	No risk	Risk of post-harvest damages due to higher humidity associated with above-normal rainfall. Increased cost of post-crop product manage-ment	Risk in the Cari-bbean planting area due to above-normal rainfall. Second planting season likely to be affected	Second crop in the Atlantic Autonomous Regions likely to be affected by above-normal rainfall. Exportations likely to be impacted	Crop loss risk in Western Cari-bbean area. Posible impacts in food security of indi-genous population.
Beans	Plantations in coastal area likely to be affected by above-normal precipitation	Risk of crop damage for plantations in Northern Plains, Cari-bbean area and South Pacific	No risk	No risk	Risk in the Cari-bbean planting area due to above-normal rainfall. Second planting season likely to be affected	Second crop in the Atlantic Autonomous Regions likely to be affected by above-normal rainfall. Yield reduction may reach 15% .	Loss of crop risk in Western Caribbean area. Posible impacts on food security of indigenous popu-lation.

Table 2. Central American Climate Outlook Forum risk assessment for two crops (maize and beans), for December 2011-February 2012.

5.2. A flood early warning system for Vargas, Venezuela

In Venezuela, heavy rains represent a significant problem for human populations, not only in rural areas where farms and crops are affected, but also in some urban settings where many inhabitants reside in poorly-constructed houses that are highly vulnerable to floods and landslides, or that are sometimes located on steep terrains and floodplains.

Figure 4. Aerial view of Caraballeda, Vargas state (Venezuela), after mudslides and debris flows event on December 15 1999.

On December 15, 1999, on the northern coast of Venezuela, torrential rains led to flash floods and debris flows that killed tens of thousands of people, destroyed thousands of households, and meant the complete collapse of the area's infrastructure (see Figure 4). The "Vargas disaster", as it has been known ever since,, is considered the worst natural disaster in Venezuela's last half century (Table 1). Even though flood prediction is an essential piece of Climate Risk Management, Venezuela did not have, at the time, a consolidated early warning system that could alert decision-makers and stakeholders about this extreme event. In this section, the implementation of an early warning flood system for Vargas state is described in some detail. It is a completely general methodology, so other regions in LAC can benefit from the experience described here.

The Vargas disaster is also an important case study on flood prediction, not only because of its unusual rainfall amounts, but also for the nature of the terrain where it took place. The basin is located in a mountainous region of metamorphic rock, coarse soils and steep slopes, making the area highly vulnerable to floods and debris flows [10].

The early warning system involves the estimation of vulnerabilities and probabilities related to heavy precipitations and mudslides. Guenni *et al.* [11] has computed the vulnerabilities using the total affected and exposed people, considering both the spatial and temporal variability, as discussed in section 4. As an example, Figure 5 presents a zoom-in for most parts of Venezuela showing months of rainfall exposure in colors and the population density as black pixels (for details see [11]). Regions near the Venezuelan coast are in general more vulnerable (due to more population density and exposure). Using map algebra in a Geographical Information System (GIS), the map is then intersected with other maps containing information about the terrain (e.g. water holding capacity or mudslide probabilities). Following equation (1), the final map, in sketching homogeneous vulnerability zones, must be written in terms of probabilities for different hazard intensities.

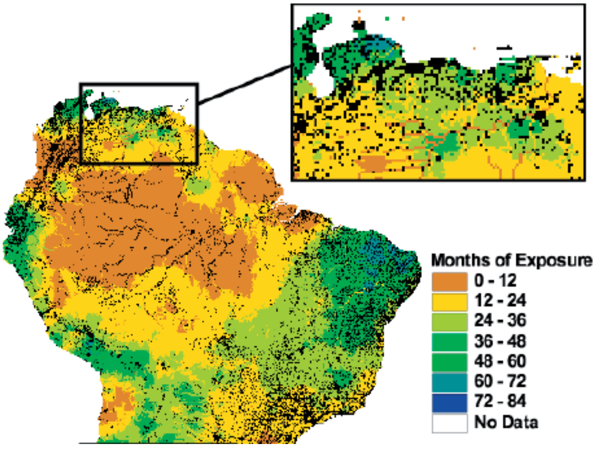


Figure 5. Map of exposition (in colors) and population density (pixels in black) for Venezuela. After Guenni *et al.* [11]

On the other hand, in order to compute flood risk probabilities, hazard probability maps must be produced (also in terms of intensities). In this case, Torres and Muñoz [12,13] have suggested a methodology based on hindcasts involving an off-line coupling of a regional climate model, a process-based hydrological model and a routing model.

The Climate Weather Research and Forecasting model (CWRF) was used to simulate 13 years (1996-2008) of data, paying special attention to the Vargas disaster atmospheric conditions. Rainfall and maximum and minimum temperatures were bias corrected using observed values. Then they were used as input data along with soil properties and vegetation parameters for the Variable Infiltration Capacity model (VIC). This energy and mass balance model was finally coupled to the Lohmann routing model to produce realistic stream flows that adequately considered the local topography (for details about the different models and procedures see [2,12,13] and references therein). Simulation outputs were consistent with the flooding event of December 1999. The study was able to reproduce how long rains, along with low evapo-transpiration due to high cloudiness, contributed to saturate soil layers. Moreover, the study was able to identify critical values (see Figure 6) in order to establish “yellow” and “red” alerts in the Vargas flood early warning system, using the statistics obtained for the 13 years simulation and local stream flow data provided by the National Weather Service. Model outputs were then used to create hazard probability maps in terms of hazard intensity.

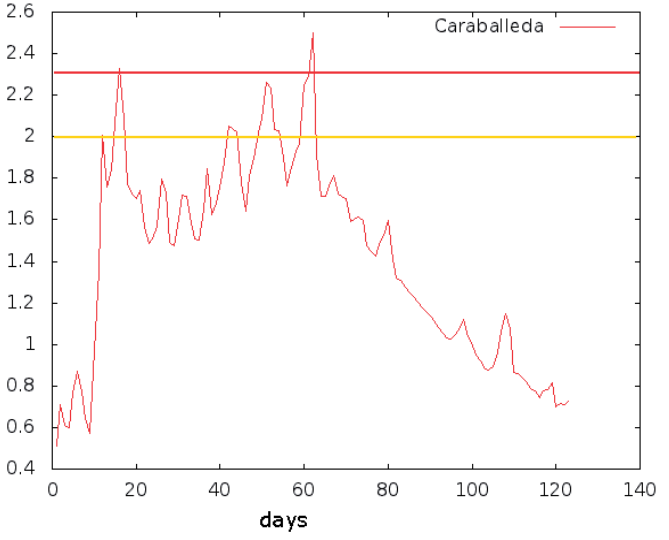


Figure 6. Caraballeda's simulated runoff time series for the Vargas disaster event using CWRF, VIC and the Lohman routing model (see main text). The runoff has been normalized using the standard deviation. The critical day (December 19, 1999) corresponds to simulation day 62 in this plot. The yellow and red lines at 2σ and 2.3σ, respectively, were defined as the "yellow" and "red" alerts in the designed early warning system. (After Torres and Muñoz [12,13])

As a final result, series of maps are provided with the hazard's probability of occurrence in terms of the hazard intensity. They are produced computing the convolution between the corresponding hazard and vulnerability maps, following equation (1). For additional details see [12,13].

This probabilistic approach has an important advantage: it takes into consideration the possible uncertainties related to each one of the processes involved in the final flood risk probability map.

5.3. Bridging the gap between climate information and health services in Colombia and Ecuador

Climatic factors play a significant role in the transmission dynamics of several infectious diseases [14]. Therefore it should be a critical priority to incorporate climate information into disease risk assessments and early warning and response systems. (See figure 7.) This, in fact, has been one of the goals of the Colombian and Ecuadorian meteorological and health services in recent years. In Colombia, in particular, the National Institute of Health (INS) recently teamed up with the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) and several universities and research centers to design and implement a proactive, collaborative, multidisciplinary, Integrated Surveillance and Control System –

ISCS [15,16]. This initiative is part of a set of measures and policy options that the Colombian adaptation strategy to climate change proposed for three areas of primary concern: the high-altitude Andean ecosystems, the insular and coastal areas, and human health (see Integrated National Adaptation Pilot [16]).

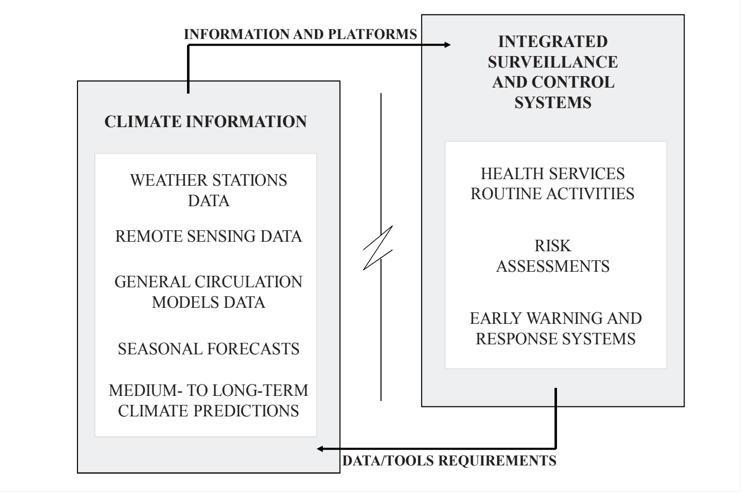


Figure 7. Climate information and human health integrated surveillance and control systems.

The aim of the adaptation strategy for the health sector is to have a better-prepared institutional arrangement for increased exposure to malaria and dengue fever, two climate-sensitive, vector-borne diseases that are still considered human health burdens in Colombia [17]. According to the scientific literature, the potential increase in the incidence of these two diseases is likely to occur not only in the already endemic malaria and dengue fever prone areas, but also on the fringes with the Andean regions where local communities have not been exposed to pathogens before. Thus they lack the immunity against these microorganisms. The approach has been to assist the health sector to better cope with current climate variability and climate-related events, as a means to make it better prepared against future climatic conditions likely to be brought by the ongoing long-term global climate change. As a result, Colombia has been working on reducing people's vulnerabilities to the negative impacts of malaria and dengue outbreaks, as well as developing an Early Warning System framework, supported by seasonal forecasting capabilities, weather and environmental monitoring, and statistical and dynamic models [18-23].

This effort required linking what had previously been two largely separate analytical domains: the field of public health, traditionally dominated by human health experts and practitioners; and weather and climate science, usually lead by meteorologists, climatologists, engineers, and environmental science experts. Several biologists, entomologists and social experts have also joined this collaborative, inter-institutional and interdisciplinary effort. The implementation of the ISCS has required, among many other activities, analyzing the local eco-epidemiological settings of various malaria- and dengue fever-prone pilot sites, implementing process-based biological and statistical models, and strengthening the local capacity of health authorities. It has also required strengthening the IDEAM capability to routinely and systemically produce and disseminate, relevant, continuous, homogenous, and reliable climate information that could support the decision-making process of health authorities. Such information includes ground-truth historical records, modeling simulation outputs, seasonal forecasts, El Niño Southern Oscillation forecasts, and climate change predictions. All this information, along with disease morbidity profiles, entomological conditions, socio-economic drivers, and impacts of interventions and control campaigns, is now steadily becoming a core part of routine activities and disease control plans of health services at regional, municipal and local levels, and is starting to facilitate a better allocation of health resources and more cost-effective preventive responses. On a broader scale, measures proposed as part of the adaptation plan have been mainstreamed into the Colombian political agenda to ensure their sustainability. They have reached, for instance, the 2010-2014 National Development Plan, the 2010-2014 Environmental Action Plan, the Colombia's Poverty Reduction Plan, and the Public Health National Plan.

In Ecuador, in turn, the National Institute of Meteorology and Hydrology (INAMHI) has conducted various research activities [24,25] on the analysis of potential increases in the incidence of malaria and dengue fever in key lowland provinces. Activities have included the implementation and coupled analysis of the 30 km WRF (Weather Research and Forecasting), dynamic downscaling regional climate model, and the Ross-Macdonald malaria infectious disease process-based model, to reproduce the spatio-temporal variability of primary positive cases reported by the National Malaria Eradication Service over a recent 13-year historical period (see figure 8). They have also included a first design of an Early Warning System for malaria and dengue fever outbreaks, which is now developing into an integrated surveillance and climate modeling for malaria and dengue fever predictability in rural and urban settings. The INAMHI has also joined efforts with several research groups to broaden the understanding of the complex transmission dynamics (in both space and time) of these vector-borne diseases, in order to improve decision-making processes in regional and local health authorities and, in a more general sense, to strengthen the already solid surveillance and control activities of infectious diseases conducted by the Ecuadorian health sector.

5.4. Air quality risk assessment for Lima and El Callao (Peru)

Poor air quality conditions can cause many respiratory problems including asthma and severe allergies and can also pose serious health problems such as cancer. In order to issue public warnings of high pollution episodes, the Peruvian National Service of Meteorology and Hydrology – SENAMHI is working on the design and implementation of an air quality forecast system for the cities of Lima and El Callao. The system maps out current and future ambient air quality conditions based on hindcast simulation runs of the BRAMS model and the Weather Research and Forecasting coupled with Chemistry (WRF-Chem) model (see figure 9), as well as on *in situ* measurements of main air pollutants, such as particulate

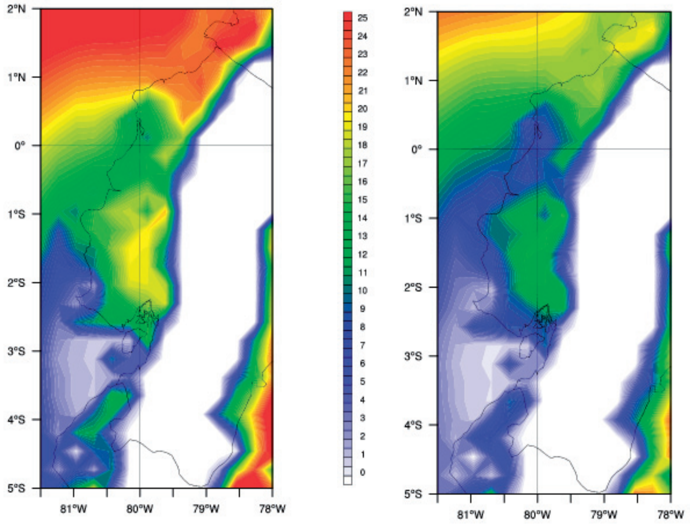


Figure 8. January *Plasmodium vivax* (left panel) and *P. falciparum* (right panel) basic reproductive rates on the Ecuadorian coast, simulated for the period 1996-2008 and for *Anopheles albimanus* mosquito species. (After Muñoz and Recalde [24]).

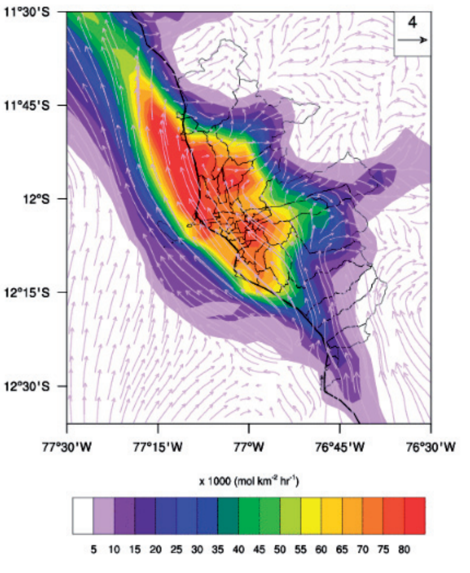


Figure 9. 4-km spatial resolution hindcast WRF-Chem model simulation outputs of NOx concentration fluxes in the geographic domain 12°30'S – 11°30'S and 76°30'W – 77°30'W. Typical NOx concentration fluxes are expressed in thousand mol/km²/hr. The reference arrow represents wind speeds of 4 m/s.

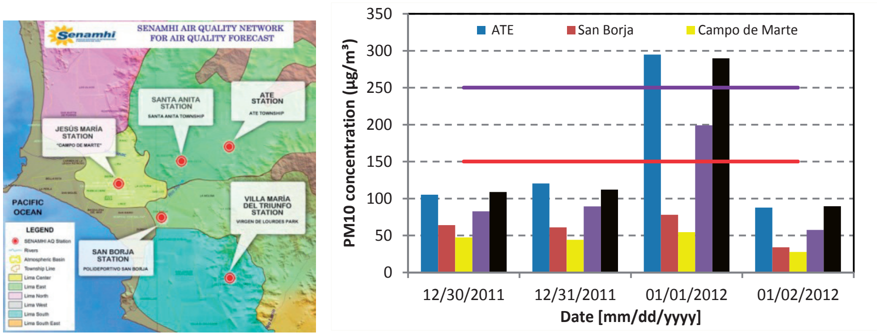


Figure 10. 10 micrometer or less particulate matter (PM10) concentrations (right panel) gathered at the SENAMHI air-quality monitoring stations Ate, San Borja, Campo de Marte, Santa Anita, and Villa Maria del Triunfo (see locations in left panel) over the period spanning from December 30, 2011 through January 2, 2012. The red solid line depicts the health-based daily air quality standard concentration of 150 µg/m³. The violet line depicts, in turn, the 250 µg/m³ daily concentration threshold above which high pollution episode warnings are issued. Air quality is considered 'good' when daily concentrations do not exceed 50 µg/m³.

matter smaller than 10 and 5 micrometers (PM10 and PM5, respectively), nitrogen oxides (NOx) and ground-level ozone (O_3). Modeling outputs and real-time information allow the SENAMHI to issue public warnings that could activate municipal plans to help keep pollution levels down and alert local residents in Lima districts to the potential health threats.

Up to date, the air quality network includes five high-precision monitoring stations located in the surroundings of the aforementioned two densely populated cities. (See figure below.) High pollution episode warnings include events such as the one that took place in the localities of Ate and Villa María del Triunfo on January 1, 2012. *In situ* air quality measurements (figure 10) suggested that PM10 concentrations reached 295 and 290 µg/m³, respectively, which exceeded in 45 and 40 µg/m³ the 250 µg/m³ daily concentration threshold. In the monitoring station Santa Anita, in turn, the PM10 concentration reached 199 µg/m³, also exceeding the health-based daily air quality standard concentration of 150 µg/m³. The air quality stations San Borja and Campo de Marte reported moderate daily concentrations of about 78 and 54 µg/m³, respectively.

5.5. The agrometeorological bulletin and improved decision-making processes in Paraguay

Extreme weather events such as river floods, severe storms, droughts and below-freezing low temperatures are strongly linked to the onset of El Niño Southern Oscillation (ENSO). Extreme events primarily affect river streamflow and cause numerous direct and indirect impacts on many key sectors of the Paraguayan economy: agriculture, ground and river transportation, potable water, construction, electricity, and recreation. Several studies have demonstrated that the potential impacts of these extreme events and their economic consequences are directly related to the magnitude of ENSO episodes. Historically, Paraguay has experienced above normal mean temperatures and rainfall amounts, as well as unusually heavy rainfall events, during El Niño (or ENSO positive phase) episodes. The resulting floods have affected thousands of individuals, damaged numerous houses, public buildings and highways, and submerged entire crops, cattle grasslands and livestock farms. Recent examples include the catastrophic events that took place in the Paraguay and Paraná rivers, particularly during the El Niño 1982-83, 1991-92 and 1997-98 years. In the flat Paraguayan savannahs, rainfall extreme events have also caused a proliferation of *Aedes aegypti* mosquito breeding sites, increasing the incidence of classic dengue fever. This has been particularly true in the central, northeastern and eastern portions of this country, where climatic conditions are suitable for the successful development of *Aedes* mosquitoes. During La Niña events (or ENSO negative phase), the country usually experiences the opposite (i.e. rainfall deficits and well below normal ambient temperatures). The concomitant droughts cause a decline in dairy production, an increase in the rate of desertification of arable land, and a rise in the occurrence of grassland and forest fires. They also decrease hydropower generation, increase the pollution of rivers and pools of stagnant water, and limit fluvial transportation, thereby increasing import/export transport fees and diminishing the trade of goods. Moreover, long dry spells affect sunflower, maize, soy, cotton, and wheat production, thus reducing the revenue from these key agricultural activities. All these impacts, although mainly the ones affecting the agriculture sector, prompted the creation of a multidisciplinary group led by the Paraguayan Meteorological and Hydrological Service, and the Risk Management Unit at the Ministry of Agriculture and Livestock Farming. Collaboratively, the group issues an agro-meteorological bulletin, every month or at other intervals depending on specific needs, following the approach presented in the decision-making information system below. Activities include (see Figure 11) the

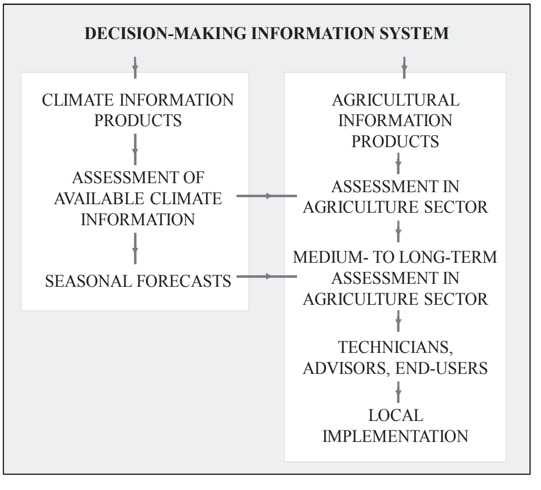


Figure 11. The Paraguayan monthly agrometeorological bulletin (available online at: http://www.meteorologia.gov.py/) and the decision-making information system.

assessment of available climate information, seasonal and ENSO forecasts, and medium- to long-term potential impacts on local agricultural production. Key information is then shared with local farmers and smallholders through technicians, advisors and on-the-ground experts.

Up to date, the total number of end-users has reached over 2,000 local, regional and international groups including, among many others, the South American Common Market (MERCOSUR) and organizations in the United States of America, Canada and European countries. Numerous technicians, decision-makers, researchers, and students also use this sectoral information in their individual routine activities.

5.6. Climate risk of droughts in Chile and its effect on agriculture

In central Chile, the *secano* zone involves a total surface of 4.362 km^2, and a human population of 54.450 habitants (6.25% of the regional population). Considered as a rural area, the *modus vivendi* depends basically on agriculture. In the coastal secano of the O´Higgins Region, the total surface designated to crops and plantations is on the order of 22.800 hectares. The majority of this land possesses forage plants (27%), cereals –especially wheat- (21%), fruits (20%), vines and vineyards (25%), and in lesser proportion legumes (5%) and vegetables (2%) [26]. Presently, the cattle production is on the order of 536.170 animals, most of them being sheep.

The secano zone shows a dry season that varies between 6 and 8 months per year. Precipitations take place between April and September, with values around the 500-600 mm/year [27]. Reports on droughts in Chile go back to the times before the Spanish Colonies, with a strong impact in the agriculture production throughout the history of the country [28]. In a period of 400 years, 25% has been reported as dry and half of them as extremely dry [29]. The mean probability is therefore on the order of one dry year every 4 years. Quintana and Aceituno [30] have recently discussed trends of decrease in precipitation for central Chile of 10-30% for the second half of the XX century.

Traditionally, the drought management in Chile has been contingent. Nonetheless, in the first decade of the XXI century, government actions in this regard have increased, orienting efforts towards an integrated management of extreme climate events. For example, this originated the National Plan for Civil Protection in 2002. The formal approval of the National Climate Change Strategy in 2006 and the respective Action Plan in 2008 have revealed the adaptation needs of the different productive sectors of Chile, but also the most vulnerable territories and populations. All this created a positive scenario for the establishment of improved policies for climate risk. Three important examples after the extreme drought of 2008 are the creation of the Comité Interministerial de Recursos Hídricos (Inter-ministry Water Resources Committee), the Comisión Nacional de Emergencias Agrícolas y Riesgos Agroclimáticos (National Committee for Agriculture Emergencies and Agro-Climate Risk) and the Sistema Nacional de Gestión del Riesgo Agroclimático (the National System for Agro-Climate Risk Management, found at http://www.minagri.gov.cl/agroclimatico/comision_nacional.php).

The public risk management initiatives implemented by the Chilean organizations involve one or several of the following mechanisms. While excellent initiatives, some of them can still be further improved. Some suggestions are pointed out in those cases.

Early Warning System

Chile possesses a modern early warning system that is able to detect and send alerts about natural disasters. It uses information produced by the Dirección Meteorológica de Chile (DMC), Dirección de Obras Hidráulicas (DOH), Dirección General de Aguas (DGA), Servicio Hidrográfico y Oceanográfico de la Armada (SHOA) and others. The Ministry of Agriculture via its Regional Centers for Agro-meteorological Information interacts with the climate information providers and generates tailored products for the corresponding sector. Nonetheless, these products do not always arrive on time to the farmers.

Agriculture Emergency Announcements

The Agriculture Emergency Zone (ZEA, in Spanish) announcements for droughts are fast and the corresponding institutional instances are efficient. However, the criteria used for announcing the ZEAs, which is based on the agriculture impact, may require modifications. The impacts are assessed in terms of a combination of factors: e.g. losses reported by a technical organism, request of a local authority and help requests from affected farmers. Specifically, ZEA announcements may require a criteria homogenization for both monitoring the evolution of the drought, as well as the emergency declaration itself. The announcement must adjust to regional realities and take into consideration the potable water supply and its effects on agriculture activities.

Budget Availability for Emergencies

Even when the budget allocation for emergencies is normally made through the utilization of funds already directed to other activities, in the last years, there has been an increasing amount of money in Chile properly directed to environmental emergencies. For instance, the emergency drought budget made available for the 2007-2008 case enabled the development of a higher number of implemented policies. In the Ministry of Agriculture, different mechanisms (e.g. emergency bonds, accident benefits, livestock health operations, agriculture emergency employment programs and a special incentive fund program for the recovery of degraded soils) were able to take care of around 161,000 beneficiaries.

A final comment can be included here about the implementation of an additional system to monitor the success of the risk management process in Chile. Presently there is neither an evaluation of the loss levels associated with a drought, nor an *ex post* diagnostic of the implementation mechanisms. It is necessary to possess adequate metrics to assess the effectiveness of applied strategies and the recovery of the affected populations. It is important to develop and to implement a protocol for the evaluation of the impacts of droughts. In addition, this will help with emergency management policies and the development of *ad hoc* prevention and mitigation measures focused on the vulnerability reduction of certain territories. A continuously updated registration of the users and affected people during an emergency is also required; this information will help to project estimates of the offer and demand under similar circumstances. The associated database will consolidate the emergency monitoring system and also will provide useful information for future studies and experience-sharing with other institutions.

6. Discussion and concluding remarks

The Latin American Observatory partnership has been able to improve several aspects of the way decision-making tools are created and provided to stakeholders and end-mile users in different institution of Latin America in the past years, increasing the dissemination of available services for assessing and –when possible- forecasting hazards and vulnerabilities. This allows advisement of authorities in taking the right course of action in order to protect human populations from the direct or indirect effects of hazards.

In this chapter, we have presented the general methodologies that are used in the OLE^2 for risk assessment and management, along with a few case studies for different countries. The main innovation is related to the way the partnership interacts to share data, products, experiences, helping institutions with lesser human resources and capacities to take advantage of the strengths of other partners, but also sharing efforts at a regional level so the work is divided among a greater number of experts, even when they don't belong to a certain institute, or even when they are not present in the same country. This interdependence has proven a key component of the long-term sustainability of the Observatory.

In terms of the methods employed, important learned lessons arise in relation to the use of probability maps when working with vulnerabilities, hazards and risk management. A probability approach involves the consideration of the related uncertainties of every step of the methodology, which are always present in this kind of analysis. This introduces an *a priori* spectrum of possible immediate actions to decision-makers, even if the real intensity of the hazard is not known until the moment of the emergency. The whole system is aimed at improving the response times and also benefit/cost ratios.

Providing services that take into consideration different scales (e.g. interannual to decadal to climate change scales, as well as from basins to nation-wide to continental scales) and the evolution in time of both hazard and vulnerability probabilities has proven in our experience to be a more adequate practice when assessing and managing risk.

The recognition of the local chains of dissemination of the information, institutional relationships and their adequate use and promotion instead of creating new ones is another lesson learned. The Observatory as a "boundary institution" is paying special attention to these issues and also to the importance of "translating" methods and products so the final services can not only be adequately understood but also be properly assimilated by decision-makers and stakeholders.

Finally, there is an increasingly growing culture on risk assessment and management in LAC, but as Baethgen has pointed out recently [31], it is also important to consider the benefits that may arise if there is also preparedness to take advantage of probable "positive" natural events (e.g. a natural event may be a hazard for certain sectors of production, but may be positive for others). The Observatory must therefore consider also Opportunity Assessment and Management policies in the near future to address such profitable scenarios.

Author details

Á.G. Muñoz[1,2,3], D. Ruiz[1,3,4], P. Ramírez[1,5], G. León[1,6], J. Quintana[1,7], A. Bonilla[1,5], W. Torres[1,2], M. Pastén[1,8] and O. Sánchez[1,9]
[1]*Observatorio Latinoamericano de Eventos Extraordinarios (OLE[2])*
[2]*Centro de Modelado Científico (CMC). Universidad del Zulia, Venezuela*
[3]*International Research Institute for Climate and Society (IRI). Columbia University in the City of New York, United States of America*
[4]*Escuela de Ingeniería de Antioquia, Colombia*
[5]*Comité Regional de Recursos Hídricos (CRRH), Costa Rica*
[6]*Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), Colombia*
[7]*Dirección Meteorológica de Chile (DMC), Chile*
[8]*Dirección Nacional de Aeronáutica Civil (DINAC), Paraguay*
[9]*Servicio Nacional de Meteorología e Hidrología (SENAMHI), Perú*

7. References

[1] Muñoz, Á.G., Núñez, A., Cova, R. Climate System Simulations: An Integrated, Multi-Scale Approach for Research and Decision-Making. In Computational Simulations and Applications, Chapter 21. Zhu, J. (Ed.), ISBN 978-953- 307-430-6, InTech Publishing. pp. 449-468. 2011.

[2] Muñoz, Ángel G., and Coauthors. An Environmental Watch System for the Andean Countries: El Observatorio Andino. *Bull. Amer. Meteor. Soc.*, 2010. 91, 1645–1652.

[3] Muñoz, Á. G., X. Chourio, S. Reverol, A. Urdaneta, C. Díaz. ANDESGRID: A Grid Insfrastructure for Geosciences in the Andes. *Proc. Latin-American Conf. on High Performance Computing*, Gramado, Brazil, Bull, Hewlett-Packard, Intel Corporation, Microsoft, Silicon Graphics International, 2010. 9–14.

[4] Mora, S. Disasters are not natural: risk management, a tool for development. From: Culshaw, M. G., Reeves, H. J., Jefferson, I. & Spink, T. W. (eds) Engineering Geology for Tomorrow's Cities. Geological Society, London, Engineering Geology Special Publications, 2009. 22, 101–112.

[5] Vaughan, C. (Climate Service Partnership's Program Manager), personal communication.

[6] Mora, S., Keipi, K. Disaster risk management in development projects: models and checklists. *Bulletin of Engineering Geology and the Environment*, 2006, 65, 155-165.

[7] Cardona, O. Indicators of disaster risk and risk management in Latin America and the Caribbean. Instituto de Estudios Ambientales (IDEA). Department of Sustainable Development, Inter-American Development Bank, Washington, DC. 2006.

[8] Beven, John L.; Avila, Lixion A., Blake, Eric S., Brown, Daniel P., Franklin, James L., Knabb, Richard D., Pasch, Richard J., Rhome, Jamie R., Stewart, Stacy R. Atlantic Hurricane Season of 2005. *Monthly Weather Review*, 2008. 136 (3): 1109–1173.

[9] Goddard, L., Baethgen, W., Greene, A., Seager, R., Gonzalez, P., Muñoz, Á. G. , Ines, A., Tebaldi, C., Ruane, A. Multi-scale Climate Information for Agricultural Planning in Southeastern South America for Coming Decades. NSF-EaSM. 2012

[10] Wieczorek, G., Larsen, M., Eaton, L., Morgan, B., Blair, J. Debris-flow and flooding hazards associated with the December 1999 storm in coastal Venezuela and strategies for mitigation. U. S. Geological Survey. Open File Report 01-0144. 2006.

[11] Guenni, L., Hernandez, A. and Fillipone, M. Modelling population vulnerability and risk to extreme rainfall events in Venezuela. *Acta Científica Venezolana*, Vol. 54, Sup. 1: 2-12. 2003.

[12] Torres, W. and Muñoz, Á.G. Simulación Computacional con Modelos Dinámicos del Evento Extremo de Vargas (Venezuela), 1999. BSc Thesis. Universidad del Zulia. 2011. 53 pp.

[13] Torres, W. and Muñoz, Á.G., Early Warning and Flood Risk Assessment for Vargas state, Venezuela. In preparation.

[14] McMichael, A.J., D.H. Campbell-Lendrum, C.F. Corvalan, K.L. Ebi, A. Githeko, J.D. Scheraga, and A. Woodward. Climate change and human health: Risks and Responses. World Health Organization, Geneva. 322 pages. 2003.

[15] Ruiz, D., S. Connor, and M. Thomson.A multimodel framework in support of malaria surveillance and control. Chapter VII, pages 101-125 in: Seasonal Forecasts, Climatic Change, and Human Health – Health and Climate / Advances in Global Change Research Vol. 30; Madeleine C. Thomson, Ricardo Garcia Herrera and Martin Beniston (ed.), Springer Science + Business Media, Dordrecht; Publisher: Springer Netherlands, ISBN 978-1-4020-6876-8, The Netherlands. 2008.

[16] Conservation Internacional. Integrated National Adaptation Pilot project results – Final report (in Spanish). 122 pages. Available online at: http://www.conservation.org.co/ . 2011.

[17] Instituto de Hidrología, Meteorología y Estudios Ambientales - IDEAM, Colombia. Segunda Comunicación Nacional ante la Convención Marco de las Naciones Unidas sobre Cambio Climático. Bogotá, Colombia. 2010.

[18] Thomson, M.C. and S.J. Connor. The development of Malaria Early Warning Systems for Africa. *Trends in Parasitology* 17(9): 438-445. 2001.

[19] Poveda, G., W. Rojas, M.L. Quiñones, I.D. Vélez, R.I. Mantilla, D. Ruiz, J.S. Zuluaga and G.L. Rúa. Coupling between annual and ENSO timescales in the malaria-climate association in Colombia. *Environmental Health Perspectives* 109: 489-493. 2001.

[20] Thomson, M.C., F.J. Doblas-Reyes, S.J. Mason, R.Hagedorn, S.J. Connor, T. Phindela, A.P. Morse, and T.N. Palmer. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature* 439: 576-579. 2006.

[21] Ruiz, D., G. Poveda, I.D. Vélez, M.L. Quiñones, G.L. Rúa, L.E. Velásquez, and J.S. Zuluaga. Modelling entomological-climatic interactions of Plasmodium falciparum

malaria transmission in two Colombian endemic-regions: contributions to a National Malaria Early Warning System. *Malaria Journal* 5:66, doi:10.1186/1475-2875-5-66. 2006.

[22] Poveda, G., M.L. Quiñones, I.D. Vélez, W. Rojas, G.L. Rúa, D. Ruiz, J.S. Zuluaga, L.E. Velásquez, M.D. Zuluaga, and O. Hernández. Desarrollo de un Sistema de Alerta Temprana para la malaria en Colombia. Universidad Internacional de Andalucía. 182 pages. 2008.

[23] Ruiz, D., A.M. Molina, M.L. Quiñónes, M.M. Jiménez, M. Thomson, S. Connor, M.E. Gutiérrez, P.A. Zapata, C. López, and A. Londoño. Simulating malaria transmission dynamics in the pilot areas of the Colombian Integrated National Adaptation Pilot project. Escuela de Ingeniería de Antioquia, 374 pages. 2011.

[24] Muñoz, Á.G., and C. Recalde. Reporte metodológico sobre el experimento de predicibilidad de malaria en el Litoral Ecuatoriano. Proyecto INAMHI-MAE-SCN-PRAA-PACC. 52 pages. 2011.

[25] Muñoz, Á.G., D. Ruiz, and C. Recalde. Malaria biological models and dynamical downscaling for northwestern South America in the Observatorio Andino framework. ICID+18 / 2nd International Conference: Climate, Sustainability and Development in Semi-arid Regions, Fortaleza, Ceara (Brazil). 2010.

[26] INE-Instituto Nacional de Estadísticas de Chile (2007) VII Censo Agropecuario y Forestal 2007.

[27] Dirección Meteorológica de Chile (DMC). Climatología Regional. Informe Técnico. 47 pages. 2001.

[28] Urrutia de Hazbun, Rosa y Carlos Lanza Lazcano. Catástrofes en Chile 1541-1992. Santiago, Editorial La Noria. 440 pp. 1993.

[29] Norero, Aldo y Carlos Bonilla (ed). Las sequías en Chile: causas, consecuencias y mitigación. Colección en agricultura, Facultad de Agronomía e Ingeniería Forestal, Universidad Católica de Chile. 128 pp. 1999.

[30] Quintana, J. and P. Aceituno. Changes in the rainfall regime along the extratropical west coast of South America (Chile): 30 - 43ºS. *Atmósfera* 25 (1), 1-22. 2012.

[31] Baethgen, W. Climate Risk Management for Adaptation to Climate Variability and Change. *Crop Sci.* 50:S-70–S-76. 2010.

Managing Climate Risk with Seasonal Forecasts

Andrew Charles, Yuriy Kuleshov and David Jones

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/2568

1. Introduction

Seasonal climate refers to average conditions in the atmosphere and ocean over time scales of the order of three months. When considering risks associated with seasonal climate we are concerned with deviations from normal conditions, or 'climate anomalies'. Summers that are hotter than usual, extended drought conditions and exceptionally active tropical cyclone seasons are examples of seasonal climate anomalies.

The countries of the Pacific Ocean are exposed to climate risk across a range of sectors, most notably in water resources, agriculture and disaster preparedness. In Fiji, the forestry industry is affected by an increased likelihood of fires in dry conditions and by access roads becoming too muddy to work on in wet conditions. In Samoa and Fiji the supply of hydroelectric power is vulnerable to rainfall deficiencies, as dams tend to be relatively small in comparison to average inflows. Extreme weather conditions threaten tourism revenue for islands such as Rarotonga in the Cook Islands. Seasonal variations of ocean temperatures, which can drive the migration of species such as Tuna and cause the bleaching of coral reefs in which fish spawn affect the productivity of fisheries which are an important economic resource for countries such as Kiribati. Seasonal variations in surface water and temperature can create more favourable conditions for host vectors of diseases such as malaria, increasing their prevalence. [1]

While many climate anomalies are essentially chaotic and not predictable, there exists large-scale coupling (feedback) between the atmosphere and the ocean, which imparts a degree of predictability to variations of seasonal climate in the atmosphere-ocean-land surface system. The most significant manifestation of this coupling, and the physical source of much of this predictability is the El Niño Southern Oscillation (ENSO), a quasi-periodic mode of variability of the equatorial Pacific Ocean [2]. The primary manifestation of ENSO is in the patterns of sea surface and sub-surface temperature in the Pacific Ocean, with cooler than normal central equatorial Pacific sea surface temperatures termed 'La Niña' and warmer than normal temperatures termed 'El Niño'. During La Niña and El Niño events, feedbacks between the ocean and atmosphere lead to changes in the dominant atmospheric patterns, which influence climatic conditions throughout the world. The ocean processes are slower and more predictable than the atmospheric processes responsible for weather, and their influence on the likelihood of atmospheric states can be used to make predictions, either through characterising this relationship empirically using historical data, or by using a physically motivated model of the coupled ocean-atmosphere system.

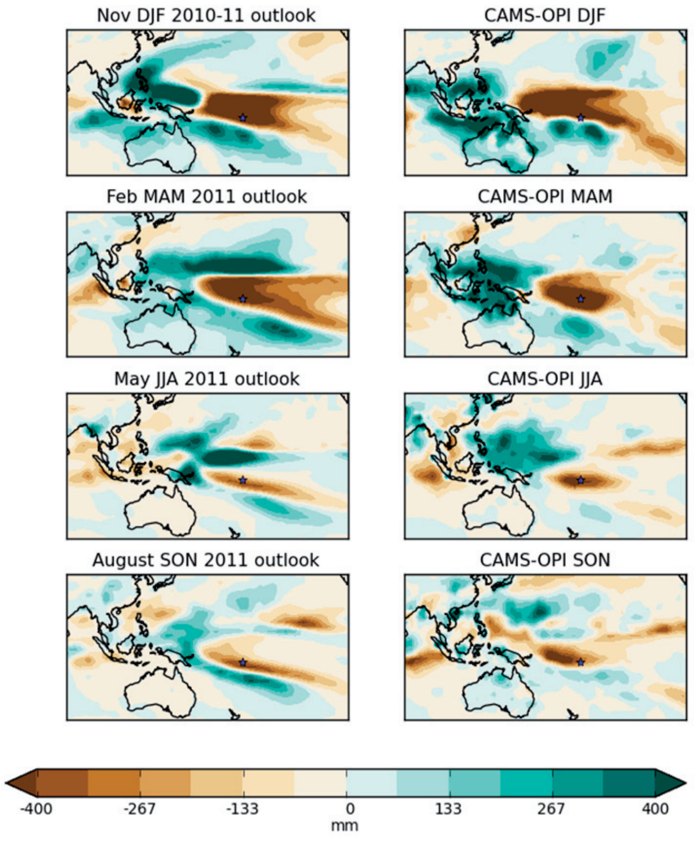


Figure 1. Left: CCGM-based predictions (left) and analysis (right) of seasonal rainfall anomalies in millimetres in the tropical South Pacific region during 2011 for the four calendar seasons starting December-January-February (DJF).

The Tuvalu drought of 2011 provides an example of vulnerability to seasonal climate risk.

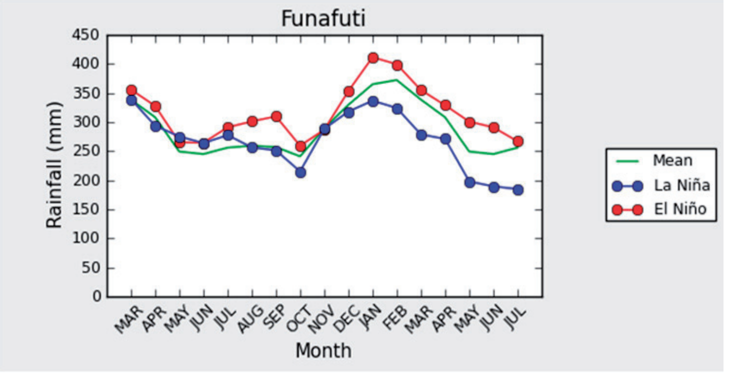


Figure 2. Funafuti (Tuvalu) rainfall in millimetres, composite, during all years, El Niño / La Niña events. Analysis: A.Cottril, Data: Pacific Climate Change Science Project, Tuvalu Meteorological Service (http://informet.net/tuvmet/).

Populations on low coral atolls such as Funafuti (located at 8 South, 179 East) rely heavily on rainwater harvesting for water resources as there are no natural streams or lakes. Rainfall from December 2010 to January 2011 was up to 600mm below normal levels for the western central Pacific region in which Funafuti is located (Figure 1)[3]. Long range rainfall outlooks for the March to May season forecast a continuation of the pattern of suppressed rainfall[1]. These outlooks turned out to be substantially correct, with analysed rainfall deficits of up to 400mm in the region for the period March to May[2]. On the 28th of September 2011, critically low water supplies caused the government of Tuvalu to declare a state of emergency. In early October the governments of Australia, New Zealand, Korea and Japan began delivering fresh water supplies and portable desalination units.

The physical cause of the lack of rainfall in Funafuti in 2011 was cooler than normal waters in the equatorial Pacific, associated with the strongest La Nina[3] episode in recent recorded history, which peaked in the Southern Hemisphere summer of 2010-2011. La Niña events typically decay in Southern Hemisphere Autumn, but in this case the event weakened and then re-established itself in the second half of 2011. The cooler than normal waters in the region of Tuvalu suppressed the rainfall generating convection of moist air, which led to

---

[1] Based on Island Climate Update, a monthly summary of seasonal climate monitoring and prediction in the tropical South Pacific Outlooks issued by the National Institute of Water and Atmospheric Research (http://www.niwa.co.nz/climate/icu) and on seasonal outlooks issued for the region by the International Research Institute for Climate and Society (portal.iri.columbia.edu).

[2] CAMPS_OPI blended rainfall analysis data from the National Center for Environmental Prediction, Climate Prediction Center USA, http://www.cpc.ncep.noaa.gov/products/global_precip/html/wpage.cams_opi.html.

[3] The La Niña phase of ENSO is associated with cooler than normal equatorial Pacific waters and suppressed rainfall in this region. The 2012 La Niña saw record values of indices used to measure the strength of such events http://www.bom.gov.au/climate/current/statements/scs38.pdf

rainfall deficiencies over a sustained period. Figure 1 illustrates that seasonal outlooks based on dynamical models provided guidance anticipating the persistence of these rainfall deficiencies throughout 2011. The tendency towards suppressed rainfall at Funafuti during La Niña events is evident from the composite time series shown in Figure 2. This event illustrates the real nature of climate risk and that, for some phenomena, we now have the capability to predict the features of the earth system that are responsible well in advance.

Many of the examples in this chapter will revolve around the island countries of the Pacific that are directly affected by ENSO and are able to benefit directly from advances in the ability to predict it. Routine seasonal outlooks are issued regularly by national meteorological agencies including the Australian Bureau of Meteorology and The United States National Oceanic and Atmospheric Administration (NOAA), as well as by organisations such as the International Research Institute for Climate and Society (IRI). The availability of seasonal outlooks for the coming seasons gives important information for governments and aid agencies to plan their assistance.

Seasonal outlooks of the likelihood of extreme, synoptic timescale events such as tropical cyclones are also of use for planning disaster preparedness. Tropical cyclones are the most destructive weather systems that impact on coastal areas in the Pacific. While individual tropical cyclones are not predictable beyond timescales of the order of one day, the distribution of tropical cyclone activity is influenced by large-scale climatic features such as ENSO [4].

Climate risk may be assessed in a historically averaged sense, by using the past distribution of extreme events such as droughts or tropical cyclones to give predictive probabilities of the events in the future. Climate change complicates this approach, because while observed changes in the mean state of the climate systems so far have been small, this small change in the mean state can lead to large changes in the frequency and magnitude of extreme events[5]. We refer to this as the influence of climate change on climate variability. The effect of climate change on weather patterns is likely to be considerably more complex than a simple shift of the existing probability distribution. As an example, a recently completed global analyses has found a near 50-fold increase in the frequency of extremely hot temperatures during the northern summer, meaning that the historical occurrence now greatly underestimates the risks of extremes[6]. It has been proposed that a change in climate forcing projects onto the existing modes of variability of the climate system, altering the frequencies and intensities of existing weather regimes[7] [8]. An example of such a mechanism is the prospect that global warming has intensified the hydrological cycle, causing more extreme flooding and droughts [9]. The current set of coarse resolution GCMs used to evaluate anthropogenic climate change may not be sufficiently detailed to capture such nuanced responses, and as such considerable uncertainties remain about the impact of climate change on weather events. In the face of these uncertainties, an effective and low cost option to reduced vulnerability to climate change is to improve the accuracy, availability and use of forecasts[10].

The aim of seasonal forecasting is to predict the average weather or aggregate weather over a long period, usually three months. By exploiting the relationship of weather systems with large scale, long time-scale coupled ocean-atmosphere processes, probabilistic forecasts can be made of the likely tendency of conditions in the coming season. Seasonal predictions are not deterministic, in other words they do not make a prediction that a single outcome will or will not happen. Rather they give a statement of risk, typically about the likelihood of wetter than normal, or warmer than normal conditions.

A range of potential applications for seasonal outlooks has been identified. As noted, in countries dependant on rainwater harvesting for water supplies, advance knowledge of drought conditions can allow pre-emptive water saving or water supply bolstering. Knowledge of the relative likelihood of fires or inaccessibility due to rainfall could be used to plan forestry activities. Rainfall outlooks can be used to estimate the availability of water for hydroelectric power generation, and to pre-emptively purchase fuel for backup generators, avoiding the payment of expensive spot rates for fuel. Tourism operators can develop forward plans that take into account changes in the likelihood of climatic disturbances. Reefs likely to suffer from elevated temperatures can be declared off-limits for fishing and tourism to reduce other sources of stress on corals [11]. Seasonal variations in surface water and temperature can increase the prevalence of certain diseases such as malaria by causing more or less favourable conditions for host vectors [1]. The beef industry in Vanuatu can benefit from forward estimates of how many head of cattle a pasture will be able to support. Seasonal forecasts have been shown to be of economic utility in the management of wheat farming in Australia by guiding changes in practice such as crop row spacing and fertilizer application [12].

2. The limitations of empirical models and the imperative for a dynamical model basis for seasonal forecasting

Empirical models (or 'statistical models') are currently used by many meteorological services for seasonal climate outlooks. These models are based on empirical relationships, usually between ENSO based indices ('the predictors') and variables such as local rainfall and temperature ('the predictands'). Using current observed values of ENSO indices these past relationships can be used to create forecasts [13].

A warming of the climate system due to greenhouse gas forcing is predicted by theory, demonstrated by numerical predictions and has been observed over the course of the past century [14]. While the empirical relationships between climate predictors and predictands such as rainfall may be robust, in a warming climate, environmental indicators used as predictors are now frequently outside of the range of historical records, meaning that relationships are being assumed for events which do not have an historical analogue. In general, empirical models cannot reliably account for aspects of climate variability and change that are not represented in the historical record. Empirical forecasting usually depends on the assumption of stationary relationships between predictors and predictands. This also renders such schemes susceptible to periodic changes in these relationships due to decadal timescale variability.

For example, outlooks for tropical cyclone (TC) activity in the Australian region are based on a regression model using values indices representing major modes of variability in the ocean. The 2010-11 TC season featured to a very strong La Niña event with an unusually hot Indian Ocean, an event without historical precedent. In this case the statistical models significantly over-predicted the number of TCs that occurred in the Australian region. Analysis shows that the environmental indicators used for tropical cyclone seasonal outlooks for the Australian region in 2010-11 and 2011-12 are outliers in the predictor phase space, in other words, outside of the range of variability for which the model was tested and built.

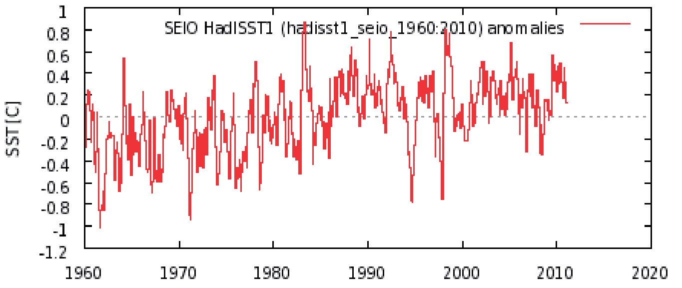


Figure 3. Time series for an Indian Ocean based climate index showing a trend. Such trends may reduce the efficacy of empirical prediction schemes. (Source: KNMI Climate Explorer: http://climexp.knmi.nl/).

Inter-annual variability in the intensity and distribution of tropical cyclones is large, and presently greater than any trends that are ascribable to climate change. However climate change impacts our ability to make skilful predictions of tropical cyclone activity using empirical models, because in the warming environment predictors such as SSTs now frequently lie outside of the range of past variability. Improved empirical methods can be developed to adjust for this, by incorporating trends and by treating predictors that lie outside the observed range of variability more cautiously [15]. However it is widely considered that dynamical models provide the best prospects for improved seasonal forecasting in the future, either through providing long range forecasts of environments favorable to cyclo-genesis, or through high resolution models that can provide an estimate of the number of cyclones expected to form.

2.1. Seasonal forecasting with dynamical models

An alternative paradigm for seasonal prediction is the use of coupled ocean-atmosphere General Circulation Models ('coupled models' or GCMs). State of the art coupled models consist of a physically based model of the ocean, usually solved using a grid based scheme, coupled to a physically based atmospheric model, often solved using a spectral spatial discretisation [16]. GCMs solve a set of dynamical equations ('the primitive equations') to project the current analysed state of the ocean-atmosphere system into the future. The term 'analysed' here is used quite deliberately to describe methods used to determine a global estimate of the state of the ocean and atmosphere based on the combination of available observations and using numerical methods based on a mix of physical and statistical relationships to infer the state of regions not subject to direct observation. The objective of the assimilation process is placing constraints on the observations to ensure they present a physically plausible set of initial conditions for the ocean-atmosphere simulation.

A number of dynamical models are run at operational meteorological centres around the world. We briefly describe the main components here with reference to the model used operationally at the Australian Bureau of Meteorology for ocean temperature forecasts. The first component is an ocean data assimilation system, which provides an estimate of the state of the upper ocean based on an analysis of ocean observations. The observations of the ocean come from a variety of sources including satellite observations of sea surface temperature and sea level height, fixed, drifting and profiling buoys (such as the TOGA-TAO array which provides real-time observations of the region of the Pacific Ocean central to ENSO) and observations taken from ships.

This ocean assimilation system initialises an ocean model though a complex process which attempts to bring the model into a state consistent with the oceanic observations but also such that it is internally balanced to minimise so called 'initialisation shock'. Ocean model resolution for the Bureau of Meteorology POAMA (Predictive Ocean Atmosphere Model for Australia) [17] is 2 degrees in longitude with a latitudinal resolution telescoping from 0.5 degrees near the equator to 1.5 degrees near the poles. The model resolves 25 vertical levels. Specialised coupling software is used to transmit surface fluxes of heat and momentum between the ocean model and an atmospheric model. The POAMA atmospheric model has a spherical harmonic horizontal structure with triangular truncation at wave number 47 (grid cells of roughly 250km by 250km when transformed) and 17 pressure levels. The atmospheric model typically has its own assimilation system to ingest data from available observations of meteorological parameters including wind, pressure and temperature. Coupled assimilation, in which the ocean and atmosphere are initialised together to reduce initialisation shock is an area of current research.

Processes with a spatial scale smaller than the model grid scale are 'parameterised', which means a statistical or process-based model is used to represent the average effect of this process on the sub-grid scale. Design and configuration of sub-grid scale processes is a specialised and active area of research, with current activity focussed on the use of stochastic models to better capture the uncertainty of the sub-grid processes.

Seasonal climate prediction is inherently probabilistic because the evolution of the climate system is highly sensitive to initial conditions. Small difference or 'errors' in the description of the initial climate state grow with time leading to very different forecast outcomes. To estimate the range of physically plausible outcomes, GCMs are typically run as an ensemble, in which a number of simulations are performed with slightly different initial conditions. The initial conditions are perturbed to realistically sample the plausible range of initial climate states.

Ensemble strategies in theory allow for better estimation of the probability of extreme, or less likely events. The nonlinear nature of the coupled ocean-atmosphere system means that these probabilities may not be well estimated from a single 'best guess' deterministic forecast. Using simple decision models which will be discussed in more detail below, Palmer [18] demonstrated that the economic value of ensemble forecasts is greater than that of individual models or simple ensemble means.

Because basic physics does not change under global warming, dynamical models are less compromised by climate change than statistical models. GCMs explicitly take into account climate processes that are important for seasonal climate prediction such as equatorial oceanic waves and atmospheric convection driven by ocean temperatures and are not constrained by what has occurred in the past. GCMs implicitly include the effects of a changing climate whatever its character or cause and can predict outcomes not seen previously.

3. Model based forecast products

Seasonal climate forecasts are inherently probabilistic due to imperfect model initialisation, instabilities in the modelled system, and model error. One approach to transform a GCM ensemble forecast into a probabilistic forecast is to define one or more event thresholds, and then take the fraction of ensemble members above this threshold as the probability forecast. This approach effectively takes the model ensemble distribution as a best guess of the probabilities of future states of the system. These can be referred to as 'ensemble relative frequency' or 'perfect model' probabilities, as they assume that the model ensemble is a perfect sample from possible futures consistent with the model initial conditions. This procedure does provide an adjustment for model biases, for example if the model tends to be biased towards warmer temperatures, because the ensemble distribution for a particular realization is measured against the model's own climatological state.

One event for which probabilities may be desired would be the occurrence of above median monthly rainfall over a region of interest. Figure 4 shows the POAMA hindcast ensemble for the year 1997 and its conversion to a probabilistic forecast of the event of monthly rainfall being above the long-term median in the Murray Darling Basin, a region of high agricultural importance [19]. This probability forecast was generated for retrospective seasonal forecasts generated with the POAMA 1.5 model for the period of 1980 to 2006. The individual ensemble members show that for each month a range of outcomes is possible including both above and below media rainfall. These retrospective forecasts are produced from the first season of model output, meaning there is no time elapsed between the model initialisation and the period being forecast for.

4. Accounting for model error

Uncertainty in probability forecasts can be divided into three distinct categories. The first category of prediction uncertainty is linked to the non-linearity of climate dynamics that causes a sensitivity to initial conditions. This is the so-called butterfly effect, which imposes

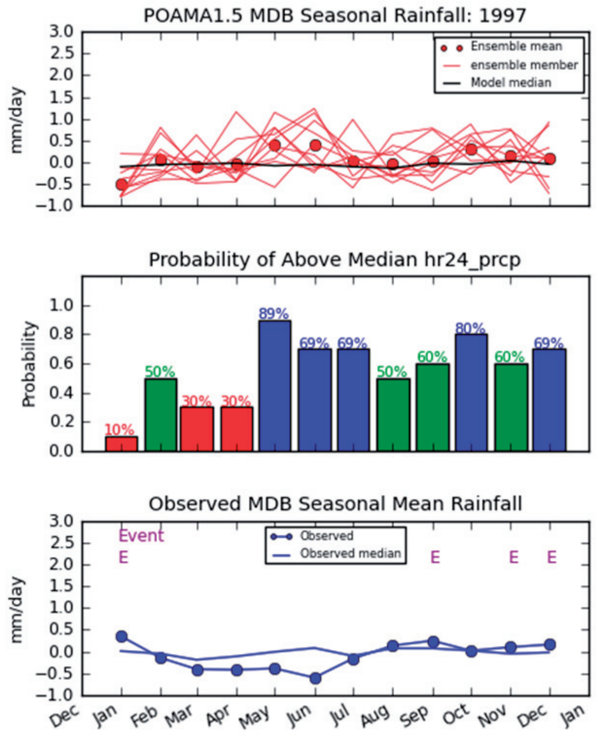


Figure 4. Retrospective forecasts of mean seasonal rainfall for the Murray Darling Basis produced using the POAMA 1.5 CGCM for 1997. Upper) Time series of the model ensemble rainfall anomaly (mm/day), Middle) Probability forecast derived from the number of ensemble members lying above the model median, Lower) Observed seasonal rainfall anomaly (mm/day), where E denotes the occurrence of the above-median event.

hard limits on our ability to make deterministic predictions of nonlinear systems. The simple fact that we do not have infinite precision means that instabilities on scales smaller than the smallest resolved model scale inevitably grow and affect the larger scale until no predictive skill remains [20]. The 'saturation time' after which the system is effectively unpredictable is longer for the ocean than the atmosphere.

The second major category of prediction uncertainty is the sparseness and imprecision of earth system observations. As discussed above, the analysed state of the atmosphere and ocean is necessarily different from its actual state, and as such model projections are projecting an imperfect estimate of the initial state forward in time. As such even with a perfect physical model, predictions would be imperfect. This source of error interacts with the first, because instabilities growing from initial conditions that are not present in nature may produce possible future states that are inconsistent with actual potential future states. Ensemble forecasting allows this initial condition uncertainty to be estimated and quantified by sampling the space of plausible initial conditions and projecting this sample forward in time. These two kinds of uncertainty can be described as 'flow dependent'[21] because their rates of growth and magnitudes are sensitive to the stability of the point in phase space characterising the flow.

The final category of uncertainty is model error, the fact that our mathematical idealisations of the climate system are not perfect. This includes errors due to imperfect physical parameterisations, errors due to unresolved processes at the sub-grid scale and differences between the mean state of the model and the true system. This class of error is widely studied and motivates research into better models with improved representation of physics, and model calibration techniques that can account for or correct the errors.

Single model ensemble forecasts only capture the components of prediction uncertainty associated with uncertain initial conditions and model-captured instability, and these are only fully captured in the ideal case of an infinite ensemble that uniformly samples initial condition uncertainty. An ensemble of a single model provides no information about the model error component of prediction uncertainty (Stephenson, 2005), and models that are structurally similar will invariably share biases.

4.1. Assessing forecast error

Forecast validation is the process of measuring the correctness of a set of issued forecasts. It can be thought of as being distinct from model validation which is about determining whether a model correctly resolves physical processes [20].

Here we give an example of forecast validation based on the definition of discrete events, for example the event of rainfall over a three month period exceeding a given threshold, and of categorical forecasts, for example low, medium and high probability of the event. For three forecast categories, the contingency table summarising the forecast- verification set has the form shown in Table 1, with forecast categories f_1, f_2, f_3 counts of observed events o_1, o_2, o_3 and counts of non-events n_1, n_2, n_3 over each forecast. The 'distributions oriented' theory of forecast verification interprets the contingency table statistics in terms of the joint, marginal and conditional probability distributions of events and forecasts [22]. In this theory, the contingency table contains all the information required to generate a standard set of verification scores.

Forecast	Observed Events	Observed Non Events
f_1	o_1	n_1
f_2	o_2	n_2
f_3	o_3	n_3

Table 1. Contingency table for a binary event with three forecast categories.

Forecast	Events	Observed Non Events	Total
Low	37	59	96
Medium	67	70	137
High	59	32	91

Table 2. Contingency table for MDB seasonal monthly rainfall hindcasts from POAMA 1.5, all start months.

The joint distribution of the forecasts in one bin f_i and observed events e is $p(f_i, e) = o_i/N$ where the total number of forecast-verification pairs is $N = \sum o_i + \sum n_i$.

The calibration-refinement factorisation of the joint distribution for a particular forecast bin,

$$p(F_i, E) = p(E|F_i)p(F_i),$$

is composed of two factors: the true positive ratio $p(E|F_i)$ and the marginal frequency $p(F_i) = (o_i + n_i)/N$, where

$$p(E|F_i) = \frac{o_i}{(o_i + n_i)}.$$

The true positive ratio $p(E|F_i)$ is the conditional probability of the event given this particular forecast, while $p(F_i)$ is the probability that the forecast system produces this category of forecasts, which indicates if the system is biased in one way or another. $p(E|F_i)$ can be considered an estimate of the expected probability of the event of above median rainfall based on the information from the forecast and its verification.

We apply this simple forecast validation scheme to the POAMA MDB rainfall forecasts discussed above. In order to compute meaningful statistics on these probability outlooks, three bins for the probability of rainfall exceeding the climatological median were used. A small number of forecast verification pairs in any particular bin reduces the statistical significance of results markedly. Larger probability bins can be used to mitigate this, but at the expense of forecast resolution and sharpness. The three bins translate into categorical forecasts of a low, medium and high probability of an above median rainfall event. The binned forecasts were verified against Australian rainfall data from the Australian Bureau of Meteorology National Climate Centre's gridded atmospheric data set [23].

Table 2 shows these counts for the MDB rainfall forecasts described above for all months in the hindcast period.

Forecast	Mean Ensemble Frequency (Model probability)	p(E∣F)	90% Probability Interval of p(E∣F)
Low (0-33%)	0.21	0.39	0.31 - 0.47
Medium (33-66%)	0.50	0.49	0.42 - 0.56
High (66-100%)	0.80	0.65	0.56 - 0.73

Table 3. Calibration table for GCM forecasts of above median seasonal rainfall, computed using data in table 2 with 90% probability interval.

If the calibration distribution in each bin is assumed to be a Bernoulli distribution, probability intervals for the parameter can be generated for the forecasts by a permutation counting method. An alternative method for larger datasets for which permutation counting is prohibitive is to use percentiles of a normal posterior distribution. Table 3 gives the true positive ratio with a 90% probability interval for the data in Table 2. It can immediately be seen that the probability distribution implied by the model ensemble is not consistent with the probability distribution implied by the verification of the forecasts. For example the mean probability for 'low probability' forecasts is 21%, but the event occurs 39% of the time for this forecast category.

The earth system is very high dimensional and the procedure used here reduces the dimensionality of the problem. Such dimension reduction may result in a loss of information about the performance of the system – we are faced with a trade-off between information contained in the model-based forecasts, and seeking to extract information from the model-reforecast dataset. In this case simple binning is used, more sophisticated methods such as principal component analysis could also be employed. We will return to this point later in the chapter when calibration is discussed.

Forecast	Events	Nonevents	p(E|F)	90% Probability Interval
Low (0-33%)	3	6	0.33	0.14 - .59
Medium (33-66%)	1	3	0.25	0.4 - 0.63
High (66-100%)	10	4	0.71	0.50 – 0.87

Table 4. True positive ratio for POAMA MDB June-July-August rainfall.

4.2. Assessment of probability forecasts

In assessment of probability forecasts the two main aspects of performance are resolution and reliability. Reliability is defined as the degree to which the observed frequency of an event coincides with its forecast probability. Reliability does not guarantee useful skill, but forecasts that are not reliable cannot be taken at face value and must be adjusted, either implicitly as occurs when a verification plot demonstrating overconfidence is published next to a forecast or explicitly by downgrading probabilities that are not justified by model performance. The term 'well calibrated' is used to describe probability forecasts that are reliable. Resolution is defined as the frequency with which different observed outcomes follow different forecast categories, in other words the degree to which the forecast system can 'resolve' different outcomes.

Figure 5 shows the reliability diagram for the POAMA 1.5 Murray Darling Basin Average monthly mean rainfall, for all months. The green bar marks the forecast 90% probability interval (ci), the purple bar marks the 90% probability interval for perfect forecasts with the same sample size. Reliability diagrams are plots of the true positive ratio (also known as the calibration function, observed relative frequency, likelihood and hit rate) against the mean probability of the forecasts in each bin. Reliability diagrams are used to assess the degree to which the model forecast probabilities agree with the observed frequencies, shown in figure 5 with the probability intervals described above. The figure shows that even when small sample size is taken into account, the forecasts are overconfident. Resolution is represented by the spread of points on the reliability diagram in the vertical – it can be seen the model has some ability to resolve between the two outcomes.

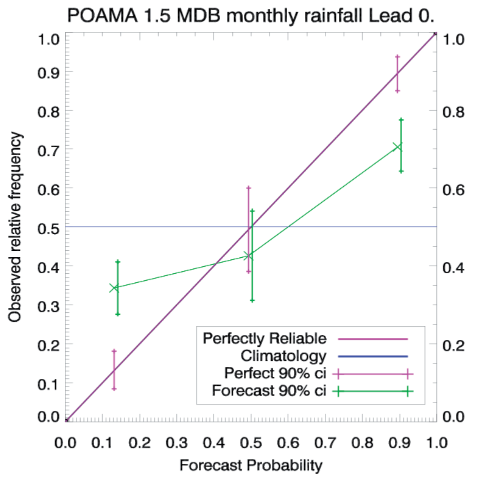


Figure 5. Reliability diagram for Murray Darling Basin monthly rainfall forecasts.

5. Understanding prediction utility: Simple decision models

In order to begin to understand potential uses of seasonal forecasts it is instructive to study simple cost-loss decision models. Such simple models provide a framework to begin to quantify the potential value of forecasts. Before proceeding, we note that real-world decisions are typically made with far more parameters and subject to greater uncertainty regarding potential costs and payoffs than the simple models studied here.

We first consider a simple binary event, binary decision model in which there are two possible outcomes – the occurrence or non-occurrence of an event - and the user makes a decision to protect, or not protect, against the event. Protection has a cost; failure to protect incurs a loss. The classic example is the decision to carry an umbrella to protect against the possibility of rain. A seasonal timescale example is the decision to apply fertiliser to a crop based on the likelihood of future rainfall over a season. An early study of these issues was made by Anders Angstrom as documented by [24].

A failure to protect with cost C results in a loss L. In this framework it only makes sense to take action given the probability of the event P if P > CL. If it is not, then the expected loss is less than the cost of taking protective action. The combination of the joint distribution of forecasts and observations and the decision-makers cost function determines the potential economic value of the forecasts. (Table 5.)

5.1. Adjusting model output: Introducing calibration

Decisions about the use of GCMs for seasonal climate forecasting are usually based upon measures of model performance over a hind-cast (retrospective forecast) period. A natural and popular extension of this idea is that GCM-based forecasts should be adjusted by this skill assessment. This motivates 'Model Output Statistics' methods[25] and ‘model calibration’ and has been widely adopted in medium range weather forecasting [26] and seasonal forecasting [27] [28] [29].

In order to make rational decisions based on quantifiable costs, losses and probabilities the end user needs the calibrated forecast probabilities, and needs to know what their costs and losses are for each contingency. Given the calibrated forecast probabilities, with reliable confidence intervals, they are in a position to use these probabilities to determine the optimum course of action to follow for their unique cost function. Given information about climatology, a model and its verification, the calibrated model probability $p(E|F)$ is this best estimate, subject to the assumptions made in determining the calibrated probability.

As a simple example of calibration, consider the true positive ratio calculated above. While crude and subject to sampling error, this represents the conditional probability of the event given the model forecast category. The true positive ratio, proposed as the best estimate of the event probability from the POAMA MDB seasonal outlooks discussed above is a conditioning of probabilistic forecasts derived from the GCM ensemble upon probabilities obtained from comparison of the hindcast set with observations. These conditional probabilities are needed for users to make optimal decisions [30]. Skill for coupled models is commonly presented as correlation plots, mean error plots and sometimes more esoteric scores for probabilistic forecasts. While these scores are useful for model diagnostics, and can quantify potential forecast value, it is not obvious how users who need to make decisions based on forecasts should convert these measures into new estimates of probability. We note that some effort has been spent into developing verification measures that do have a direct relationship to economic value such the ROC (Receiver Operating Characteristic) score and the logarithmic score based on the information content of a forecast.

Resolution can be degraded by calibration and it is expected that the application of calibration techniques will involve some trade-off in which resolution is traded for reliability. It is also the case that cross-validation methods used on the application of calibration in order to avoid ‘artificial skill’ can also result in artificial reduction in skill scores, and thus in the assessment of such methods it can be difficult to disentangle cross validation artefacts from true reduction of model skill due to calibration.

This simple calibration framework can be extended: similar methods can be applied to parametric probability density functions [28]. Below we discuss different calibration methods, but first we turn to more sophisticated decision models.

Event	Action	No Action
Yes	C	L
No	C	0

Table 5. Simple binary cost-loss model.

5.2. Extending the binary cost-loss model

In the simple cost-loss model the cost of taking protective action is the same whether the event occurs or does not occur. While this may be true for many economic decisions, when social and political dimensions are considered there is a clear penalty, in terms of confidence in the forecasting system and reduced possibility of action in the future, for false alarms. The binary cost-loss model can be developed further to include such a false alarm or 'cry wolf' effect. Such an extension is effectively an adjustment for the deviation from perfect rationality of forecast users.

The above model can also be extended to more sophisticated decisions based on event probability thresholds, with different actions to be taken at different probability thresholds, depending on the users attitude to risk. We present a hypothetical example of an agriculturalist making a decision about whether to apply additional fertilizer, at a cost, with a potential payoff depending on the probability of expected rainfall being above median. In this example 20% rainfall probability is the threshold at which the cost of applying fertilizer is less than the expected payoff (Table 6). The decision thresholds in Table 6 provide a way of mapping from a given forecast to an action, again in relation to a binary yes/no event. Such tables are dependent on the details of individual enterprises and must be determined with regard to their operating costs and potential losses. The premise for Table 6 is the decision by wheat farmers to apply top-dressed fertiliser in order to benefit from expected rainfall[12], however the numbers selected are arbitrary and shown for illustration. Another management decision that could be studied using this methodology is choice of cultivar, for example to decide whether to plant a drought tolerant strain of wheat or one with a higher potential yield in the event of good rains.

Forecast Probability	Action	Outcome if Sufficient Rainfall	Outcome if Insufficient Rainfall
0-20%	No fertilizer application	Missed profit	Minimal loss
20-70%	Normal fertilizer application	Normal profit	Moderate loss
70-100%	Maximum fertilizer application	Bumper crop	Greatest loss

Table 6. Example probability thresholds for a decision about whether to apply no fertilizer, a normal amount, or a maximum amount to take advantage of expected seasonal rainfall.

Using the true positive ratio we calculated for our sample rainfall forecasts in Table 3, the farmer would find that the calibrated 'low probability' forecasts from POAMA are not sufficient to justify the 'no fertilizer' action, because the observed frequency of above median rainfall events is above the 20% threshold. In other words, while the forecast have skill they do not have value to this particular decision.

Another simple decision model is the theory of Kelley betting, which deals with cost-loss scenarios in the context of gambling. In this theory a gambler bets a fraction w_i of their wealth on an outcome i, where $\sum_i w_i = 1$. The ratio of the gamblers post bet wealth to his pre bet wealth is $W = \sum_i p_i \log_2 o_i w_i$ where o_i is the wealth multiplier, or odds, assigned to the outcome and p_i is the event probability.[31]

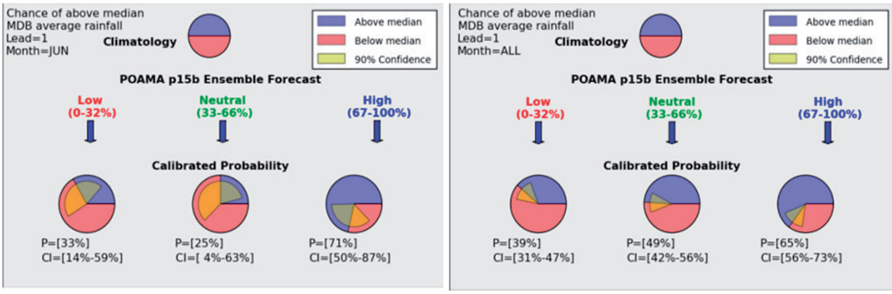


Figure 6. Posterior probability of above median seasonal rainfall for all forecasts (right) and June-July-August forecasts (left) initialized at the end of March.

5.3. Presenting probability outlooks

We now turn our attention to ways of presenting information about forecasts and their skill-based calibration. The actual contingency table (Table 2) has the advantage of containing almost all the usable information (assuming the stationarity of the marginal distributions), but the disadvantage of requiring knowledge of verification methods to translate it into usable probabilities. A plot of the actual ensemble of past forecasts (Figure 4) allows users to eyeball the agreement and spread between forecasts and observations. However it provides no quantitative information about how much credibility to assign to a particular forecast. The reliability diagram (Figure 5) provides this information, but it is not intuitive to interpret for most users. A simple pie chart can also be used to present relative probabilities. Figure 6 shows visually how the model forecast adjusts the model estimated probabilities, and what the credible intervals based on the size of the sample are. It shows the prior climatological probability of the event and the updated probabilities, with 90% credible intervals for each forecast category. This plot is designed to communicate to end users how much the forecast ought to affect their estimate of the event's probability, based on the rate of event occurrence for previous forecasts.

Coupled model skill varies strongly by month, but using the simple binning calibration method this information is difficult to resolve. Table 4 shows the contingency table and true positive ratio for June-July-August seasonal forecasts. The true positive ratio suggests that the forecasts have reasonable skill and that we ought to take the forecast of a high probability of above median rainfall as increased from a 50:50 climatological odds to 9:2 in favour of the event. Unfortunately the small sample size in each probability bin results in very large probability intervals as shown in Figure 6 (left). The wide probability intervals around our estimate of skill by month are troubling, because we know that skill varies strongly by month but are unable to quantify this adequately for these forecasts. Pooling forecast-verification pairs in order to increase confidence is one way to increase sample size, by aggregating forecasts at different locations and times. Both procedures will reduce the size of our credible intervals, but risk increasing the autocorrelation of the forecast data. A similar sample size problem affects the statistical significance of attempts to calibrate forecasts for individual grid points.

A question for forecast users is how the probability range should affect the decision. The wider the interval, the less evidence exists that the forecast probability corresponds to a repeatable relationship between model and reality. Decision makers may prefer to assume climatological probabilities until this information can be sharpened. Theoretical work or modeling could determine optimum forecasts for selected decision making cost functions.

5.4. Adjusting for Model Error in Continuous Forecasts

Calibration methods can be considered to adjust the probability distribution produced by the model by using information about its past performance, with the aim of providing unbiased and reliable forecasts. A straightforward approach to the generation of probability outlooks is to build a linear regression model for predictand y using the GCM ensemble mean x as a predictor:

$$y = \alpha x + \beta + \epsilon$$

where α and β are regression coefficients which may be computed by the least-squares method such that $\alpha = r\frac{\sigma_x}{\sigma_y}$ and $\beta = \frac{1}{N}\sum y - \alpha x$, with correlation coefficient r. The random errors ϵ are typically assumed to be normally distributed with a standard deviation equal to the mean of squared regression residuals $\sigma_\epsilon = \frac{1}{N}\sum(y - \alpha x - \beta)^2$. When applied to synthetic data it can be shown that this scheme produces probability forecasts that are reliable in the sense discussed above. A limitation of this method is that it assumes the data are normally distributed, significant deviations from normality may require that data are transformed or that different methods be chosen to compute the regression coefficients. Analytical methods can be used to assess the errors in the regression parameters but again these are usually based on distributional assumptions.

Such regression-based approaches can be made more robust to small sample size using Bayesian methods in which model parameters θ are given by Bayes theorem as

$$p(\theta|H,O) = \frac{p(H,O|\theta)}{p(H,O)}\,p(\theta)$$

with hindcast data H and observations O. The likelihood function $p(H,O|\theta)$ estimates the probability of observing the hindcast-observation series given a set of model parameters. $p(\theta)$ is the prior probability for the model parameters. Probability density functions for model parameters θ can be determined using Markov Chain Monte-Carlo sampling [32].

5.5. Variance inflation

Johnson and Bowler (2009) outline a variance inflation technique which adjusts the ensemble forecast to meet two conditions: a) that ensemble members have the same variance as observations, and b) that the root-mean-square error of the ensemble mean be equal to the spread of the ensemble. A major difference between this and the previous method of linear regression with residual errors is that the ensemble spread remains a major determinant of forecast uncertainty. The first condition is designed to achieve the statistical indistinguishability of the first two moments between ensemble members and observations. The second condition is designed to ensure that the ensemble spread accounts for the expected model error. These conditions are achieved by increasing (or decreasing) the perturbations of the ensemble members from the mean while keeping the correlation between model and truth is unchanged (except in the case of a negative correlation between model and truth, in which case the sign of the correlation is reversed).

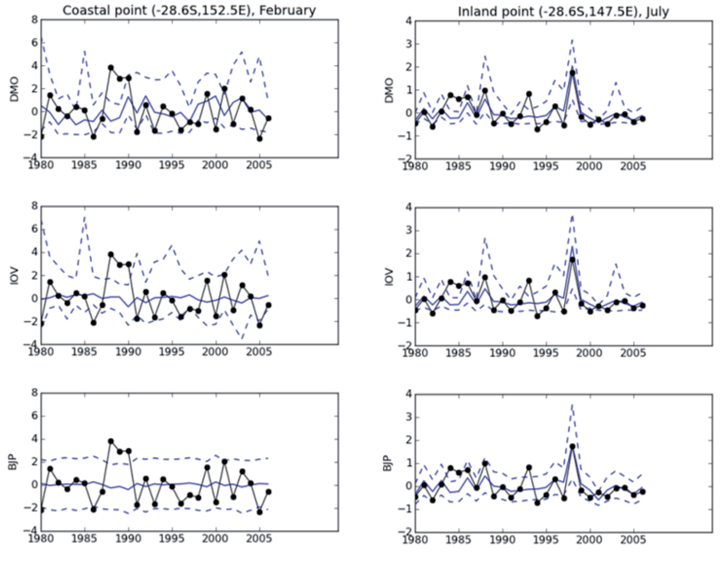


Figure 7. Effect of calibration procedures on model time series. Left: A model grid point with high hindcast correlation. Right: A model grid point with low hindcast calibration. Black line: observations, blue solid line: model mean, blue dashed line: 10% and 90% model probability intervals.

Given ensemble mean $\bar{f}$ and ensemble member perturbations ϵ_i, adjusted ensemble members g_i are constructed by

$$g_i = \alpha\bar{f} + \beta\epsilon_i$$

Coefficients α and β are computed as

$$\alpha = \rho \frac{\sigma_x}{\sigma_{\bar{f}}}$$

and

$$\beta^2 = (1 - \rho^2) \frac{\sigma_x^2}{\sigma_\epsilon^2}$$

with observed variance σ_x, ensemble mean variance $\sigma_{\bar{f}}$, correlation between observations and ensemble mean ρ and time average of ensemble variance σ_ϵ. Leave-one-out cross validation is used for the calculation of correlation and standard deviation when constructing a calibrated hindcast set. Typically the time series for each GCM grid point is calibrated independently. Johnson and Bowler show that under the assumption of normally distributed model predictions and observations, this procedure minimises the root-mean-square error.

5.6. Regression estimate of event probability

Another method of adjusting probability forecasts is to regress the forecast probabilities directly against the observed events/non-events frequencies. While having the drawback that it is computed directly on ensemble-derived probabilities, it has the advantage that it makes no distributional assumptions, and estimates only one parameter.

5.7. General remarks on calibration

In the case of overconfident forecasts, calibration procedures reduce the amplitude of the probabilities, adjusting for this overconfidence by reducing the resolution. Conceptually, this calibration step can be considered the application of a statistical model to the direct model output in which the forecasts are corrected for mean state bias and over-confidence in the ensemble distribution. Figure 7 presents the application of two calibration methods to model time series which exhibit a high and low correlation with the verifying observations respectively. The central panel demonstrates the effect of the variance inflation adjustment described above, while the lower panel shows a regression adjustment with Bayesian parameter estimates.

It could be argued that such procedures degrade, or corrupt model outputs, because they make use of only limited information from the model reforecast set and available observations. This may be the case, but if such information can be specified it can be included in the calculation of calibration factors. If it cannot be specified and measured, then we are hardly in a position to use it to inform our estimates of future probabilities!

In seasonal forecasting, calibration is complicated by the short length of the hindcast verification data set, typically 15 to 30 years, which imposes hard limits on how much information we can reliably say we have about the model. This paucity of data makes model skill assessments and model adjustment difficult because parameters calculated from the verification dataset will necessarily have large sampling error. For this reason it is desirable that calibration models have a small number of parameters.

The problem is even thornier, because some circulation regimes such as strong El Nino events are thought to be more predictable than others, so there is every chance other variables may be strongly related to the predicted accuracy of a given forecast. Indeed the practice of ensemble forecasting is designed to reflect such changes in potential predictability. For example the influence of strong El Niño and La Niña events leads to greater predictability of climate anomalies in affected regions during such events. Given this knowledge, information about the state of ENSO should in theory be used to estimate the certainty of seasonal outlooks. Empirical outlooks do just this, but schemes using this information for dynamical model based outlooks are not yet common.

5.8. The multi-model approach

Another approach to the quantification of model error is to combine forecasts from a number of different yet plausible dynamical models. Multi-model combination aims to benefit from a better representation of uncertainty in model physics, model configuration and initialisation strategy. The multi-model approach is widely used in operational weather prediction (out to 7 to 10 days ahead). Model combination is complicated by varying grid resolutions, ensemble sizes, different model skill and mean biases between models, as well as unresolved questions about model weighting. The multi-model approach has been criticized on the grounds that combining a forecast from a bad model with a forecast from a good model may result in a less skillful forecast if one does not weight models to reflect their level of skill.

5.9. Downscaling

Another family of model adjustments is motivated by the mismatch between the resolved scale of GCMs and the scale at which most decisions are made. GCMs can provide useful forecasts of atmospheric fields at seasonal timescales but are typically run at coarse spatial resolution such that the direct model output represents spatial averages over thousands of square kilometres (typically grid cells some 100 km in size). This coarse resolution poses a problem for applications that require forecasts at a finer spatial scale, especially in regions where the real topography causes local rainfall to diverge significantly from model grid averages. Where the errors to be corrected are primarily a result of the spatial scale of the GCM, the correction is called ‘downscaling’. Downscaling is desired for those Pacific islands where the interaction between the prevailing winds and local topography is a significant driver of variability, but the GCM does not resolve local topography.

The primary goal of downscaling is to replace the large-scale grid box climate variable, in this case rainfall, with rainfall that is better representative of the local situation. One method of downscaling is that of meteorological analogues. In this approach, large-scale synoptic meteorological fields are used as predictors for small scale variables. The output of a seasonal timescale GCM is used to generate forecasts of the large-scale fields. The analogue methods has been shown has been shown to produce good results for Twentieth Century South Eastern Australian rainfall in the context of downscaling for climate change projections[33]. As with most statistical downscaling techniques, analogue downscaling is computationally cheap, in contrast to resource-intensive dynamical downscaling using nested atmospheric models.

Figure 6 shows the topography resolved by a high-resolution numerical weather prediction model, and the topography resolved by a coarse resolution seasonal prediction model.

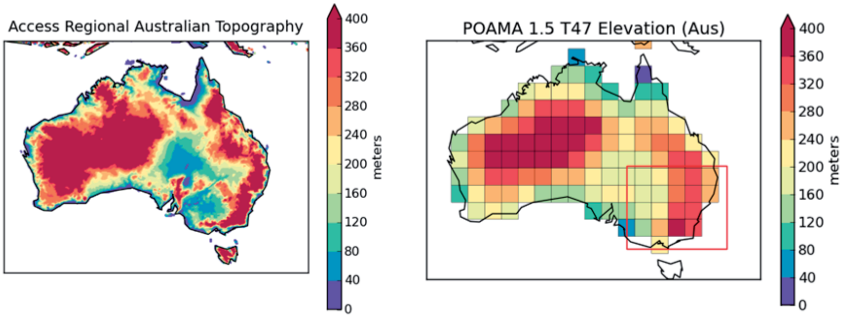


Figure 8. Left: topography resolved in a high resolution weather model. Right: topography resolved in a coarse resolution seasonal prediction GCM.

6. Software architecture: From models to systems

The design of systems for the generation and distribution of GCM based outlooks is architecturally complex. It is here that the interdisciplinary nature of the seasonal forecasting activity becomes clear. In addition to more traditional earth system science involved in understanding coupled ocean-atmosphere processes, the tasks of data processing, data modelling and information system architecture require advanced computing skills. We now outline a general pattern for the design of systems for the delivery of seasonal forecasts to end users which is a generalisation of the implementation described above and in [34].

Four distinct layers can be defined as components of the overall process of turning the outputs of GCMs into seasonal outlooks suitable for use by decision-makers.

The **model layer** comprises the GCM simulating the evolution of the coupled ocean-atmosphere system. This component is a complex software system in itself, integrating the ingestion of data analyses (themselves based on multiple networks of observations), the assimilation of these observations into the model integration cycle, and the output of variables of interest. This layer is the domain of earth system scientists and experts in numerical computation. GCMs are typically the result of the combined efforts of a large number of such scientists and engineers working over a long period of time.

In the **forecast generation layer**, forecast products are generated from dynamical model output. This process involves statistical corrections for model biases and may involve the integration of outputs from a number of different models. Decisions about which model outputs to use will be based on the analysis of model performance over an historical period. The resulting derived forecast products are typically stored in self-describing files with additional metadata to support the clients that deliver the outlooks. By storing the generated forecasts in an accessible, metadata rich format they are easily ingestible by downstream clients, whether these are simple viewers or more complex models that use the calibrated model data as one input among many. GCM model outputs may be used to drive other models (for example hydrological models). In general metadata is preserved as data is processed and new metadata added to describe transformations. This enables downstream users of the data to understand its provenance. Metadata curation, while tedious, should be considered a best practice if data is to be made public and its use promoted. Practitioners in this domain will typically be data scientists, statisticians, and climate scientists who work closely with forecast users.

At the **data service layer** forecast data is exposed via a data server, which makes the forecast data available using standard interfaces such as OPEnDAP[35]. At this stage, the generated outlooks are data products, not graphical products. The format of the output is not dependent on the particular dynamical model, or even that the model is dynamical: the forecast is simply a time series of gridded data with descriptive metadata. This layer is the domain of information architects and software engineers with expertise in moving data across networks efficiently.

The **product service layer** provides the means for the majority of forecast users to access the products they require, typically in the form of maps and graphs presented as images, data tables and expert commentary. Such products are developed by climate scientists and associated professionals with expertise in data visualisation, usually in close consultation with forecast users. This layer may take the form of pre-generated images and tables, or of complex applications that obtain and process data directly from the data service layer using web services.

The use of open standards, interoperable systems and simple, clean interfaces simplify the challenge of integrating data from multiple streams into usable seasonal forecasts. Systems need to be interoperable to reduce the cost of exporting and ingesting data, a procedure which is required at all stages of the process from the modelling (where analyses must be ingested for the initialisation of models) to the product services (where potentially large volumes of image and web page requests must be serviced). The use of open standards supports this interoperability. Open standards arise in communities of practice over periods of time, and generally become enshrined in documentation and formally supported by inter-institutional bodies. They are to be preferred over the creation of new *ad hoc* formats and interfaces. Clean interfaces means that coupling between system modules should be kept to a minimum, and that system modules communicate with each other as far as possible using the standards described above. The integration of model outputs into arbitrary decision-support systems and downstream models is supported by providing the model output, and post-processed model-based forecasts in standard formats, and exposing web services that provide access to data and meta-data via clearly defined protocols.

6.1. Designing seasonal outlook products and tools

Agile software development methodologies allow technical development to proceed simultaneously with the gathering of user feedback and refinement of designs. They are characterised by short development cycles with clearly defined goals and sub goals. Beginning development early ensures technical issues are solved, avoiding delays if system requirements cannot be completely specified in advance, or scientific results that underpin the forecast products cannot be anticipated. More traditional software development lifecycles, such as the so-called 'waterfall' model, depend on system requirements and features being specified early and held static throughout the development period.

In the agile model of software development, regular user testing takes place at each development increment with user feedback incorporated into the next iteration. A test system may be made available for the use of developers and other project team members. An agile approach suits small and specialised project teams, as the flexibility of the agile approach may hold management risks for a larger development group. This approach enables responsiveness to the requirements of the end users of the system.

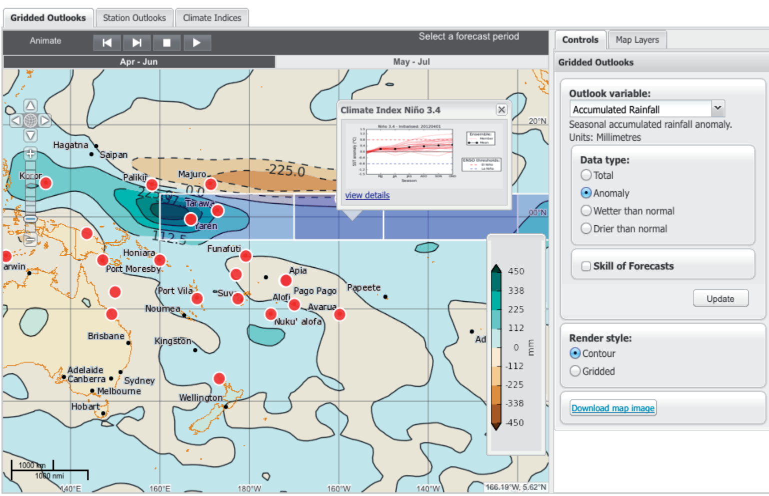


Figure 9. An example of a web-based tool providing seasonal climate forecasts.

6.2. A case study: The pacific seasonal forecast portal

The development of web based tools integrating model-based outlooks with climatological information and other contextual information is one means of communicating information about climate risk to end users. [36]

A system was developed to ingest output data from the Predictive Atmospheric-Ocean Model for Australia (POAMA) GCM [17]. A user-facing component was developed, based on a rich web-based interface that provides a one-stop shop for access to dynamical model-based outlooks. The purpose of the tool is to provide a specialised point of access to CGCM based seasonal outlooks for the national meteorological services of Pacific Island countries.

This project was supported by the Pacific Adaptation Strategy Assistance Program (PASAP), a component of the International Climate Change Adaptation Initiative - an Australian Government Initiative of $328 million over five years, 2008-2013 to assist with high priority climate adaptation needs in vulnerable countries in the Asia-Pacific region. As part of this program, the Australian Bureau of Meteorology lead a project to strengthen climate prediction capacities in the national meteorological and hydrological services of Pacific Island countries, including countries both north and south of the equator: Papua New Guinea, Tuvalu, Kiribati, Fiji, Marshall Islands, Federated States of Micronesia, Palau, Nauru, Cook Islands, Samoa, Tonga, Niue, Solomon Islands, Vanuatu and East Timor. A key element of this work was the development of a web-based application providing access to dynamical model-based seasonal outlooks. As previously described, one means to reduce vulnerability to climate change is by improve preparedness to anomalous climatic events.

Graphical displays of seasonal forecasts of broad-scale, point and climate driver forecasts are generated with an example shown in Figure 9. The web application displays the contextual information provided as meta-data by the data service layer, consumes the outputs of web services that produce figures and tables. It displays model-based outlooks as overlays on dynamical maps using geospatial web services. Access is given not just to application graphics but also to outlook data. User-friendly options for data extraction from the web portal are provided to support users of the range of tools from Excel to R.

An agile, iterative approach to the development of the web portal user interface (UI) included testing of early development versions of the portal with users at a project workshop, and in a series of country visits. These sessions validated the overall UI design and provided valuable feedback for improvements.

While much work went into the web front end, an equal amount of work was spent ensuring that the forecast generation layer is decoupled from this specific client. The access to the data provided by the data service layer allows for the future design of web clients that perform computational value adding using processing services, for example the ingestion and subsequent combination and calibration of multiple selected models.

The integration of data into a dynamical mapping tool provides opportunities for data mash-up in which data from different sources is displayed in composite. The provision of geospatial information in such a way that data from multiple sources can be integrated opens the way for new and interesting applications. For example, one potential future application for seasonal forecasting might be the display of agricultural or fishery yield data overlaid with outlook reliability data.

Over the course of the project several workshops were held bringing representatives of partner countries together with scientists and service developers. These workshops provided training in the use and interpretation of dynamical seasonal predictions and introduced the software tools developed to provide the seasonal outlooks. Communication and training are essential elements in the development of outlook products: producing well-calibrated outlooks is not effective unless the end users are equipped to use them correctly. A particular challenge is that of communicating seasonal forecasts that are couched in terms of probabilities, and which may differ from model to model. Both formal and informal user assessment of outlooks must be carried out to ensuring that what the forecast provider thinks is being communicated is what is being communicated. An iterative design process in which users are consulted early and often reduces the risk of miscommunication, and allows for learning to proceed over a period of time.

In workshops scientists presented lectures on the physical basis of seasonal predictability; the historical skill of the POAMA GCM in predicting ocean and atmospheric conditions across the Pacific; software tools developed to provide access to the latest seasonal forecasts based on the coupled models. In one workshop participants also engaged in a series of exercises using the portal to generate seasonal outlooks for their local region which they described to the group in a series of successful presentations. Such hands on exercises are highly effective at developing skills in using seasonal forecasts and associated tools, in assessing the knowledge and level of engagement of participants, and in testing whether the tool works properly under realistic conditions.

In discussions participants highlighted the importance of climate studies focused on improving the understanding of climate variability in Pacific Island Countries, noting that climate variability interacts with climate change leading to many of the first felt impacts of climate change. Improved knowledge of extreme climatic events, with the assistance of tailored forecast tools, will help enhance the resilience and adaptive capacity of communities affected by climate variability and change.

7. Conclusion

From the physical basis to the complexities of applications to specific industries and decisions it is clear that seasonal prediction is a large-scale enterprise requiring coordinated work across a range of scientific and technological disciplines. Steady improvements in GCM resolution and physics, coupled with ever increasing understanding of the physical mechanisms of predictability, will ensure that seasonal predictions become an important component of adaptation to a changing and more variable climate.

Author details

Andrew Charles*, Yuriy Kuleshov and David Jones
National Climate Centre, Bureau of Meteorology, Australian Government

* Corresponding Author

Acknowledgement

The Australian Agency for International Development (AusAID) and the Australian Government Department of Climate Change and Energy Efficiency (DCCEE) provided support for the PASAP project. The Water Information Research and Development Alliance (WIRADA) supported work into analogue downscaling methods. Considerable thanks are due to Kay Shelton, Andrew Cottril, Roald de Wit, Salesa Kaniaha, Q.J. Wang, David Robertson, Andrew Watkins, Oscar Alves, Claire Spillman, Debra Hudson and Harry Hendon for contributions to the work outlined here.

8. References

[1] Abawi Y, Llanso P, Harrison M, Mason SJ. Water, Health and Early Warnings. In: Troccoli A, Harrison M, Anderson DLT, Mason SJ. (eds.) Seasonal Climate: Forecasting and Managing Risk. Dordrecht, Springer; 2008. p351–95.

[2] McPhaden MJ, Busalacchi AJ, Cheney R, Donguy J-R, Gage KS, Halpern D, et al. The Tropical Ocean-Global Atmosphere observing system: A decade of progress. J. Geophys. Res. 1998;29;103(C7): 14,169–14,240.

[3] Janowiak JE, Xie P. CAMS–OPI: A Global Satellite–Rain Gauge Merged Product for Real-Time Precipitation Monitoring Applications. Journal of Climate. 1999;12(11):3335–42.

[4] Solow A, Nicholls N. The Relationship between the Southern Oscillation and Tropical Cyclone Frequency in the Australian Region. Journal of Climate. 1990;3(10):1097–101.

[5] Rahmstorf S, Coumou D. Increase of Extreme Events in a Warming World. PNAS. 2011 1;108(44):17905–9.

[6] Hansen J, Sato M, Ruedy R. Public Perception of Climate Change and the New Climate Dice. arXiv:1204.1286 [2012 5 (accessed 2 July 2012). http://arxiv.org/abs/1204.1286

[7] Corti S, Molteni F, Palmer TN. Signature of recent climate change in frequencies of natural atmospheric circulation regimes. Nature. 1999;29;398(6730):799–802.

[8] Palmer TN. A Nonlinear Dynamical Perspective on Climate Prediction. Journal of Climate. 1999;12(2):575–91.

[9] Huntington TG. Evidence for intensification of the global water cycle: Review and synthesis. Journal of Hydrology. 2006 15;319(1–4):83–95.

[10] Nicholls N. President's Column: Climate change and the weather. Bulletin of the Australian Meteorological and Oceanographic Society. 2011;24:53–4.

[11] Spillman CM, Alves O, Hudson DA. Seasonal Prediction of Thermal Stress Accumulation for Coral Bleaching in the Tropical Oceans. Monthly Weather Review. 2011;139(2):317–31.

[12] McIntosh PC, Pook MJ, Risbey JS, Lisson SN, Rebbeck M. Seasonal climate forecasts for agriculture: Towards better understanding and value. Field Crops Research. 2007;104(1–3):130–8.

[13] Drosdowsky W, Chambers LE. Near-Global Sea Surface Temperature Anomalies as Predictors of Australian Seasonal Rainfall. J. Climate. 2001;14(7):1677–87.

[14] Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, et al., (eds.) IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, United Kingdom and New York NY, USA.

[15] National Research Council NRC. Assessment of Intraseasonal to Interannual Climate Prediction and Predictability [Internet]. The National Academies Press; 2010. Available from: http://www.nap.edu/openbook.php?record_id=12878 (accessed 21 April 2012).

[16] Bourke W. An Efficient, One-Level, Primitive-Equation Spectral Model. Monthly Weather Review. 1972;100(9):683–9.

[17] Wang G, Hudson D, Alves O, Hendon H, Liu G, Tseitkin F. SST skill assessment from the new POAMA-1.5 System. BMRC Research Letters. 2008;8:1–6.

[18] Palmer TN, Råisånen J. Quantifying the risk of extreme seasonal precipitation events in a changing climate. Nature. 2002 31;415(6871):512–4.

[19] Charles A, Hudson D, Alves O. Combining verification with probabilistic forecasts for decision making. In Anderssen, R.S., R.D. Braddock and L.T.H. Newham (eds) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009. Cairns: p. 4184–90. Available from: http://www.mssanz.org.au/modsim09/J1/charles_J1.pdf

[20] Lorenz EN. Deterministic Nonperiodic Flow. Journal of the Atmospheric Sciences. 1963;20(2):130–41.

[21] Palmer T, Hagedorn R. Predictability of weather and climate. Cambridge University Press; 2006.

[22] Jolliffe IT, Stephenson DB. Forecast Verification: A Practitioner's Guide in Atmospheric Science. Wiley; 2003.

[23] Jones D, Weymouth G. An Australian monthly rainfall dataset. 1997 p. 19. Report No.: 70.

[24] Liljas E, Murphy AH. Anders Ångström and His Early Papers on Probability Forecasting and the Use/Value of Weather Forecasts. Bulletin of the American Meteorological Society. 1994;75(7):1227–36.

[25] Glahn HR, Lowry DA. The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. Journal of Applied Meteorology. 1972;11(8):1203–11.

[26] Wilks DS, Hamill TM. Comparison of Ensemble-MOS Methods Using GFS Reforecasts. Monthly Weather Review. 2007;135(6):2379–90.

[27] Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA. Current approaches to seasonal to interannual climate predictions. International Journal of Climatology. 2001 1;21(9):1111–52.

[28] Coelho CAS, Pezzulli S, Balmaseda M, Doblas-Reyes FJ, Stephenson DB. Forecast Calibration and Combination: A Simple Bayesian Approach for ENSO. J. Climate. 2004;17(7):1504–16.

[29] Palmer TN, Doblas-Reyes FJ, Weisheimer A, Rodwell MJ. Toward Seamless Prediction: Calibration of Climate Change Projections Using Seasonal Forecasts. Bulletin of the American Meteorological Society. 2008;89(4):459–70.

[30] Murphy AH, Winkler RL. A General Framework for Forecast Verification. Monthly Weather Review. 1987 1;115(7):1330–8.

[31] Roulston MS, Smith LA. Evaluating Probabilistic Forecasts Using Information Theory. Monthly Weather Review. 2002 1;130(6):1653–60.

[32] Wang QJ, Robertson DE, Chiew FHS. A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. Water Resour. Res. 2009 9;45:18 PP.

[33] Timbal B, Jones D. Future projections of winter rainfall in southeast Australia using a statistical downscaling technique. Climatic Change. 2008;1;86(1):165–87.

[34] Charles A, McClymont D, de Wit R, Jones D. A software architecture for seasonal climate forecasts in the Tropical Pacific. In Chan, F., Marinova, D. and Anderssen, R.S. (eds.) MODSIM2011: 19th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011. p.1188–94. http://www.mssanz.org.au/modsim2011/C4/charles.pdf

[35] Cornillon P, Gallagher J, Sgouros T. OPeNDAP: Accessing data in a distributed, heterogeneous environment. Data Science Journal. 2003;2:164–74.

[36] Kuleshov Y, Jones D, Hendon HH, Charles A, Cottril A, Lim E, et al. Pacific Adaptation Strategy Assistance Program: Strengthening the capacity for seasonal prediction services in Pacific Countries. Bulletin of the Australian Meteorological and Oceanographic Society. 2012;25(1).