PROBABILISTIC SAFETY ANALYSIS OF DAMS
Methods and Applications

By

Negede Abate Kassa

Dissertation

submitted to Faculty of Civil Engineering
of Technische Universität Dresden
in partial fulfillment of the requirements
for the degree of

Doctor of Engineering (**Dr.-Ing.**)

TECHNISCHE UNIVERSITÄT DRESDEN

Dresden, Germany 2009

Rigorosum, Demonstration und Disputation: 29.04.2010

Vorsitzender der Promotionskommission:
Prof. Dr.-Ing. Peer Haller (TU Dresden)

Gutachter:
Univ.-Prof. (em.) Dr.-Ing. habil. Hans-B. Horlacher (TU Dresden)
Prof. Dr.-Ing. Jürgen Jensen (Universität Siegen)

PROBABILISTISCHE SICHERHEITSANALYSE VON DÄMMEN
Methoden und Anwendungen

Dissertation

vorgelegt von

Negede Abate Kassa, M.Sc.

an der Fakultät Bauingenieurwesen
der Technischen Universität Dresden
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)

TECHNISCHE UNIVERSITÄT DRESDEN

Dresden, Germany 2009

I dedicate this dissertation to my wife Abrehet.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Univ.-Prof. Dr.-Ing. habil. Hans-B. Horlacher for giving me the opportunity to do this PhD under his supervision, for his invaluable advices and continuous supports of all kind throughout the research period. It would not have been possible for me to complete this PhD with out his resourcefulness and help – I owe you a big thank you Prof. Horlacher! My thanks also go to my second supervisor Prof. Dr.-Ing. Jürgen Jensen for volunteering to give his expertise as a co-supervisor and for his constructive comments. Thanks are also due to members of the PhD examination committee for offering their expertise and time.

This dissertation is a result of the research work that was carried out mainly at the Institute of Hydraulic Engineering/Institut für Wasserbau und Technische Hydromechanik (IWD), Faculty of Civil Engineering/Fakultät Bauingieurwesen of Dresden University of Technology/Technische Unibesrsität Dresden (TUD), where I have benefited a lot from the inspiring and friendly atmosphere. Thanks are due to all members of the Institute. I, especially, acknowledge the support I got from Dipl.-Ing. Torsten Heyer, Dipl.-Ing. Herbert Martin, Dipl.-Ing. Holger Haufe, Dipl.-Ing. Sophia Stoebenau, Mrs. Kerstin Winkel and Mrs. Carola Luckner.

This research work is financially supported mainly from The United Nations Development Program Second Country Co-operation Framework (UNDP-CCF2) project under the framework of the support to Ministry of Water Resources (MoWR), Ethiopia. The field work got financial assistance from Gessellschaft für Technische Zusammenarbeit/German Technical Cooperation (GTZ). Financial support for attending conferences was availed by IWD-TUD. Moreover, during the last year of the study funding was possible through a research assistant employment I got at IWD-TUD. The funding from all sources is deeply acknowledged.

Data used for the case study of this research were obtained from National Meteorological Service and MoWR, Ethiopia. Their co-operation is deeply acknowledged. I am grateful to Ato Girma Mekonnen and Ato Abreham Assefa for hosting and supporting the laboratory work in the Construction Design S.Co. Materials Testing and Foundation Investigation Laboratory. Thanks are due to Wt. Yodit Ayallew, Ato Dercho Duelo, Ato Melaku H\michael, and Ato Yeshiyanew Alemu for their help in field and at Arba Minch University soil mechanics laboratory.

Especially, I cannot say enough thank you to my hero and love- my wife Abrehet, for her continuous support and encouragement all along. My gratitude goes also for her help in reviewing the dissertation draft and to her useful advice on mathematics and computing. Thanks Abisho! My lovely daughter Liyou you are my inspiration. I am sure this will double with Ruh's birth. My parents have always a special place in my life. I cannot thank you enough Bitish and Abate for being good and loving parents to me. If I possess any good qualities, it is all because of you. My brothers and sister, I owe you a big thank you for your supports and love all along.

I should not finish without expressing my gratitude to my Ethiopian and German friends in Dresden (Anteneh, Taddele, Bahlbi, Mesfin, Doreen, Nigussie, Bini, Dr. Webtaye, Kelemwa, Alemayehu, Alemash, Falk, Bederu and Tsedeke). So much fun and good times I have shared with you. Thank you for all your supports and comforts. You made my stays in Dresden lively.

Negede Abate, August 2009.

PREFACE

The safety of dam is inseparable from its foundation performance, its spillway capacity and hydraulic characteristics and its integrity against seepage and destructive floods. In the past, when sites were favorable and resources plentiful, the shortcomings of science were overcome by generous budgets. Massive dams stand today as monuments to that era. Many of those works may outlive the present civilization, surviving probable maximum floods, maximum earthquakes, and an inevitable measure of neglect. They were sound investments in their time. Today, greater emphasis must be placed upon the economical use of resources. Design criteria and rehabilitation decisions must be scrutinized to eliminate excess use of resources whilst the provision of adequate safety is assured. Enhancing dam safety increases investment costs and its decrease boost potential risk, thus some balance must be found between the two conflicting interests (economy and safety). It is necessary to determine levels of acceptable risk for different sets of conditions; criterion for the design of proposed dams and rehabilitation of old dams must be objectively justified. This calls for analytical risk and dam safety analysis.

The safety evaluation of aging dams and risk-based (probabilistic) design of new dams is getting increased attention more than ever before. The public and decision makers now demand transparency and accuracy in design decisions regarding safety issues of dams. Engineers are more and more required to explicitly quantify risk-levels associated with designed dimensions of new dams and upgrading or rehabilitation recommendations of old dams. They are required to quantify how safe a dam is and how well the balance between safety and economy is kept; as specified by engineering, societal, environmental, political and economical standards or regulations. It is required to base such evaluations on numerical justifications derived from transparent procedures that are well founded on theory than subjective judgment. However, this is a complex undertaking because of uncertainties associated with parameters involved in dam design; such as uncertainty in occurrence of loads and hazards, material property, geology, models and boundary conditions, operators' inputs, maintenance etc. Most physical and operational variables in dam design are uncertain parameters. Nominal magnitudes assumed in design are likely to differ in time and/or space and thus performance (capacity and safety) of dams is too an uncertain parameter.

Traditionally uncertainty in design parameters is assumed to be accounted for through the use of safety factors. By assigning generous safety factors dam performance (capacity and safety) is assumed to remain in an acceptable range and it is customarily presumed that dams never fail. Nowadays, in most modern engineering codes and society, the appropriateness of the safety factor approach is being questioned. This approach does not allow for transparent accounting of uncertainties and for numerical quantification of safety. It does not permit for optimizing safety and economy uniformly across the system by associating a quantified level of risk to alternative design options. Sometimes, if design is complex, safety factors can compound to cause over design in certain parts of dams safety chain, with still one or more weak links and uncertain reliability left unseen in other elements of the safety chain, ultimately living the system only as safe as the weakest element in the chain. The safety factor method is only capable of giving qualitative measurement. If engineers want to deal with variability of design parameters head-on and attempt to give quantitative evaluation of performance, they need to use more information in their calculations than mean or nominal values of design parameters. This requires a shift from

classical deterministic design approaches to modern probabilistic approaches with an integrated process that affords explicit recognition and treatment of uncertainties.

Nonetheless, though it is believed that probabilistic approaches give enhancement to the traditional dam safety evaluation concept it is not extensively practiced. That is because techniques, tools, and standards that support the implementation of probabilistic methods are still at rudimentary stage. Moreover, the profession is not provided with sufficient examples demonstrating how a coherent whole exists in the application of probabilistic methods in dam design and safety analysis.

The subject of this Ph.D. dissertation is the determination of failure probabilities of dams, which are fundamental ingredients in dam safety analysis, through the use of a multitude of rigorous probabilistic and mathematical approaches and updated design procedures. The research addresses techniques of doing probabilistic design and safety analysis. It conceives a variety of approaches for customizing classical design equations, which are originally set for deterministic design applications, so that they suit the intended probabilistic dam safety analysis. It provides codes and tools for doing the necessary computations on the major dam failure mechanisms. It suggests implementation architectures and techniques, and demonstrates applications of engineering, mathematical software and supporting technologies.

The research basically tested the applicability of three classical probabilistic methods- Monte Carlo Simulation Method (MCSM), First Order Second Moment (FOSM) and Second Order Second Moment (SOSM) and one new analytical method of quantifying failure probabilities (performance) from analytically transformed distributions. The latter method is called Analytical Solution for finding Derived Distributions (ASDD) method. It is new for its practical applications in dam engineering, although its underlining theories are known for quite a while in the domain of advanced probability theory. The method's basic principles are derived, integrated and presented in such a way that they are understandable by practicing engineers and are convenient for implementation. An implementation architecture that is easy to follow is prepared for all methods taking into account the dominant dam failure mechanisms. Using a variety of carefully selected practical case studies the practical applicability of all methods is demonstrated and their effectiveness is compared. It is shown that how a coherent whole exists in the proposed methodologies. It is tried to implement the methodologies in a context that exposes insufficiencies in existing deterministic design practices.

The major prepositions when starting this research was (1) the classical deterministic design procedure is poor as an option. It is not accurate and transparent. It does not adequately describe reality, it potentially obscure variability in design parameters and thus provide vague performance estimates. Using this method economic optimization is not possible. And thus generally it is unsuitable to 21st century societal requirements and standards. (2) Until the knowledge gained through research for dealing with uncertainties in design, risk and safety analysis of dams become adequate enough to permit good quantitative descriptions of performance and safety issues, design practices will remain in the realm of subjective judgments. This mostly result over designs due to compounding factor of safeties with still non-uniform reliability distribution in the system safety chain, i.e. with still one or more weak links in the safety chain that put the entire system only as safe as the weakest element in the chain. (3) Dam engineering as a profession is not adequately presented with demonstrated analytical and

probabilistic techniques, application tools and implementation strategies to meet the growing demand for transparency and precision of dam safety analysis procedures and related issues. (4) It is a thorny issue to analyze uncertainty in design parameters and integrate that in risk based design formulation. (5) The introduction of probabilistic approaches will not make design practices error proof and absolutely accurate. But, even a tiny improvement may help. (6) The risk-based dam design approach, which for decisions bases on finding economic optimum points between investment cost and cost of damage in case of failure, will have to go a long way to become an attractive approach when dealing with large dams where rupture would endanger a downstream population (especially if there is a large town or city for which "absolute safety" will be demanded- though no method guarantee this). But, it can be attractive approach when there is no risk of loss of life, as with dams in remote areas, near seas or lakes; or those small dams impounding small reservoirs; or those lower dams in wide valleys where rise in water level due to dam failure is not significantly larger when compared with the same flood occurring with out the dam; or when there is reliable warning and evacuation system for the downstream population. The later cases, those which make the approach attractive at its current level of detail and precision, are prevalent in most developing countries with scarce financial resources, countries like Ethiopia, where extensive dam construction projects are kicking off in remote areas and where a number of dam upgrading and rehabilitation projects are foreseen.

This research has managed to publish several papers whose bibliographic information is given in the references section.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

a.m.s.l	Above Mean Sea Level
AAU	Addis Ababa University
AMU	Arba Minch University
ANCOLD	Australian Committee on Large Dams
ARI	Average Recurrence Interval
ASDD	Analytical Solution for finding Derived (quality variable) Distributions
cdf	Cumulative Distribution Function (or simply probability distribution)
CDSco	Construction Design Share Company
CI	Confidence Interval
CLT	Central Limit Theorem
DEM	Digital Elevation Model
FEM	Finite Element Methods
FH	Flood Hydrograph
FOSM	First Order Second Moment Method
FRL	Full Retention Level
FRV	Function of Random Variable (Random Vector)
GSE	Geological Survey of Ethiopia
HPEM	Harr Point Estimate Method
ICE	Britain Institute of Civil Engineers
ICOLD	International Commission on Large Dams
LHS	Latin Hypercube Sampling
MAM	Moment Analysis Method
MCSM	Monte Carlo Simulation Method
MDDL	Minimum Drawdown Level
MoWR	Ministry of Water Resources, Ethiopia
MWL	Maximum allowed Water Level
NAMSA	National Meteorological Service Agency, Ethiopia
OMC	Optimum Moisture Content
pdf	Probability Density (distribution) Function
PE	Point Estimation method
PMF	Probable Maximum Flood
PMP	Probable Maximum Precipitation
RPEM	Rosenblueth Point Estimate Method
RV	Random Variables
SFH	Surface Flow Hydrograph (direct flow hydrograph)
SOSM	Second Order Second Moment Method
UH	Unit Hydrograph
USACE	US Army Corps of Engineers
WWDSE	Water Works Design and Supervision Enterprise, Ethiopia

S	strength or capacity of a system
L	load or hazard on a system
Z	reliability of a system ($Z = S\text{-}L$)
μ_Z	mean reliability of a system

σ_Z	standard deviation of reliability of a system
β	reliability index (μ_Z/σ_Z)
Z_s	reliability against sliding
Z_{OT}	reliability against overtopping
F	factor of safety
P_f	probability of failure
F_s	factor of safety against sliding
F_{OT}	factor of safety against overtopping
$P_{f,s}$	sliding failure probability
$P_{f,OT}$	overtopping probability
R	risk
D	damage

L(m)	length of failure surface (slip surface length, base length of an active soil mass)
$\bar{\tau}$ (kN/m^2)	apparent shear strength
$\bar{\tau}'$ (kN/m^2)	effective (true) shear strength
τ (kN/m)	effective shear strength per meter width of a slice of known length l_n
W(KN)	weight of an active soil mass
W_n(KN)	soil slice weight
$\overline{G}'$ (kN/m^2)	effective shear stress (gravitational driving force due to the component of an active soil mass of weight W above a failure surface of length L)
G (kN/m)	effective shear stress per meter width of a slice of known length l_n
u_n (kN/m^2)	pore water pressure at the base of a slice
U_n (kN/m)	pore water force at the base of a unit width soil slice of known length l_n
α_n (Degrees)	inclination of slice base
b_n (m)	slice width
l_n (m)	slice base length ($l = b/cos\alpha$)
τ (kN/m)	shear strength per meter width of a slice of length l
σ (kN/m^2)	total normal stress
σ' (kN/m^2)	effective normal stress
c' (kN/m^2)	effective (true) cohesion
c (kN/m^2)	apparent cohesion
ϕ' (degrees)	effective (true) angle of shearing resistance
ϕ (degrees)	apparent angle of shear resistance
C_v (m^2/s)	coefficient of consolidation
m_v (m^2/kN)	coefficient of volume compressibility
c_α	coefficient of secondary consolidation
PL (%)	plastic limit
LL (%)	liquid limit
PI (%)	plasticity index
MDD (gm/cm^3)	maximum dry density
γ (kN/m^3)	unit weight of soil
γ_w (kN/m^3)	unit weight of water
h_n (m)	average height of soil (slice)
h_w (m)	average height of water above dam surface
z (m)	depth of overburden ($h + h_w$)

n_f	flow channels
K (m/day)	permiability
n_d	number of equipotential drops
h_d (m)	height of dam or significant wave height as appropriate
H (m)	average head or reservoir depth above spillway crest, outlet intake or ground level as appropriate
h_o (m)	initial water level
h_f (m)	rise in water level due to flood routing
h_s (m)	wind setup
H_s (m)	significant wave height
h_r (m)	wave run up
h_i (m)	seiche
Q_b (m^3/s)	base flow
Q_P (m^3/s)	peak flood discharge
t_p (h)	time to peak
t_g (h)	time-to-centroid
F_e (km)	effective fetch
U (m/s)	Wind speed
T	return period
Q_{PMF} (m^3/s)	probable maximum flood discharge
$\boldsymbol{X}$	random vector (bold upper case letter)
X	random variable (regular upper case letter)
$x_1, x_2, \ldots, x_n$	realization of random vector $\boldsymbol{X}$ (bold lower case letters)
x	realization of random variable X (regular lower case letter)
$\overline{A}$	set A complement
$A \cup B$	union of event A and B
$A \cap B$ or AB	intersection of event A and B

ABSTRACT

Successful dam design endeavor involves generating technical solutions that can meet intended functional objectives and choosing the best one among the alternative technical solutions. The process of choosing the best among the alternative technical solutions depends on evaluation of design conformance with technical specifications and reliability standards (such as capacity, environmental, safety, social, political etc specifications). The process also involves evaluation on whether an optimal balance is set between safety and economy. The process of evaluating alternative design solutions requires generating a quantitative expression for lifetime performance and safety. An objective and numerical evaluation of lifetime performance and safety of dams is an essential but complex undertaking. Its domain involves much uncertainty (uncertainty in loads, hazards, strength parameters, boundary conditions, models and dam failure consequences) all of which should be characterized. Arguably uncertainty models and risk analysis provide the most complete characterization of dam performance and safety issues. Risk is a combined measure of the probability and severity of an adverse effect (functional and/or structural failure), and is often estimated by the product of the probability of the adverse event occurring and the expected consequences. Thus, risk analysis requires (1) determination of failure probabilities. (2) probabilistic estimation of consequences.

Nonetheless, there is no adequately demonstrated, satisfactorily comprehensive and precise method for explicit treatment and integration of all uncertainties in variables of dam design and risk analysis. Therefore, there is a need for evaluating existing uncertainty models for their applicability, to see knowledge and realization gaps, to drive or adopt new approaches and tools and to adequately demonstrate their practicability by using real life case studies. This is required not only for hopefully improving the performance and safety evaluation process accuracy but also for getting better acceptance of the probabilistic approaches by those who took deterministic design based research and engineering practices as their life time career. These problems have motivated the initiation of this research.

In this research the following have been accomplished:

(1) Identified various ways of analyzing and representing uncertainty in dam design parameters pertinent to three dominant dam failure causes (sliding, overtopping and seepage), and tested a suite of stochastic models capable of capturing design parameters uncertainty to better facilitate evaluation of failure probabilities;
(2) Studied three classical stochastic models: Monte Carlo Simulation Method (MCSM), First Order Second Moment (FOSM) and Second Order Second Moment (SOSM), and applied them for modeling dam performance and for evaluating failure probabilities in line with the above mentioned dominant dam failure causes;
(3) Presented an exact new for the purpose analytical method of transforming design parameters distributions to a distribution representing dam performance (Analytical Solution for finding Derived Distributions (ASDD) method). Laid out proves of its basic principles, prepared a generic implementation architecture and demonstrated its applicability for the three failure modes using a real life case study data;

(4) Presented a multitude of tailor-made reliability equations and solution procedures that will enable the implementations of the above stochastic and analytical methods for failure probability evaluation;
(5) Implemented the stochastic and analytical methods using real life data pertinent to the three failure mechanisms from Tendaho Dam, Ethiopia. Compared the performance of the various stochastic and analytical methods with each other and with the classical deterministic design approach; and
(6) Provided solution procedures, implementation architectures, and Mathematica 5.2, Crystal Ball 7 and spreadsheet based tools for doing the above mentioned analysis.

The results indicate that:

(1) The proposed approaches provide a valid set of procedures, internally consistent logic and produce more realistic solutions. Using the approaches engineers could design dams to meet a quantified level of performance (volume of failure) and could set a balance between safety and economy;
(2) The research is assumed to bridge the gap between the available probability theories in one hand and the suffering distribution problems in dam safety evaluation on the other;
(3) Out of the suite of stochastic approaches studied the ASDD method out perform the classical methods (MCSM, FOSM and SOSM methods) by its theoretical foundation, accuracy and reproducibility. However, when compared with deterministic approach, each of the stochastic approaches provides valid set of procedures, consistent logic and they gave more realistic solution. Nonetheless, it is good practice to compare results from the proposed probabilistic approaches;
(4) The different tailor-made reliability equations and solution approaches followed are proved to work for stochastic safety evaluation of dams; and
(5) The research drawn from some important conclusions and lessons, in relation to stochastic safety analysis of dams against the three dominant failure mechanisms, are.

The end result of the study should provide dam engineers and decision makers with perspectives, methodologies, techniques and tools that help them better understand dam safety related issues and enable them to conduct quantitative safety analysis and thus make intelligent dam design, upgrading and rehabilitation decisions.

1 INTRODUCTION

Dam design is preoccupied with generating alternative technical solutions that can meet intended functional objectives. The process of choosing among potential technical alternatives requires not only a detailed analysis of technological aspects but also critical evaluation of design conformance against specifications (environmental, economic, social, political criteria etc). Moreover, assessments have to be made on whether there is conformance with safety standards and whether an optimal balance is set between safety and economy. Such an assessment is vital in design of proposed dams and in making intelligent dam upgrading and rehabilitation decisions. It also helps in explaining problems encountered during dam operation.

The value of each design parameter involved in dam design, upgrading or rehabilitation studies is likely to differ in time and space from the nominal magnitudes used in deterministic designs. The differences, among many other things, are caused by the inherent uncertain nature of design parameters (loads, hazards, material properties, geometry, construction practices, and boundary conditions) and by the time dependent nature of structural reliability (deterioration). Consequently, performance of dams at any time is directly affected by the uncertainty associated with variables listed above, which can broadly be categorized as uncertainty in loads and resistance mechanisms. Design and analysis models adequacy and subsequent operation and maintenance practices will have an additional role by influencing either of these groups.

Traditionally, uncertainties are assumed to be accounted for through the use of global safety factors. Based on this assumption it is taken for granted that performance of designed dams will remain in an acceptable range and they never fail (or are "adequately safe"). Nowadays, in most modern engineering codes and society, the appropriateness of such black box safety factors (factor of ignorance), and thus the resulting fuzzy safety of dams, are being questioned. The safety factor based approach does not allow for transparent accounting of the various inherent uncertainties and for assuring uniform distribution of safety across the system. Moreover, it does not allow for optimizing safety and economy through associating a quantified level of risk to alternative design options.

To work and think naturally, by acknowledging variability of parameters, uncertainties in loads, hazards, strength parameters, boundary conditions, models and dam failure consequences should be characterized. Furthermore, instead of assuming "absolute safety" based on phony and incomplete premises, randomness in performance and safety of dams shall be quantitatively characterized. Uncertainty models and risk analysis provide the most complete characterization of dam performance and safety issues. The most highly developed uncertainty model to date is probabilistic uncertainty model (Möller, 2004). However, when parameters are characterized by informal uncertainty (e.g. when only a small number of observations are available, when boundary conditions are subject to arbitrary fluctuations, when measurement and observations are coarse or when system overview is incomplete) or when they are characterized by vague uncertainty (e.g. uncertainty quantified using linguistic variables, experts statement, for example, phrases of type, high flow, low flow) the use of fuzzy set theory based uncertainty models might be more suitable. Möller (2004) wrote on the state of development of uncertainty models. Topics of uncertainty modeling are also discussed and illustrated in (Ayyub, 1998) and (Elishakoff,

1999). The current research mainly utilizes probabilistic uncertainty models in combination with advanced probability theories and mathematical statistics.

1.1 Background

Most physical variables in dam design (hydrologic, hydraulic, material, geotechnical, geometry etc variables) are in fact random variables. Consequently, performance (capacity and safety) of dams is too random parameter. Standard design calculations are calculations based on nominal values of design variables. This kind of standard approach that uses mean or nominal values, which has been and still is in use by engineers, is called deterministic design approach. Deterministic design approach assumes that dams shall not fail and shall stay under top operating conditions throughout their design life. In deterministic design approach either no significant attention is given to variability of design parameters or some factors of ignorance (factor of safety) is used to keep the uncertainties out of sight. In addition, no significant attention is given to deterioration and variations of performance with time, and little effort is made towards integrating inputs of maintenance and operation in design.

As mentioned above, classical deterministic design approaches apply safety factors to allow for uncertainties. Because such methods are not based on transparent accounting of parameters interaction and associated uncertainty within failure mechanisms framework, it could not allow for uniform distribution of safety across the system. In some important cases, where there is an upper and lower specification for functional limits, safety factor methods can not be used at all. Such methods are liable to resulting either of two unwanted proceedings: either safety factors can compound to cause over design with unevenly distributed reliability especially when design is complex or, on the contrary, when little consideration is given to desired reliability of hydraulic systems and when too much is expected of them, while being ignorant of parameters uncertainties and when they are left unquantified, there might be under designs against one or more of failure mechanisms resulting devastating failures. History thought us in a very horrible way that this is a possibility and the consequences of uncertainty in hydraulic systems can be very costly and often tragic. Examples of the August 2002 Elbe River flooding around Dresden, Germany and the most recent Hurricane Katrina of the US Gulf Cost (29 August 2005) are enough to mention. The officially registered total damage cost of the Dresden flooding is estimated to be around 11.3 Billion Euros (IKSE 2004). The Hurricane Katrina failures has resulted an estimated total economic damage cost of 200 Billion USD and 100 Billion of the damage cost is left to the Gulf Cost (Wikipidia 2009). The two damages claimed lives of 18 and over 1,130 people respectively (CNN 2005).

Design criteria and rehabilitation decisions must thus be scrutinized to provide adequate safety and to eliminate excess. Enhancing dam safety increases costs, and some balance must be found between the two. Therefore, it is necessary to determine levels of acceptable risk in dam design, upgrading and rehabilitation endeavors. Criterion in design of proposed dams and in rehabilitation or upgrading of old dams must be quantitatively justified. This demand for analytical risk and dam safety analysis methods, which use more information in their calculation algorithms than mean or nominal values of design parameters, and that are capable of capturing uncertainties. This means a shift from classical deterministic approach to modern probabilistic design approaches.

Probabilistic Design method studies how to make calculations with probability distribution functions (pdf) of design parameters, instead of nominal or mean values only. This will then allow engineers to design systems that meet specific standard conformance and hence optimize safety and economy. In other words structures can be designed for a specific failure probability (probability of functional or structural failure) so that they face a pre-defined volume of failure.

If design parameters variability in any system are measured; like hydraulic, hydrologic, geotechnical, environmental, structural etc systems, data on the frequency of occurrence of the values of the parameters will be found. If there are sufficient data values the frequency can be rescaled to give probabilities (probability distributions). All design parameters may thus be viewed as random variables. Seeing flow magnitudes in hydraulic systems design as random variables is a common practice. That is partly because data on flow magnitudes could easily be generated or collected and their variability is apparent. But, taking other design parameters than flow, even other flow characteristics than flow magnitude, as random variables is not frequently seen phenomenon. Even the statistically evaluated peak flow magnitudes are usually used with deterministic design settings and standards.

During probabilistic design, dealing with variability (uncertainty) comprises the core of the design engineering work content. However, refined skills, techniques, tools and standards that are required to effectively do the job are currently rudimentary. As a result, probabilistic design, risk and safety analysis is not widely practiced in dam design, upgrading and rehabilitation studies. There is a growing awareness that the lack of such refined inputs is greatly impacting system performance, economics, risk and safety.

Most of the stages in probabilistic design, risk and safety analysis process still need further researching. Among the inadequacies preventing a wider application of the method are:

- Determination of pdfs of design parameters is not a trivial exercise. Often there is little information about design parameters pdfs (which most likely has not been built yet). In this regard, beside the common statistical methods, some fascinating techniques could be applied, like the use of fuzzy logic to arrive at conclusion based on vague, ambiguous, imprecise, noisy or missing input information. Advanced topics on methods of describing design parameters uncertainty as fuzzy and fuzzy random variables, together with remarkably advanced state-of-the art applications, are provided in (Möller 2004);
- If it is required to do better than MCSM, computation of pdfs of quality variables that are functions of multiple random variables is difficult. Quality variables imply geometry, dimension, strength capacity, performance, reliability etc characteristics of dams or their components which are determined by interactions of random design parameters. Sometimes just doing computations with moments, other times use of MCSM might be satisfactory. But, to get a more comprehensive, transparent, theoretically justifiable and accurate results, there is a need for getting theoretically founded fully probabilistic design approach that is capable of analytically driving quality variables pdf from design parameters pdfs and their functional relation;
- Development of stochastic data models and quantitative risk assessment techniques, which including analytical representation of failure mechanisms (development of customized

reliability functions and solution approaches) and probabilistic damage assessment, are areas that require further research and development of extensive analytical techniques;
- Calculation of risk costs (especially for risks involving non-monetary quantifiable damage, like lose of life) is a researchable area;
- Methods of doing cost optimal design decisions that impact different stakeholders under uncertainty are still notorious;
- Determination of influence of maintenance and dosing optimal inspection and repair intervals with an intent of integrating maintenance in design requires development of deterioration and method for modeling maintenance, which is also a researchable area;
- The uncertainty of probabilistic design approach by itself. i.e. effect of the models' uncertainty on its prediction capability is something that requires extensive demonstrations to clear doubts and convince rather firmly rooted deterministic mind sets;
- Computer methods development and demonstration for implementation of more rigorous analytical methods is also an absent requirement.

Therefore, the PhD research attempted to address selected gaps among the inadequacies listed above. Through the use of a multitude of rigorous probabilistic and mathematical approaches and updated design procedures it attempts to conceive a more transparent, comprehensive and precise approaches for calculating failure probabilities of dams, which are fundamental ingredients in dam safety analysis. The research addresses techniques of doing probabilistic design and safety analysis of dams using stochastic, numerical and analytical methods. It conceives a variety of approaches for customizing classical design equations, which are originally set for deterministic design applications, so that they suit the intended probabilistic analysis. It compares results from the different stochastic, analytical and deterministic methods and discusses their performance. It provides codes and tools for doing the necessary computations on major dam failure mechanisms. It suggests implementation architectures and techniques, and demonstrates applications of engineering, mathematical software and supporting technologies. The scope, methodology and main research question of the PhD work are presented in the next two sections.

1.2 Scope and methodology of the research

This research attempts to present potential probabilistic methods and the associated mathematics and required techniques for computing failure probabilities. It presents probabilistic methods that help to think and work naturally (and correctly) with variability of variables in dam safety evaluation. It is tried to put probabilistic design, risk and safety analysis methods in a context that exposes insufficiencies in existing deterministic dam design principles and safety standards. Fundamental errors in much discussed safety aspects and safety analysis methods are pointed out.

Basically, computation of probability distribution of quality/performance variables and the ultimate determination of failure probabilities for alternate designs using four probabilistic methods is presented: FOSM and SOSM Methods, MCSM and Analytical ASDD method. In addition, the intermediate process of determination and representation of design parameters pdfs, implementation architectures of the probabilistic methods and tailor-made reliability equations

for computing dam performance (functional and/or structural failure probabilities) is presented. The entire process presented takes into account uncertainty in engineering parameters (uncertainty in load and resistance parameters).

However, the ability of dams to survive hazards and loads does not only depend on capacities provided for its components during design but also on certain other features that come later in its life cycle; such as construction, operation, surveillance and maintenance. For a properly designed dam to remain in safe condition some reliance has to be made on construction, operation, surveillance and maintenance. These factors are also uncertain parameters but they are not dealt with in this research. Moreover, although basic concepts related to the follow up risk and safety analysis and standards is presented, detailed probabilistic analysis of damages or consequences of failure is not in the scope of this research. This area is researchable field with equal breadth and depth as the problem of probabilistic evaluation of dam failures itself. Thus, the scope of the research is limited to analysis of failure probabilities due to interaction of uncertain strength and load design parameters.

The research will emphasize on firing up decisive thoughts towards achieving a full probabilistic design, risk and safety analysis procedures and probabilistic safety standards for dams. It focuses on showing and filling selected gaps and inadequacies in dam safety (failure probability) evaluation methods. Emphasis is given to demonstration of theories, procedures and techniques through practical application of approaches in embankment dam design. Moreover, application of computer methods and computational software in the proposed solution procedures is demonstrated.

A desire for clarity of principles has led to this dissertation being relatively concise. If a reference for further details on theories and methodologies is required the author and his publications can be consulted.

1.3 Problem statement and primary research objectives

From the discussions thus far what apparently follows are the main research question and the primary objectives of the research. The main research question boils down to the following statements:

"How could the lifecycle performance (service, safety and investment) of alternate dam designs against decisive failure mechanisms, for those systems without failure data (or with limited failure data), be predicted using (time sensitive) stochastic relations of design parameters (probabilistic/analytical methods, implementation techniques and supporting technologies)? And which design standards are appropriate (deterministic with assumed 'absolute safety' or risk based with quantified safety and minimized overall life cycle cost)?"

The primary objectives of the research, thus, are:

- To conceive a more transparent, comprehensive, and accurate method of calculating failure probabilities of dams. This will be done for selected dominant failure mechanisms, through provision of multitude of probabilistic/mathematical

approaches and updated design equations (reliability functions) and calculation procedures,
- To provide implementation architectures, tools and techniques for that will facilitate implementation of the proposed approaches,
- To demonstrate the proposed idea on real life case study, to compare results from the different approaches and draw conclusions and recommendations,
- To outline outlooks and future research ideas.

1.4 Organization of the thesis

Apart from the introduction in chapter 1, the dissertation is composed of the following parts.

The fundamental principles of probabilistic design and safety analysis of dams are laid in chapter 2. Towards the end of this chapter extensive review of related works is given.

Chapters 3 and 4 focuses on stochastic and analytical methods for accounting uncertainty in design parameters, with the ultimate intention of estimating performance randomness and calculating failure probabilities. In chapter 3 underlying principles of one statistical trial method (Monte Carlo Simulation Method-(MCSM) and details of other two approximate Moment Analysis Methods (MAM) - FOSM and SOSM methods and their implementation architectures are presented. In Chapter 4 underling principles, detailed derivations of principles and implementation architecture of a new and relatively precise Analytical method of finding Solutions for Derived Distributions - (ASDD method) is presented.

In chapter 6 and 7 the stochastic and analytical methods will be applied for calculating failure probabilities/performance of a case study dam. The case study considers three failure mechanisms (stability, seepage and overtopping). In all the case studies, tailor-made reliability equations, uncertainty modeling, analysis and representation methods and an integrated failure probability estimation approaches are presented. In each case, results from the different probabilistic methods are compared with each other and with that from deterministic method, and issues of accuracy, suitability, ease and meaning of computed probabilities are discussed. The salient features and data on the case study dam, together with broad probabilistic perspective of dam engineering, are provided in chapter 5.

For basic theories on statistics and probability theory Chapter 3, which is dealing with the MCSM and the MAMs, refers to appendix 9.1. Generally, reading appendix 9.1 first might facilitate the understanding of the chapters that follow. However, for a reader with sound knowledge on statistics and probability theory reading appendix 9.1 first is not a necessity. The coverage of underling statistical and mathematical theories is moved to the appendix in order not to distract such readers focus on practical application of the theories. Finally, in chapter 8 conclusions and some suggestions for future research are presented.

The thesis has appendices giving summary of some basic statistical and mathematical theories, list of case study data and demonstration of calculation procedures. In the main body of

the thesis reference to the different sections of the appendix will be made as appropriate. Moreover, a CD-ROM containing the soft copy of the data used, Mathematica codes written to solve the different problems and analysis results is attached. Particularly codes prepared for doing reservoir stochastic routing and wave calculations (chapter 7) are provided in the CD-ROM.

2 BACKGROUND

Many of our traditional and modern built hydraulic structures are imperfect and we increasingly have to live with potential failure of hydraulic structures of different functions. The reason is often not insufficiency of engineering concepts or understanding of physics of the systems during design. But, frequently we remain ignorant (in quantitative terms not qualitatively) of the fact that design parameters as well as operating conditions are subject to many deviations and uncertainty. Quantitatively specifying potential inherent imperfections of structures is essential. In this regard, probabilistic design and safety analysis should play great role. It permits the description of variations (uncertainty) in design and operation parameters. It ultimately allows restricting design by giving quantified performance (capacity, reliability, risk, safety). The question that remains is thus: how to design for required (specified) performance (capacity, risk, safety, reliability)?

This chapter gives brief theoretical perspectives and concepts towards answering this question. Moreover, a general risk analysis implementation framework is provided. The chapter's last section provides review of related works. Succeeding two chapters' present details of proposed stochastic and analytical techniques to accomplish the above task.

2.1 Terminologies and guiding principles of probabilistic design, risk and safety analysis

It is said time and again that there are huge uncertainties embedded in any design. These uncertainties are caused by the randomness in loads, material characteristics, geometry, boundary conditions, construction, maintenance, and operation inputs, time dependent nature of strength (deterioration), etc. The randomness of these parameters is both in space and time. The temporal value of each design parameter is, therefore, likely to be different from the value used in the deterministic design. If values of design parameters distributed either in time (such as flood, wind velocity) or distributed across a space (like, sediment flow, dry bulk density or fill material angle of friction (ϕ')) is measured data on frequency of occurrence of values of respective parameters will be found. If there are sufficient data the frequency values can be rescaled to give probability. Design parameters may then be viewed as random variables and their uncertainty could be modeled using pdfs.

Random variables (RV) are numerical values representing every possible outcome of an experiment or a sample space. Most physical variables used in engineering design in fact are random variables. For example, sections 5.5 and 5.6 discuss the randomness in strength and load parameters involved in stability and overtopping analysis of Tendaho Dam. Figure 5-14 to Figure 5-17 shows how the recorded data frequency values are rescaled to probability. Since performance or capacity of designed systems are functions of design parameters, and design parameters are random variables, performance or capacities of systems is thus a *function of random variables*. Evidently, therefore, performance too is random variable. For each type of a random variable a probability distribution (pdf) may be produced or assigned (see 9.1.1.3 and 9.1.1.5).

Functions of random variables and the concept of reliability functions can mathematically be defined as follows:

A functions, say $Y = (X_1, X_2, \dots, X_n)$ is a *function of random variable* over the domain of $X_1, X_2, \dots, X_n$ if $X_1, X_2, \dots, X_n$ are independent random variables and that can be represented by a probability density functions, say $f(x_1), f(x_2), \dots, f(x_n)$.

Probabilistic design studies how to make calculations with probability distributions of design parameters. It uses probability distributions, instead of nominal or mean values, while designing a system to determine capacity and calculate failure probability (safety). A central tool in probabilistic design is, therefore, the ability to calculate functions of random variables and analytically model the physics (relationships between loads and strengths that define failure mechanisms). The relationships between load and strength elements of a system are usually modeled using the concept of *reliability functions*. Reliability functions are computational models on the basis of which failure mechanisms are defined as a difference between set of pertinent loads and strengths Eq. (2-1) *Reliability function* (Z) is thus expressed as a difference between strength and load:

$$Z = S - L \qquad (2\text{-}1)$$

Where, S refers to strength or capacity of a system and L refers to load or hazard or demand on a system. Zero value of Z defines a limit state condition. Negative values of Z correspond to failure and positive values represent survival. Failure occurs when load exceeds or equals strength of a designed system. At $Z = 0$ the system is on the verge of failure and this limit state condition is usually categorized under failure probability. Hence, $Z \leq 0$ defines failure. Consequently, failure probability (P_f) is defined as the probability that $Z \leq 0$, which can be given as:

$$P_f = \int_{-\infty}^{0} f(z) \qquad (2\text{-}2)$$

However, as discussed earlier many of load (L) and strength (S) parameters encountered in dam safety analysis are uncertain parameters (for example see sections 5.5 and 5.6). And thus, modeling failure mechanisms using reliability functions and the calculation of failure probabilities require to do computations using pdfs of the uncertain design parameters by considering Eq. (2-1) as a function of random variable. Since the arguments of Eq. (2-1), i.e. the set of relevant L and S parameters are random so, too, is Z. The pdf of Z (which is represented by $f(z)$) can be derived from the pdfs of those of L and S parameters and the functional relation between them in the reliability function as:

$$P_f = \iint_{LS} f_Z(L,S)\,dL\,dS \qquad (2\text{-}3)$$

However, for complicated (multivariate) models with vector values of L and S, which are typical in practice, finding analytical solution for Eq. (2-3) is believed to be impossible. Consequently, numerical/statistical simulations method (like MCSM) or approximation methods (like FOSM and SOSM) are employed. For example, Hartford (2004) states that the analytical

solution of Eq. (2-3) is not possible. As a result, the analytical solution has never been presented for use by engineers with its detailed theory, application procedure, and solution strategies. In this dissertation both the approximate methods and the analytical solution are explored, discussed, applied and compared. Chapter 3 deals with the theories of the numerical/statistical simulations and approximation methods. The analytical method is discussed in detail in chapter 4. Additional background probabilistic theories, their derivation and integration towards getting a generic equation that enable us for solving problems involving multivariate function of random variables is provided in appendix 9.1. The presentation is made in such a way that is understandable by practicing engineers. The theory with regard to analytical solution of Eq. (2-3) for uni-variate case is given in section 4.3.1 and for the multi-variate case it is given in 4.3.2. The solution implementation steps for uni- and multi-variate problems are illustrated in detail using a case study in 6.3.1 and 6.3.2/6.3.3 respectively. The implementation of all the methods in practical dam safety analysis and their comparison is given in chapters 6 and 7. The topic of uncertainty modeling using pdfs is covered in appendix 9.1.1.5.

Frankel (1988) defines *reliability* as the freedom from failure of a component or system while maintaining specific performance. System *reliability* is also expressed as the probability that a system has not failed, i.e. the probability that Z is positive. Reliability of a system, a variable defining a dimension or magnitude, resistance parameter defining safety etc can also be defined as the probability that the system, the variable, the parameter perform as expected. In simple terms in this thesis reliability would mean probability of conformance to specifications. Whereas, *safety* (the sate of being safe) refers to the degree of freedom from the occurrence of risk. Safety is about the quality of averting or not causing injury, danger, or loss. Therefore, the freedom from failure (reliability) of a system is not quantitatively equivalent to safety, as not all failures result injury, damage or risk (nor all devastating failures cause same degree of injury, damage or risk). The numeric characterization of safety issue is an essential but complex task. Risk analysis provides the most complete characterization of safety issue. *Risk* (R) is a combined measure of the probability and severity of an adverse effect. It is often estimated by the product of the probability of the adverse event occurring, also called failure probability (P_f), and the expected consequences or damage (D) Eq.(2-4). In case of dams, an adverse effect refers to a functional and/or structural failure of dam or one or more of its components due to a specific failure mechanism (physical process representing a particular performance) regardless of whether complete damage on the dam's integrity (collapse) occurs.

$$R = P_f \cdot D \qquad (2\text{-}4)$$

Where P_f refers to probability of failure and D refers to range of adverse consequences (damage) given failure put in monetary terms. Therefore, risk analysis requires (1) determination of failure probabilities. (2) probabilistic estimation of consequences.

In this thesis *reliability analysis* and *probabilistic design* are taken to be synonyms as they both deal with analysis of failure of structures or their components. But, customarily these two methods followed two different approaches. Reliability analysis traditionally studies failure or reliability of systems based on historic failure data but probabilistic design analyzes failure or reliability based on stochastic evaluation of interaction of design parameters (load and strength parameters) by describing the physics of the system using reliability functions. Their final result

actually gives the quantitative characterization of system performance in terms of probability of failure, i.e. $P(L \geq S)$ or in terms of the probability of the "safe" event (probability of survival), i.e. $P(L < S)$. Thus, both reliability analysis and probabilistic design methods end up with determinations of failure probabilities. And failure probabilities computed using either of these two methods can be used in calculating expected damage (risk) given failure. The risk computed thus gives quantitative characterization of safety.

The words risk, reliability, safety, failure, collapse and damage are sometimes used in various definitions, at times in a confusing way. Among the various definitions for risk are (1) risk as probability of unwanted event, (2) consequence (damage) of an unwanted event, (3) product of failure probability and consequence of unwanted event. In this thesis (and usually in risk analysis) the third definition of risk is used. Safety is the degree of freedom from occurrence of damage. Damage is the consequence of dam's functional failure and/or collapse usually given in monetary terms. Thus, based on the third definition, risk gives the expected damage. It is therefore a quantitative characterization of safety. Reliability, as defined above, is the freedom from failure of a component. It is complementary to probability of P_f.

In case of dam engineering the use of reliability analysis in its traditional form, i.e. based on failure data, is not practical. That is because dams do normally have long life time and often statistically adequate failure data is simply not available. It is also particularly difficult to test dams and most other hydraulic systems to life or to do accelerated failure tests in laboratories and generate artificial failure data, as is done in electronic, electrical and mechanical systems. Probabilistic design approach helps to arrive at reasonable conclusions based on theoretically justifiable procedures and physics that can be proved where as reliability analysis mostly relies on empirical proves with out explaining and understanding intrinsic mechanics of systems.

The other confusion in the use of terms is between *failure* and *collapse*. A dam or its component fails if it can no longer perform one of its principal functions without automatically implying collapse. Dam or its component collapses if it undergoes deformation of such magnitude that the original geometry and integrity are lost.

In conclusion, probabilistic design, risk or safety analysis approach requires mathematical modeling of failure mechanisms (physical process) using reliability functions in which design parameters are expressed in terms of their respective pdfs. Failure occurs when load exceeds or equals strength of a designed system. Hence, $Z \leq 0$ defines failure. Consequently, failure probability (P_f) is defined as the probability that $Z \leq 0$. Risk (R) is expressed as an expected damage, which is given by the product of failure probability (P_f) and damage (D). Enhancing dam safety or reliability, i.e. reducing risk, increases investment costs, and some balance must be set between the two. It is necessary to determine levels of acceptable risk and investment. The central idea is to lower lifecycle costs of structures, accounting risk cost, without scarifying performance, i.e. selecting the design with list computed overall lifecycle cost (see Figure 2-1). Figure 2-2 gives the steps in risk analysis process schematically.

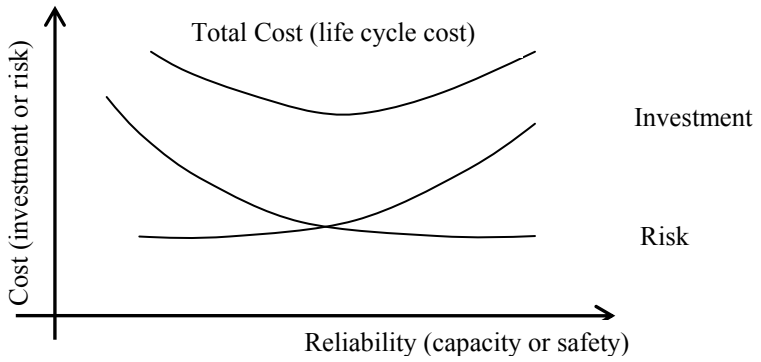


Figure 2-1. Risk analysis logic.

Performance is not a static characteristic. Therefore, risk analysis requires determination of performance across time horizon and estimating damage costs across the same time horizon. Such a procedure engages problems like modeling deterioration (condition of a structure across a time horizon), modeling effects of maintenance and operation inputs. These later issues are beyond the scope of the current research.

2.2 Risk analysis methods

The methods for conducting probabilistic design/risk/safety analysis are classified into different groups by different authors based on different criteria. For instance, Vrijling and Verhagen (2000) classify them in to four classes based on the approach employed for handling uncertainty. These classes are:

(1) *Deterministic approach (level 0 approaches):* here design is based on average situations and an appropriate safety factor is included to obtain safe structure. Actually, this approach level is not probabilistic. The approach simplifies the problem by assuming design parameters as primarily deterministic variables and uncertainty is accounted through the use of empirical safety factors. Safety factors are derived based on past experience. The problems associated with this approach includes one it does not absolutely guarantee safety or satisfactory performance. Moreover, it does not provide any information on how the different parameters of the system influence safety. Therefore, it is difficult to design a system with a uniform distribution of safety level across the different components.

(2) *Semi-probabilistic approach (level I approaches):* here characteristic value of selected design parameters among the suite of uncertain design parameters are used in design, like the load which is not exceeded in 95% of the case, or the strength or material property which is available for 95% of the construction material. Another example could be a dyke height which is available for a flood that is not exceeded in 95% of the time. In this example the flow and the dyke height are taken to be the two most important characteristic values in the load and strength side respectively. However, in reality there are many other uncertain parameters which determine even the interaction between these two parameters dyke, i.e. height and flow.

(3) *Probabilistic approach (level II and III approaches)*: this approach considers the full statistical distribution of parameters. There are two possible directions within the probabilistic approach:

a. *Level-II approach:* this comprises a number of approximate methods in which the problem is linearized, and where all probability density functions are replaced by probability density functions following normal distribution. However, normally one can not model every design parameter using the normal distribution. The safety index method, point estimate methods, FOSM and SOSM methods, etc lie under this category (refer chapter 3 for these methods).

b. *Level-III approach*: this approach considers best fit probability density functions of all relevant stochastic design parameters from both strength and load sides. In this approach, the possible non-linear character of the reliability function is taken into account. The MCSM and ASDD methods can be categorized under this group.

Möller (2006) classifies risk analysis methods into four classes based on the approaches employed for handling uncertainty, in a more or less similar way with Vrijling and Verhagen (2000). According to Möller (2006) the four classes are: (1) *conventional (classic) deterministic* design that uses global safety factor γ. (2) the *semi-probabilistic (quasi deterministic)* design that uses partial safety factors (γ_F, γ_m, γ_{sys}, etc) by splitting the global safety factor to load, material, geometry etc safety factors. (3) *probabilistic approximation* that uses safety (reliability) index β, and (4) the *probabilistic "exact"* that calculates failure probabilities P_f. The safety index method is discussed under section 3.3.

Hartford and Gregory (2004) classifies risk analysis methods based on the method employed to represent the logic of systems, i.e. how the interrelationship among initiating events, state of nature and conditions of the system are formulated. Based on this the principal methods are: (1) failure modes and effect analysis and associated methods; (2) event tree analysis (ETA); and (3) fault tree analysis (FTA). Hartford and Gregory (2004) underpin that the task of assigning probabilities to branches (failure modes), for instance in event tree analysis, committed a lot of attention and discussion and they advocate that probabilities be assigned using one or a combination of the following four methods:

(1) Statistical estimates based on empirical data;
(2) Engineering models based on physical processes;
(3) Fault tree analyses based on logical constructions;
(4) Judgment by experts

The second and the third of these methods may be grouped together under engineering reliability models. Actually, the scope of this thesis is primary focused on this important problem of assigning probabilities to failure modes based on physics of systems. Integration of the failure modes using fault tree the subsequent consequence analysis are beyond the scopes of this research. The approaches followed in this thesis could be grouped in level III probabilistic analysis according to Vrijling and Verhagen (2000) classification and in probabilistic exact category according to Möller (2006).

2.3 Risk analysis process

Dam risk analysis seeks to address several fundamental issues including: what is the cause of dam failure or collapse? How probable is it? What are the various consequences and their associated probability? What are the probability weighted consequences, or risk. The entire risk analysis process is schematized in Figure 2-2. It begins with preparation of an inventory of loads, hazards, and strength (capacity or resistance) and their interaction mechanism. A mechanism is defined as the manner in which the dam responses to loads and hazards. A combination of hazards/loads and dams strength in a mechanism (physics of the system) leads, with a particular probability, to failure or collapse of a dam or its component parts. The relationships between load and strength elements are then modeled using reliability functions that clearly depict failure and non-failure situations. After formulating the reliability function and setting limit state conditions what follows is conducting probabilistic analysis for the determination of failure probabilities. For the calculation of failure probability any of the approximate stochastic models discussed in chapter 3 or the analytical method discussed in chapter 4 can be utilized.

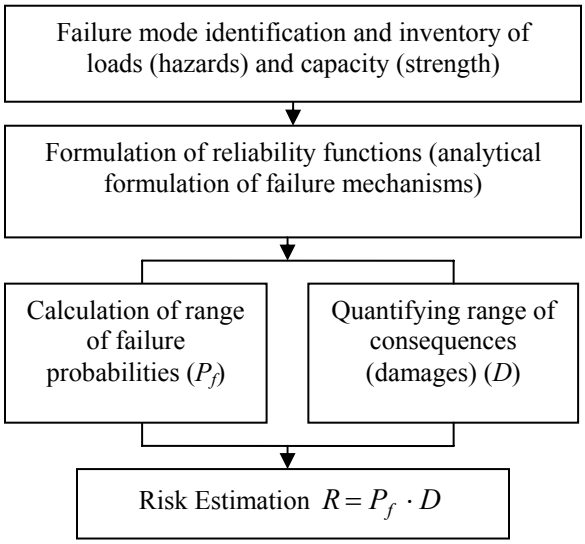


Figure 2-2. Risk analysis process.
(modified from Hartford and Gregory (2004) and Vrijling and Verhagen (2000))

Kaare HØeg in the forward of the book by Hartford and Gregory (2004) writes that dams seldom fail due to a simple fault in design or construction. A failure is generally a result of a complex linking of circumstances, conditions, faults or errors, when combined result in failure (a failure mode). Dams are composed of many components and functional units, each of which may be prone to many hazards and mechanisms. It is very important to consider a functional unit as a whole. The failure of some components may lead directly to a failure of the entire system or to a failure of one or more functional unit ("series connection"). A functional unit is to mean principal functions such as flood defense, storage of adequate water for power, irrigation or water supply, navigation, flow control etc, which is accomplished by a component of a dam or its entirety. If demands of a principal function are not adequately satisfied it will result one or more consequences (damage). In other cases components may compensate for one another ("parallel connection") and failure of one component might not result damage. A useful aid to establish an ordered pattern in the many hazards, mechanisms and components is provided by diagrams such

as *failure mode and effect tables*, *fault trees* and *event trees*. Hartford and Gregory (2004) underpin failure mode identification and establishing an ordered pattern is an essential step in the risk estimation process as it lays the foundation on which the remainder of the study is built.

Failure mode and effect trees analysis (FMEA) is a method of analysis whereby the effects or consequences of individual component failure modes are systematically identified and analyzed. It is an inductive analysis (i.e. is based on the question "what happens if a component or element fails?"). To conduct this analysis it is necessary first to break the system down into its individual components or elements.

Even tree analysis (ETA) is a technique that is used to identify the possible outcomes and, if required, their probabilities given occurrence of an initiating event. It is an inductive type analysis where the basic question that is addressed is "what happens if…..". It starts with formulation of initiating events and ends with assessment of response of the dam system. Whereas, a *fault tree analysis* (FTA) is based on the opposite procedure. In fault tree analysis factors that can contribute to a specified undesired event (called the top event) are deductively identified, organized in a logical manner and represented pictorially. Starting with the top event, the possible causes of failure modes on the next lower functional system level are identified. In drawing a fault tree symbols such as AND-gates and OR-gates are used. Fault tree analyses (FTA) are sometimes used within an event tree. FTA analysis is common in industrial applications to mechanical and electrical systems. For dams, FTA is therefore often associated with tree branches having to do with spillway gates and turbines. Details of FMEA, ETA and FTA techniques and example applications are given in (Hartford and Gregory 2004).

A risk analysis process is concluded with the determination of the probability weighted consequence of failure (risk). The topic of dam failure consequence analysis is extensively treated in an innovative way in (Hartford and Gregory 2004).

2.4 Related Works

Uncertainty modeling, probabilistic design, risk and safety analysis are getting increasing importance in engineering science. The intensive research world wide is a sure sign of wider applications in the future. Researches in this field are numerous. Some selected works related to the current research are outlined under two aspects (1) on researches and reference materials related to development and dissemination of basic theories, and (2) applications of risk analysis methods in hydraulic engineering and related fields.

2.4.1 On basic theories of probabilistic design, risk and safety analysis

Material on details of approximate reliability analysis methods can be found in books by (Thoft-Christensen and Baker, 1982), (Ang and Tang, 1984), (Thoft-Christensen and Murotsu, 1986), (Melchers, 1987) and (Arora, 1997), among others. Hartford and Gregory (2004) contains a through review of the sate-of-the-art dam risk assessment and management issues. This book covers topics like guiding principles of risk assessment and theories on assessment of probability values to be used in risk analysis procedures. Moreover, it presents a new approach and comprehensive sections on consequence analysis and assessment which are necessary for the

estimation of risk. (Mays, 1999) has a chapter on risk/reliability-based hydraulic engineering design. He discusses approximate analytical and probabilistic techniques for risk-based design of hydraulic structures. Topics on advanced state-of-the art of uncertainty modeling are given in books (Elshakoff, 1995), (Ayyub, 1998), (Möller and Beer, 2004), (Möller and Reuter, 2007). The later two books focus on describing uncertain design parameters as fuzzy and fuzzy random variable. With the introduction of the theory of fuzzy random variables a comprehensive modeling of uncertainty is now possible. Both randomness (stochastic uncertainty) and fuzziness (non-stochastic uncertainty-informal and lexical uncertainty) can now be considered simultaneously (Möller and Beer, 2004), see also (Kratschmer, 2001), (Kwakernaak, 1978 and 1979), (Coubi, 2001). Applications of fuzzy approach in reliability study are, for example, shown in (Chou and Yuan 1993, Cheung 1997). Möller and Beer (2004) give extensive discussion on the phenomenon of uncertainty and state of development of uncertainty models.

2.4.2 On applications to hydraulic engineering

Lin and Yen (2003) has done a comparative study of the accuracy and efficiency of various reliability analysis methods using an example culvert. The five basic risk analysis methods used in this study are: the Mean-value First-Order Second Moment Method (MFOSM), advanced first-order second moment method, Point Estimate (PE) method, Latin Hypercube Sampling (LHS) and MCSM. Manache and Melching (2004) applied the LHS technique, in combination with regression and correlation analyses, as a sensitivity analysis technique to the DUFLOW model developed for the Dender River Belgium. Kuo (2007) applied Rosenblueth Point Estimate Method (RPEM), Harr's Point Estimate Method (HPEM) and MCSM for assessing dam overtopping risk. Silliman et al. (1990) utilized first order reliability theory (reliability index) and MCSM to investigate variability in groundwater flow. Ramly (2002) used MCSM for probabilistic slope stability analysis using @Risk software and the method is illustrated on probabilistic slope analysis of dykes. Most of the methods applied by these researchers are discussed in chapter 3 (section 3.3). Additional citation of related works is given in the application chapters (chapters 6 and 7). Also, several international conferences (e.g. ICOLD symposium on uncertainty assessment in dam engineering 1 - 6 May Tehran, 2005; two symposiums on dam safety organized by dam safety association of Turkey, 2007, 2009) give valuable papers on extensive issues of uncertainty and risk analysis in dam engineering. In addition, two international journals (structural safety and probabilistic engineering mechanics) deal extensively with structural reliability concepts, methods, and applications.

3 MONTE CARLO SIMULATION AND APPROXIMATE MOMENT ANALYSIS METHODS

3.1 The Monte Carlo Simulation Method (MCSM)

As its name signifies the MCSM is merely a simulation method. It is a method of statistical trial. This method has been in use for long to slackly model randomness in probabilistic design and risk analysis applications. The implementation architecture of MCSM is presented in Figure 3-1. In MCSM design parameters (X_i's) are treated as random variables bound by certain pdf and simulation (repeated random experiment) is run to randomly generate arrays of X_i from the possibilities defined by the respective bounding pdf's of the X_i's. At each simulation step a corresponding value for Y_i is calculated using functional relations in design equations (multivariate functions of random variables) and the generated X_i's. In this way, the simulation calculates numerous scenarios for outputs of the FRV (outputs of pertinent design, performance, capacity or reliability equation), i.e. Y_i's (see Figure 3-1). The user selects sufficient number of simulations so that the solution converges, i.e. until a point beyond which increasing the number of simulations has minimal effect on the pattern of pdf of Y.

In MCSM random design variables have to be described using continuous distributions (pdf's). Distributions that best describe a design variable's randomness are determined from distribution fitting process run on data found from laboratory, for example, in case of material characteristics or from recorded observations, like in case of flow and metrological variables. The distribution-fitting process success is dependent up on the quality and quantity of available data. Once a distribution is fitted to a data set, in the simulation it will be assumed that the data comes from the distribution. This assumption must be examined; i.e. how well the distribution fits the data must be scrutinized. To prioritize suitability of standard distributions most of the time classical statistical goodness-of-fit tests are used. Among the numerous statistical goodness-of-fit test methods available Chi-square test and Kolmogorov-Smirnov test are the once most often used in hydrology and the top two in order of importance (ICOLD, 1992). However, simple graphical goodness-of-fit test can too be used (see appendix 9.1.3). This involves simple plotting of frequency charts of data and theoretical pdf's on same axis and checking how well the theoretical distributions fits the data. For more discussions on goodness-of-fit tests one can see classical statistical or reliability analysis texts, for instance in (DeGroot, 1975), (Leitch, 1995). (O'Connor, 1991).

The thought behind MCSM can be briefly expressed and illustrated using the following simple case. Assume that a design equation for dam freeboard is defined by a simple relation $F = A + B$; where, A is, fore example, flood surcharge, B is wind generated wave. F is the freeboard required to accommodate the combined flood and wave surcharge. In this simple example, A and B are equivalent to the uncertain design variables $\boldsymbol{X} = (X_1, X_2) = (A, B)$ of the general formulation in Figure 3-1 and F is equivalent to the FRV Y.

Consider that a parallel (simultaneous) records of A and B is available and we know that A is 95 per cent of the time less than a given value a and B is 90 per cent of the time less than value

b (here capital letters represent a random variable and small case letters represent respective realization). In the best case, such statements come from analysis made on recorded flood (*A*) and wave (*B*) data, in which 5 and 10 per cent of them gave results greater than the limits *a* and *b* of the respective records. For instance, there may have been 20 records for *A* with one data greater than the limit *a* and 20 records for *B* with two data greater than the limit *b*; or there may be 100 records for *A* with 5 greater than the limits and 100 for *B* with 20 greater than the limit, or 1000 with 50 and 100 greater than the limits respectively. For instance, take the first case where there are 20 records for each of *A* and *B* that are taken parallel (simultaneously). Obviously we can add concurrent records of *A* and *B* to get 20 values for *F*. Using these 20 computed values of *F* any quintile of interest can be evaluated, say 5% exceedence limit. Note that this is not equivalent to adding the 5 per cent quintiles of *A* and *B*.

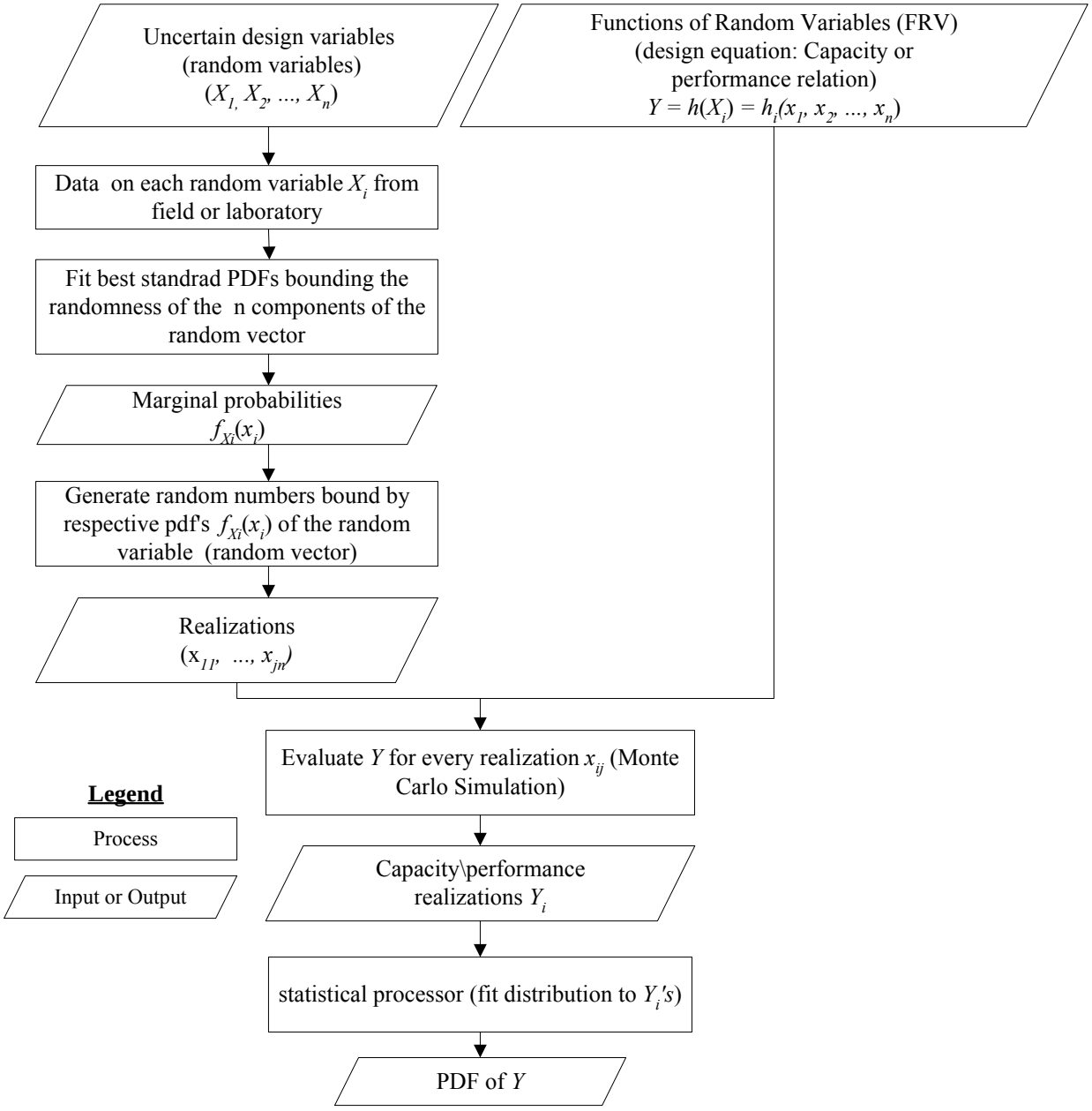


Figure 3-1. MCSM implementation architecture.

It would be possible to produce data of a similar nature, like the actual records of *A* (flood) and *B* (wave), by the use of two 20-sided dies in which one of the side of dies *A* labeled $>a$ and two sides of die *B* labeled $>b$ and the other 19 and 18 sides labeled $<a$ and $<b$ respectively. These two dies would be thrown simultaneously as many times as required with the appropriate concurrent values being recorded according to the face that fell uppermost on the respective die. The list of *a* and *b* produced, while not being identical to the real recorded data, would resemble it in a number of ways. Within limits allowed by statistical variability probabilities of exceedence for *a*, *b* and any functions of *a* and *b* would be similar. It would be impossible to tell which data set was which from any statistical test that could be applied. Thus, by adding the concurrent die generated *a* and *b* values a similar synthetic *F* record can be generated. It is this fact that is used in MCSM. In practice, dies are no longer used to generate random data for simulation. Random number generators in statistical packages and calculators can be used instead. In real life the realizations of *A* and *B* could either be equally likely over their domains (truly random- uniform distribution) or somehow peaked or skewed (biased- other continuous distributions). For example, say the wave height might range between 0 to 3 m but heights around 1m are more frequent and thus has more chance than the other heights on the domain. In such cases we can chose a random number generator with similar bias as the recorded data. This is usually done by taking random numbers generated from distributions that best describe the characteristics of available data record. That is why distribution fitting is required in MCSM. This is a functionality that is not achievable with fair dies but computers.

The drawback of MCSM includes its potential to turnout non-unique value at small number of simulation runs (converge slowly). Usually, over 10,000 simulations are run. Moreover, because it depends on simulation runs rather than a one go solution of mathematical expressions, it does not easily lend itself for adoption in analytically defined reliability functions and risk analysis formulations. The technique has the advantage that it is relatively easy to implement and can deal with a wide range of multivariate functions; including those that cannot be expressed conveniently in an explicit form (it is very versatile). An example of the use of Monte Carlo method can be found in (Negede and Horlacher, 2008b, 2009), (Horlacher and Negede, 2008), (Joos et al., 2005), (Satoh H. et al., 2005), (Lian et al., 2003), (Melih et al., 2004), and (Negede, 2002).

3.2 The Moment Analysis Approximation Methods (MAM)

Moments of distributions are ways of summarizing important characteristics of distributions as single numbers (see appendix 9.1.1.6). The first few lower order moments, like mean, variance, skew, and kurtosis are the most import in probabilistic design and risk analysis. MAMs are approximate ways of solving distribution problems involving multivariate FRVs.

The basic principle of MAM is specification of randomness of uncertain variables that are arguments of a design equation by their moments- mean, variance, skew and kurtosis, i.e. $\mu_{X_i}, \nu_{X_i}, s_{X_i}, k_{X_i}$ etc. Then the moments (μ_Y, ν_Y etc) of the outputs of the design equations are given as a function of moments ($\mu_{X_i}, \nu_{X_i}, s_{X_i}, k_{X_i}$ etc) of the uncertain variables (see Figure 3-2).

In the general case, MAM can only estimate moments μ_Y, ν_Y etc of output of a design equation (like FRV Y in Figure 3-2). It does not, however, give a distribution of known standard type for outputs of FRV- Y. Nonetheless, distributions of Y can be approximated using the derived moments μ_Y and ν_Y. MAM mostly assumed that normal distribution characterizes the Y randomness and this distribution is constructed using the computed mean μ_Y and standard deviation $\sigma\sqrt{\nu_Y}$ and by applying central limit theorem (CLT). CLT states that the distribution of the sum of n random variables approaches normality as n becomes large (Feller, 1967). To use the CLT the first two moments are enough. However, the skew, kurtosis and other higher moments can be calculated to asses whether the assumption of the CLT is plausible or not. For example, using the computed skew one can asses how valid it is to assume that the distribution of $\boldsymbol{Y}$ is symmetric and hence perhaps approaximatable by normal distribution. The computed kurtosis can also be used to compare the expected squatness of Y pdf with that of squatness of the approximated normal distribution.

For the simple cases of FRV, such as linear functions, functions made of products, and functions made of positive integer powered random variables, the computation of the moments of Y (μ_Y, ν_Y etc) can be exact and it uses particular forms of expressions (discussed in sections 3.2.1 to 3.2.3). However, in many practical cases design equations take complex forms than simple cases and thus an approximation is needed. The most common approximate moment methods are first-order second moment (FOSM) and second order second-moment approximations (SOSM) (discussed in sections 3.2.4). The founding theories of moment analysis method are given in appendix 9.1.1.6 and 9.1.1.7.

3.2.1 MAM for linear FRV

Based on discussions in appendix 9.1.1.7, for the special case of a linear function of several random variables, an expression that take particularly simple form can be derived for uses in calculating moments of outputs from such function. The general form of the moment's relations for simple cases involving linear sum are derived below.

Expectation (mean) of a linear function of random variable:

If X_1, X_2, X_3,…, X_n are independent random variables, a_1, a_2, a_3,…, a_n are constants, and $Y = a_1 \cdot X_1 + a_2 \cdot X_2 + a_3 \cdot X_3 + ... + a_n \cdot X_n = \sum_{i=1}^{n} a_i \cdot X_i$, then from Eq.(9-36) that is given in appendix 9.1 the expectation of Y is expressed as:

$$\mu_Y = a_1 \cdot \mu_{X1} + a_2 \cdot \mu_{X2} + a_3 \cdot \mu_{X3} + ... + a_n \cdot \mu_{Xn} = \sum_{i=1}^{n} a_i \cdot \mu_{Xi} \qquad (3\text{-}1)$$

Variance of a linear function of random variable

Let X_1, X_2, X_3,…, X_n are independent random variables, a_1, a_2, a_3,…, a_n are constants, and $Y = \sum_{i=1}^{n} a_i \cdot X_i$, then the variance of a linear function of random variables is given as:

$$\nu_Y = \sum_{i=1}^{n} \left(a_i^2 \cdot \nu_{X_i} \right) \qquad (3\text{-}2)$$

Proof:

$$\nu_Y = E\left[(Y-\mu_Y)^2\right], \text{ by Eq.(9-34) given in appendix 9.1,}$$

Replacing Eq.(3-1) and $Y = \sum_{i=1}^{n} a_i \cdot X_i$ in this,

$$\nu_Y = E\left[\left(\sum_{i=1}^{n} a_i \cdot X_i - \sum_{i=1}^{n} a_i \cdot \mu_{X_i}\right)^2\right]$$

$$= E\left[\left(\sum_{i=1}^{n} a_i \cdot \left(X_i - \mu_{X_i}\right)\right)^2\right]$$

$$= \left[E\left[\sum_{i=1}^{n}\left(a_i^2 \cdot \left(X_i - \mu_{X_i}\right)^2\right)\right]\right]$$

By Eq.(9-37) which is given in appendix 9.1,

$$= \sum_{i=1}^{n}\left(E\left[a_i^2\right] \cdot E\left[\left(X_i - \mu_{X_i}\right)^2\right]\right)$$

From Eq.(9-34) and Eq.(9-35) this takes the form,

$$\nu_Y = \sum_{i=1}^{n}\left(a_i^2 \cdot \nu_{X_i}\right)$$

This shows that always linear combination of random variables results variation buildup. The derivation of Eq.(3-1) and (3-2) assumes that the random variables X_1 to X_n are mutually independent. The derivation for a more general case where there is dependency among the random variable is given in (Harris, 1966).

The expression for skew s_Y of linear sum of random variables can be driven using the same set of principles and following similar procedures and it is given as:

$$s_Y = \sum_{i=1}^{n}\left(a_i^3 \cdot s_{Xi}\right) \qquad (3\text{-}3)$$

3.2.2 MAM for FRV involving products

Similar simple expressions for calculating moments of outputs from especial case of multivariate functions made of products of random variables can be derived as follows.

Expectation (mean) of product of random variables:

Suppose Y is a product of n independent random variables $Y = \prod_{i=1}^{n} X_i$. Then, from direct application of Eq.(9-37) given in appendix 9.1,

$$E[Y]=\prod_{i=1}^{n}E[X_i]$$

Applying (9-33),

$$\mu_Y=\prod_{i=1}^{n}\mu_{Xi} \qquad (3\text{-}4)$$

Therefore, mean of a product of independent random variables is simply product of their means.

Variance of product of random variables

The variance of a product is obtained by taking the expectation of the square of *Y*, i.e. taking expectations of $Y^2=\prod_{i=1}^{n}X_i^2$. Applying Eq.(9-37) which is given in appendix 9.1,

$$E\left[Y^2\right]=\prod_{i=1}^{n}E\left[X_i^2\right]$$

and from (9-41),

$$\left(\mu_Y^2+\nu_Y\right)=\left(\prod_{i=1}^{n}(\mu_{X_i}^2+\nu_{X_i})\right) \qquad (3\text{-}5)$$

To calculate the variance ν_Y of product of independent random variables, first the right hand side of Eq.(3-5) should be computed and then subtracting the square of the mean (μ_Y^2), which is calculated using Eq.(3-4), from the result for the right side gives the variance ν_Y.

In similar way, the expression for the skew of a product of random variables can be derived by taking the expression of the third power of *Y*, $Y^3=\prod_{i=1}^{n}X_i^3$. Applying Eq.(9-37), $E[Y^3]=\prod_{i=1}^{n}E[X_i^3]$ and from (9-41) follows:

$$\left(\mu_Y^3+3\cdot\mu_Y\cdot\nu_Y+s_Y\right)=\prod_{i=1}^{n}\left(\mu_{Xi}^3+3\cdot\mu_{Xi}\cdot\nu_{Xi}+s_{Xi}\right) \qquad (3\text{-}6)$$

To calculate the skew s_Y of product of independent random variables, first the right hand side of Eq.(3-6) and then subtract $\mu_Y^3+3\cdot\mu_Y\cdot\nu_Y$ whose elements calculated using Eq.(3-4) and (3-5).

3.2.3 MAM for FRV involving positive integer powers

Similarly, particularly simple exact expressions can be derived to calculate moments of outputs from the multivariate function made of positive integer powered random variables.

Expectation (mean) of positive integer powered random variables

Eq.(9-42) gives an expression for calculating expectation of a positive integer powered random variable. Since, this expectation gives a mean value it immediately follows that for $Y = X^n$.

$$\mu_Y = E\left[X^n\right] = \sum_{i=0}^{n}\binom{n}{i} \cdot E\left[\left(x - \mu_X\right)^i\right] \cdot \left(\mu_X\right)^{n-i} \qquad (3\text{-}7)$$

From this equation, for example, the relations for the mean values for the first two positive integers can be given as in Eq.(9-41) in appendix 9.1.

Variance and skew of positive integer powered random variables

The relation for variance and skew of positive integer powers can be derived following the same procedure as in the case of deriving the relation for variance and skew of product of random variables.

3.2.4 The FOSM and SOSM approximate moment analysis methods

As mentioned above most practical design and performance equations in dam safety analysis take complex forms than simple linear sums, simple products and positive powers. In such cases the above simple and exact methods of determining the moments μ_Y, ν_Y etc of *Y* does not work. Thus an approximation is needed. The commonly used approximate methods used are the FOSM and the SOSM approximations that use Taylor's series expansion truncated at the first and second orders respectively.

In these approximate methods the procedure for calculating the mean (μ_Y) for outputs of a FRV (design equation or performance function), say $Y = h(\boldsymbol{X})$, where $\boldsymbol{X}$ is a random vector with *n* components $\boldsymbol{X} = (X_1, X_2, ..., X_n)$, involves expanding the FRV $Y = h(\boldsymbol{X})$ as Taylor series about mean values $\left(\mu_{X_1}, \mu_{X_2}, ..., \mu_{X_n}\right)$ of the random variables in *h* and determining μ_Y by calculating expectations of the terms in the expansion. Similarly, the determination of the central moments $\left(\nu_Y, s_Y, k_Y\right)$ involves expanding the function $(Y - \mu_Y)^n$ as Taylor series about the same mean values and calculating expectations of the terms in the expansion and then replacing μ_Y calculated before. The mean and central moments resulting from such calculations are exact provided all terms of the Taylor's series to infinity are used or if the series terminates. But, in many practical applications μ_Y is near $h(\mu_{X_i})$, so higher terms in the series become small and can be truncated after only a few terms. If the truncation is done after the first-order term the approximation is called FOSM approximation and if a better precision is required and truncation is made after the second-order term the approximation is called SOSM approximation. Figure 3-2 presents the implementation architecture of these approximate methods.

To avoid the requirement of doing Taylor's expansion for different design equations every time generic relations for calculating FOSM and SOSM approximations of the first two moments of a universal FRV can be derived. Consider a general function of random variable $Y = h(\boldsymbol{X})$ where $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ are independent random variables. A general relations for the

second-order and first-order approximations of the mean (μ_Y) and variance (ν_Y) of a generic FRV $Y = h(\boldsymbol{X})$ can be divided as follows.

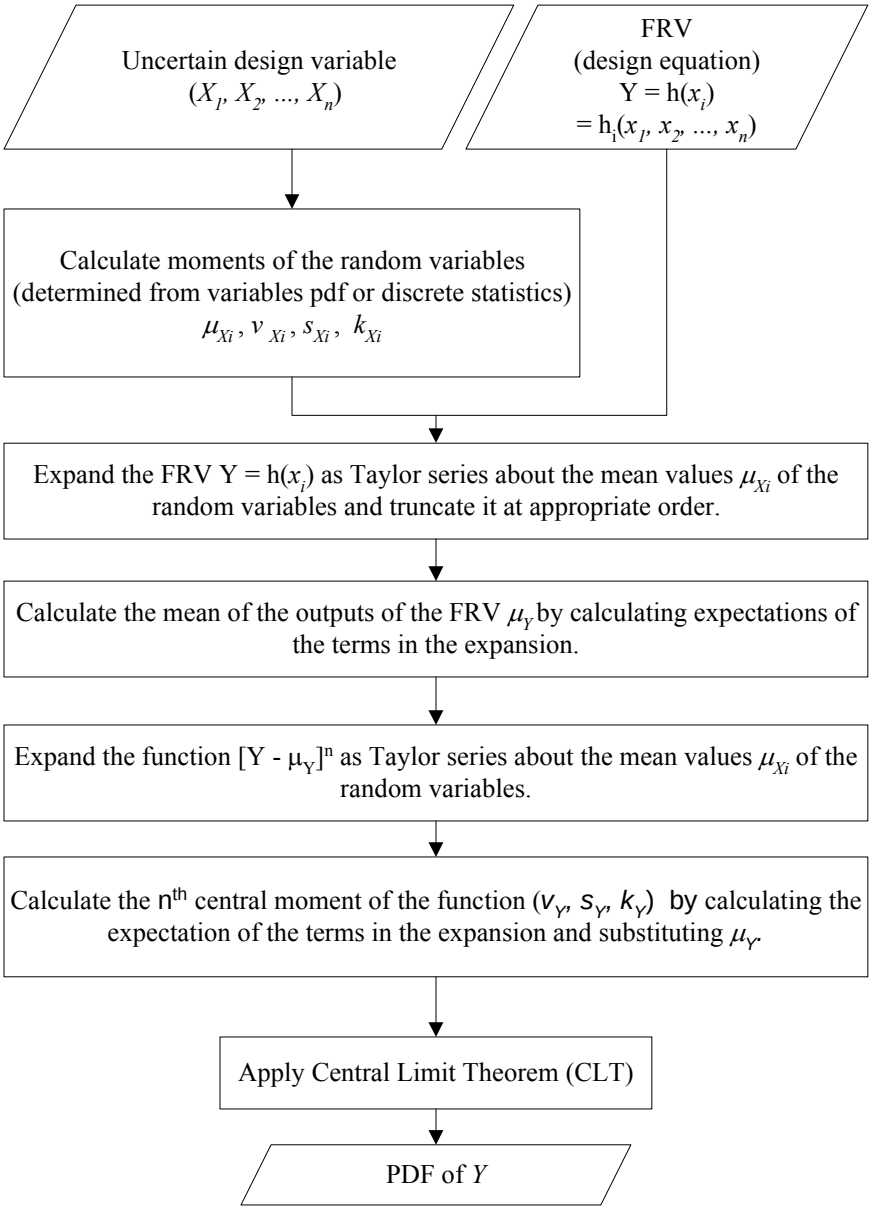


Figure 3-2. FOSM-SOSM implementation architecture.

In the especial case where $Y = h(X)$ with only one independent variable X (uni-variate FRV's) and when the value of h is given for some value of X, say $x = \mu_X$, then $Y = h(X)$ can be found for any other value of X in the neighborhood of μ_X using Taylor's series is defined as:

$$Y = h(X) = h(\mu_X) + \frac{1}{1!}\cdot(x-\mu_X)\cdot\frac{dh}{dx} + \frac{1}{2!}\cdot(x-\mu_X)^2\cdot\frac{d^2h}{dx^2} + \frac{1}{3!}\cdot(x-\mu_X)^3\cdot\frac{d^3h}{dx^3} + \ldots \qquad (3\text{-}8)$$

Where *n*! denotes the factorial of *n* and the derivatives are evaluated at the point $x = \mu_X$.

However, in practice most design functions do have more than one independent variable (most design equations are multivariate FRV's). Therefore, a generalization is needed. There are several equivalent forms of writing Taylor's series for multiple variables. One of the commonly used versions is Eq.(3-9).

Thus for a general FRV $Y = h(\boldsymbol{X})$, where $\boldsymbol{X} = (X_1, X_2, ..., X_n)$, where the X_i's are independent random variables, if the value of $h(\boldsymbol{X})$ is known for some values of $\boldsymbol{X} = (x_1, x_2, ..., x_n)$, say $\boldsymbol{X} = (\mu_{X_1}, \mu_{X_2}, ..., \mu_{X_n})$, then $Y = h(\boldsymbol{X})$ can be found for any other combination of x_i's using Taylor series approximation for multiple variables Eq.(3-9).

$$Y = h(x_1, x_2, ..., x_n) = h(\mu_{x_1}, \mu_{x_2}, ..., \mu_{x_n}) + \frac{1}{1!} \cdot \sum_{i=1}^{n} (x_i - \mu_{x_i}) \cdot \frac{\partial h}{\partial x_i} + \frac{1}{2!} \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \mu_{x_i}) \cdot (x_j - \mu_{x_j}) \cdot \frac{\partial^2 h}{\partial x_i \partial x_j}$$

$$+ \frac{1}{3!} \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} (x_i - \mu_{x_i}) \cdot (x_j - \mu_{x_j}) \cdot (x_k - \mu_{x_k}) \cdot \frac{\partial^3 h}{\partial x_i \partial x_j \partial x_k} + ... \qquad (3\text{-}9)$$

Where n! denotes the factorial of *n* and the partial derivatives are evaluated at $\mu_{X_1}, \mu_{X_2}, ..., \mu_{X_n}$. Further readings on Taylor series can be found in (Greenberg 1998) pp (630-638, 1209-1215).

Ignoring the terms above the second order (truncating all terms above the second order) yields the second order Taylor's series approximation of $h(\boldsymbol{X})$,

$$Y = h(x_1, x_2, ..., x_n) = h(\mu_{x_1}, \mu_{x_2}, ..., \mu_{x_n}) + \frac{1}{1!} \sum_{i=1}^{n} (x_i - \mu_{x_i}) \cdot \frac{\partial h}{\partial x_i}$$

$$+ \frac{1}{2!} \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \mu_{x_i}) \cdot (x_j - \mu_{x_j}) \cdot \frac{\partial^2 h}{\partial x_i \partial x_j} \qquad (3\text{-}10)$$

From Eq.(9-32) the expected value (mean) of *Y* can be found from integrating the product of $Y = h(x_1, x_2, ..., x_3)$ and joint pdf of the variables x_1 through x_n from -∞ to ∞. Each terms of Eq.(3-10) can be integrated in turn and the results added. The term $h(\mu_{x_1}, \mu_{x_2}, ..., \mu_{x_n})$ is a constant and from Eq.(9-35) expectation of a constant is the content itself, thus,

$$\mu_Y = h(\mu_{x_1}, \mu_{x_2}, ..., \mu_{x_n}) + \sum_{i=1}^{n} \int_{-\infty}^{\infty} (x_i - \mu_{x_i}) \cdot \frac{\partial h}{\partial x_i} \cdot f_X(x_1, x_2, ..., x_n) d(x_1, x_2, ..., x_n)$$

$$+ \frac{1}{2} \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} \int_{-\infty}^{\infty} (x_i - \mu_{x_i}) \cdot (x_j - \mu_{x_j}) \cdot \frac{\partial^2 h}{\partial x_i \partial x_j} \cdot f_X(x_1, x_2, ..., x_n) d(x_1, x_2, ..., x_n) \qquad (3\text{-}11)$$

Each of the elements in the second term of Eq.(3-11) must be identically zero as they represent set of first central moments (see appendix 9.1.1.6 number (ii) and Eqs.(9-34)). Thus eliminating the second term yields,

$$\mu_Y = h\left(\mu_{x_1}, \mu_{x_2}, ..., \mu_{x_n}\right) + \frac{1}{2} \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} \int_{-\infty}^{\infty} \left(x_i - \mu_{x_i}\right) \cdot \left(x_j - \mu_{x_j}\right) \cdot \frac{\partial^2 h}{\partial x_i \partial x_j}$$

$$\cdot f_X\left(x_1, x_2, ...x_n\right) d\left(x_1, x_2, ...x_n\right) \qquad (3\text{-}12)$$

Expanding the remaining terms,

$$\mu_Y = h\left(\mu_{x_1}, \mu_{x_2}, ..., \mu_{x_n}\right) + \frac{1}{2} \cdot \left[\left(x_1 - \mu_{X1}\right)^2 \frac{\partial^2 h}{\partial^2 x_1} + \left(x_2 - \mu_{x_2}\right)^2 \cdot \frac{\partial^2 h}{\partial^2 x_2} + \left(x_3 - \mu_{x_3}\right)^2 \cdot \frac{\partial^2 h}{\partial^2 x_3} \right.$$

$$\left. + ... + \left(x_i - \mu_{x_i}\right) \cdot \left(x_j - \mu_{x_j}\right) \cdot \frac{\partial^2 h}{\partial^2 x_i x_j} \right] \cdot f_X\left(x_1, x_2, ..., x_n\right) d\left(x_1, x_2, ..., x_n\right) \qquad (3\text{-}13)$$

All terms of the sort $\left(x_i - \mu_{x_i}\right) \cdot \left(x_j - \mu_{x_j}\right) \cdot \partial^2 h / \partial^2 x_i x_j$ are first central moments thus are identically zero. Noting Eq.(9-26) and doing proper replacing in Eq.(3-13) gives the ultimate expression for the second order approximation of the mean μ_Y of a general FRV $Y = h(\boldsymbol{X})$;

$$\mu_Y = h\left(\mu_{X_1}, \mu_{X_2}, ..., \mu_{X_n}\right) + \frac{1}{2} \cdot \left(\nu_{X_1} \cdot \left[\frac{\partial^2 h}{\partial x_1} \right]_\mu + \nu_{X_2} \cdot \left[\frac{\partial^2 h}{\partial x_2} \right]_\mu + ... \right) \qquad (3\text{-}14)$$

Note that all the partial derivatives are constants as they have been evaluated at the means of the individual variables. The notation $[...]_\mu$ means that the bracketed expression is evaluated at the points $x_1 = \mu_{X_1}$, $x_2 = \mu_{X_2}$, ... etc.

Similarly, the determination of the n^{th} central moments of the function Y (ν_Y, s_Y, k_Y) involves expanding the function $(Y - \mu_Y)^n$ as Taylor series about the mean values $(\mu_{X_1}, \mu_{X_2}, ..., \mu_{X_n})$ of the random variables and calculating the expectation of the terms in the expansion and replacing μ_Y, which is calculated from Eq.(3-14). As an example the expression for the second order approximation to the variance ν_Y of $Y = h(\boldsymbol{X})$ is derived below.

Since $(Y - \mu_Y)^2$ is till a function of $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ we can write $Z = (Y - \mu_Y)^2$. $\nu_Z = E[(Y - \mu_Y)^2]$ then becomes μ_Z thus directly from Eq.(3-14).

$$\nu_Z = E[(Y-\mu_Y)^2] = \mu_Z = [Z]_\mu + \frac{1}{2}\cdot\left(\nu_{X_1}\cdot\left[\frac{\partial^2 Z}{\partial x_1}\right]_\mu + \nu_{X_2}\cdot\left[\frac{\partial^2 Z}{\partial x_2}\right]_\mu + \ldots\right) \qquad (3\text{-}15)$$

Evaluating Eq.(3-15) in parts makes the steps clear. First evaluating the term $Z\left(\mu_{X_1},\mu_{X_2},\ldots,\mu_{X_n}\right)$ by noting that μ_Y is a constant and $Z=(Y-\mu_Y)^2$,

$$[Z]_\mu = \left[(Y-\mu_Y)^2\right]_\mu = ([Y]_\mu - \mu_Y)^2$$

Noting that $Y = h(\boldsymbol{X}) = h(X_1, X_2, \ldots, X_n)$ thus $[Y]_\mu = h\left(\mu_{X_1},\mu_{X_2},\ldots,\mu_{X_n}\right)$ and substituting Eq.(3-14) in the place of μ_Y and simplifying gives,

$$[Z]_{\mu_{x_i}} = \left([Y]_\mu - \mu_Y\right)^2 = \frac{1}{4}\cdot\left(\nu_{X_1}\cdot\left[\frac{\partial^2 h}{\partial {x_1}^2}\right]_\mu + \nu_{X_2}\cdot\left[\frac{\partial^2 h}{\partial {x_2}^2}\right]_\mu + \ldots\right)^2 \qquad (3\text{-}16)$$

Evaluating the second derivative term of Eq.(3-15), note that $Z=(Y-\mu_Y)^2$,

$$\left[\frac{\partial^2 Z}{\partial {x_1}^2}\right]_\mu = \left[\frac{\partial^2\left([Y]_\mu - \mu_Y\right)^2}{\partial {x_1}^2}\right] = \frac{\partial}{\partial x_1}\left[\frac{\partial\left([Y]_\mu - \mu_Y\right)^2}{\partial x_1}\right]$$

Applying chain and product rules of derivative,

$$= 2\left[\left(\frac{\partial Y}{\partial x_1}\right)^2 + (Y-\mu_Y)\frac{\partial Y}{\partial x_1}\right]_\mu \qquad (3\text{-}17)$$

Similarly, the second-second derivative term of Eq.(3-15) takes the form,

$$\left[\frac{\partial^2 Z}{\partial {x_2}^2}\right]_\mu = 2\left[\left(\frac{\partial Y}{\partial x_2}\right)^2 + (Y-\mu_Y)\frac{\partial Y}{\partial x_2}\right]_\mu \qquad (3\text{-}18)$$

Again noting that $[Y]_\mu = h\left(\mu_{X_1},\mu_{X_2},\ldots,\mu_{X_n}\right)$ and substituting Eq.(3-14) in place of μ_Y In Eqs.(3-17) and (3-18) respectively yields,

$$\left[\frac{\partial^2 Z}{\partial {x_1}^2}\right]_\mu = 2\left[\left[\frac{\partial h}{\partial x_1}\right]^2 - \frac{1}{2}\left(\nu_{x_1}\cdot\left[\frac{\partial^2 h}{\partial x_1}\right] + \nu_{x_2}\cdot\left[\frac{\partial^2 h}{\partial x_2}\right] + \ldots\right)\right]_\mu \qquad (3\text{-}19)$$

and

$$\left[\frac{\partial^2 Z}{\partial {x_2}^2}\right]_\mu = 2\left[\left[\frac{\partial h}{\partial x_2}\right]^2 - \frac{1}{2}\left(\nu_{x_1}\cdot\left[\frac{\partial^2 h}{\partial x_1}\right] + \nu_{x_2}\cdot\left[\frac{\partial^2 h}{\partial x_2}\right] + \ldots\right)\right]_\mu \qquad (3\text{-}20)$$

Substituting Eqs. (3-16), (3-19), and (3-20) in Eq.(3-15) collecting terms and simplifying,

$$\nu_y = \left(\nu_{X_1}\cdot\left[\frac{\partial h}{\partial x_1}\right]_\mu^2 + \nu_{X_2}\cdot\left[\frac{\partial h}{\partial x_2}\right]_\mu^2 + \ldots\right) - \frac{1}{4}\cdot\left(\nu_{X_1}\cdot\left[\frac{\partial^2 h}{\partial {x_1}^2}\right]_\mu + \nu_{X_2}\cdot\left[\frac{\partial^2 h}{\partial {x_2}^2}\right]_\mu + \ldots\right)^2 \qquad (3\text{-}21)$$

Eq.(3-21) gives the ultimate expression for the second order approximation of the variance ν_Y of the general FRV $Y= h(\boldsymbol{X})$. The notation $[...]_\mu$ means that the bracketed expression is evaluated at the point $x_1 = \mu_{X_1}, x_2 = \mu_{X_2}, \ldots$.

The procedures for applying Eq. (3-14), (3-21), (3-22), and (3-23) is demonstrated in appendix 9.4.

If the truncation of the Taylor's series Eq.(3-9) has been made just after the first-order term, from Eq. (3-14) and (3-21), respectively the resulting first-order second-moment approximation (FOSM) for the mean (μ_Y) and variance (ν_Y) would obviously look,

$$\mu_Y = h\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_n}\right) \qquad (3\text{-}22)$$

$$\nu_y = \nu_{X_1}\cdot\left[\frac{\partial h}{\partial x_1}\right]_\mu^2 + \nu_{X_2}\cdot\left[\frac{\partial h}{\partial x_2}\right]_\mu^2 + \ldots \qquad (3\text{-}23)$$

In the general case all the MAM's (the simple exact forms, FOSM and SOSM) provide estimates of the first few moments of the output of a FRV. However, they do not provide a distribution of known standard type for characterizing the randomness of outputs of FRV, say $Y = h(\boldsymbol{X})$. They do not either give arrays of possible values of Y, like MCSM, on which distributions could be fitted. Results from this method are mostly presented using normal distribution for characterizing the randomness in outputs of Y. This normal distribution is constructed using the computed first two moments of Y (mean μ_Y and standard deviation $\sqrt{\nu_Y}$). Nonetheless, in special case of linear combination of normally distributed random variables, the outcome of the combination itself is normally distributed. In these unique cases the MAM could give an exact result.

The main advantage of the moment analysis methods (FOSM and SOSM) is its computational eases. It is not a requirement to fit pdf for random variables. It is possible to start from first few moments of random variables determined from desecrate statistics. Moreover, the MAM is ideally suited to being programmed using symbolic programming language. And it lends

itself for easy adoption in analytically defined reliability functions and risk analysis. The accuracy of the result from MAM is dependent up on the degree of linearity of the FRV near the mean values of its random variables, the number of moments used in the analysis and truncation level of the Taylor series. The most linear the function and the more the moments are used the accurate the result will be.

3.3 Summary of other classical approximate methods

Some researchers suppose that determination of pdfs of design parameters is not a trivial exercise to do well. They believe that doing system decomposition for modeling relations between load and resistance parameters using reliability equations is exceedingly complex. Their believe is aggravated by inadequacy of accessible demonstrated solution tools. As a result, they assume pdfs for load (L) and strength (S) parameters are alleged to be difficult to find or simply unknown and doing level III probabilistic analysis (i.e. application of probabilistic exact methods-see section 2.2) are assumed to be particularly difficult, incomplete, imprecise or simply impossible.

For example, Harr (1987) argues more direct probabilistic methods that employee MCSM or truncated Taylor series (FOSM and SOSM methods) are exceedingly difficult, if not impossible, for all but a very few uncorrelated random variables. This thesis argues the contrary, at least for the case of design of dams against the dominant failure mechanisms; nowadays, for most projects across the world there is sufficient data to reasonably characterize leading random design parameters statistically. Moreover, system decomposition and analytical representation of failure mechanisms can be done effectively if there is a mind that thinks a little bit outside the realms of deterministic mind set and if engineers liberate themselves from design equations primarily designed for deterministic analysis. I believe that there is an extraordinary growth in uncertainty analysis and risk-based design theory; what is causing the imbalance between the extraordinary advances in theory and the modest professional practices is the lack of accessible and handy solution tools, shortage of illustrated real life case studies, deficiency in fine-tuning theories so that they suit understanding by practicing engineers and fit professional practice norms (implementation ease, computational efficiency, theory transparency) and absence of design standards that support risk-based design. The rapidly increasing interest for research in this field bears testimony to this speculation. In this thesis it is suggested that an important step forward could be taken by filling these gaps.

Based on the propositions discussed above, which suppose implementation of level III probabilistic approaches is complex or at times impossible, a number of simplified approximate analysis and design formats have been proposed, for example by (Cornell, 1969), (Rosenblueth and Esteva, 1972), (Ang and Cornell, 1974), (Rosenblueth, 1976), (Hasofer and Lind, 1978) and (Harr, 1987), among others. Owing to space limitations, it is not possible to include in thesis detailed review of all such classical approximate risk (reliability) analysis methods. However, brief descriptions of the most commonly known methods and their underling concepts are incorporated. In addition, an alternative way of doing MCSM using Latin Hypercube sampling is discussed. Considerable additional background material on details of approximate reliability analysis methods can be found in books by (Thoft-Christensen and Baker, 1982), (Ang and Tang, 1984), (Thoft-Christensen and Murotsu, 1986), (Melchers, 1987), (Arora, 1997) and (Mays,

1999), among others. In addition, several journal publications on applications of such methods are available. Reference for selected applications will be given under each method below and some of which were already indicated in section 2.4.

3.3.1 The method of reliability index (β)

Reliability index method (also called "second moment method") has been utilized extensively in structural analysis. This simplified approach was first proposed by (Cornell, 1969) and later by (Ang and Cornell, 1974). A convenient theoretical background on the meaning and estimation of reliability index is provided in (Ang and Tang, 1984). This method is a design scheme that proposes to maintain a minimum safety margin; in stead of requiring that P_f be below a specified probability level that gives an optimal balance between safety and investment as in level III approaches. Here, what is required is the mean reliability (μ_Z) to be at least a code-specified number of standard deviations (σ_Z) above zero or in other words the mean factor of safety (μ_{Fs}) to be at least a code-specified number of standard deviations (σ_{Fs}) above one:

$$\mu_Z \geq \beta^o \cdot \sigma_Z \text{ or } \mu_{F_S} \geq \beta^o \cdot \sigma_{FS}$$

Where β^o is code specified value of reliability index,

$$\mu_Z = \mu_S - \mu_L \text{ and } \sigma_Z = \sqrt{\nu_S + \nu_L} = \sigma_S + \sigma_Z$$

are, respectively, the mean and standard deviations of the reliability (safety margin). μ_S and ν_S are mean and variance of the system resistance (strength) and μ_L and ν_L, respectively, refers to the mean and variance of load on the system. These values are conventionally found from either expert's opinion or from statistical estimates based on empirical data. The issue of probability or moments estimations based on these methods is deliberated in (Hartford and Gregory, 2004). Alternatively, they can be found from other level II methods (FOSM and SOSM approximations) (see section 3.2.4).

Therefore, reliability index is a number that represents the number of standard deviations which separate the mean reliability (or factor of safety) from the critical reliability (or factor of safety); i.e. from the points where $Z = 0$ and $F_S = 1$ respectively. It can therefore be is given as:

$$\beta = \frac{\mu_Z}{\sigma_Z} \qquad \text{(when reliability is used for the analysis)}$$

$$\beta = \frac{\mu_{F_s} - 1}{\sigma_{F_S}} \qquad \text{(when factor of safety is used for the analysis)}$$

On the basis of these formulations the design checking equation is:

$$\beta = \frac{\mu_Z}{\sigma_Z} = \frac{\mu_S - \mu_L}{\sqrt{\nu_S + \nu_L}} = \frac{\mu_S - \mu_L}{\sigma_S + \sigma_L} \geq \beta^o \qquad (3\text{-}24)$$

or

$$\beta = \frac{\mu_{Fs} - 1}{\sigma_{FS}} \geq \beta^o \qquad (3\text{-}25)$$

Where β is the actual reliability index and β^o is the code specified reliability index. As a rule of thumb to have reasonable assurance of a safe design $\beta^o = 3$ or greater is recommended in most designs. If the type of distribution for β is known the formula or table for its cumulative distribution function can be used to calculate failure probability P_f corresponding to the code specified β^o. However, usually the distribution for β is not known and simply the *S* and *L* are both assumed to be normal distributed and a normal distribution table is used to relate the code specified value of the reliability index β^o to failure probability P_f.

Table 3.1: Relations between β and P_f based on normal distribution.

Reliability Index β	P_f (values in bracket refer number of failures per million)		Reliability Index β	P_f (values in bracket refer number of failures per million)	
0.00	0.5	(500,000)	2.33	0.01	(10,000)
0.67	0.25	(250,000)	3.10	0.001	(1,000)
1.00	0.16	(160,000)	3.72	0.0001	(100)
1.28	0.1	(100,000)	4.25	0.00001	(10)
1.65	0.05	(50,000)	4.75	0.000001	(1)

Although reliability index was originally developed for normal distributions, similar calculations can be made if *S* and *L* are lognormally distributed (i.e. when the logarithms of the basic variables follow normal distribution). For $P_f > 10^{-3}$ there is only small difference between the use of normal and lognormal distribution in reliability analysis (Yoon, and Michael, 2002), (Lumb, 1966 and 1974), (Ang and Ellingwood, 1971). Hasofer and Lind (1974) found that there is very little change in reliability between lognormal and normal distributions when the two basic variables, capacity and demand, follow both normal and lognormal distributions. They suggest, however, that assumptions of normal distribution of the reliability function (limit state function) can lead to unacceptable results if the reliability index, β, is too large, which may result in negative values of in capacity and demand. Consequently, in reality studies, reliability index (factor of safety) distribution is often best fit by lognormal rather than normal distribution. This also consistent with lognormal distribution characteristics as it is only applicable for variables which are always positive, this is valid for reliabilities and factor of safeties and also for most engineering random variables.

If it assumed that the distribution of safety factor, after a probabilistic analysis, is lognormally distributed, the following Eq.(3-26) is applicable for the calculation of reliability index. To relate the calculated reliability index to P_f a log normal distribution table could be used.

$$\beta = \frac{ln\left(\mu_{Fs} / \sqrt{1 + CV_{FS}^2}\right)}{\sqrt{ln\left(1 + CV_{FS}^2\right)}} \geq \beta^o \qquad (3\text{-}26)$$

Where μ_{Fs} is mean factor of safety CV_{FS}^2 is coefficient of variation of factor of safety = μ_{Fs} / σ_{FS}. Factor of safety is given by ratio between strength and load (S/L) and thus the computation of μ_{Fs} and σ_{FS} from μ_S , ν_S , μ_L and ν_L is not straight forward (section 9.1.1.7). Usually, for their computations Taylor series approximations are employed (section 3.2.4). Application of reliability index method for structural design and its theoretical background is outlined in (Arora, 1997).

The reliability index method does not take performance randomness in its natural terms. It simply assumes that loads and resistance parameters, and thus β , are either normal or lognormal distributed, which is hardly the case in reality. It can be seen from the Table 3.1 that the P_f is quite sensitive to small changes in β , especially for β greater than about 2, which makes it difficult to deal with in design endeavors and decisions. A doubling of β from 2.4 to 4.8 decreases the probability of failure by a factor of 10,000 but mostly the cost implications on the designed system are enormous.

3.3.2 Point estimate (PE) methods

PE methods generally provide a direct computational procedure to obtain moment (particularly mean and variance) estimates for outputs from function of random variables, such as performance or reliability functions, from known moments of random variables. The particular shape the pdf used to describe random variables is not critical to the analysis because the pdf is represented by means and two hypothetical point mass located at plus and minus one standard deviation (σ_{X_i}) from the mean (μ_{X_i}), where X_i represent random variable *i*.

PE method was first proposed by (Rosenblueth, 1975) - Rosenblueth point estimate method (RPEM). Later (Harr 1989) proposed another point estimation method - Harr point estimate method (HPEM), using orthogonal axis transform, to circumvent the computational burden of RPEM when there are many random variables. A brief presentation of these methods is given below further details can be found in (Rosenblueth, 1975) and (Harr, 1989). Tsai et al. (2005) give evaluation of PE methods in uncertainty analysis for environmental engineering applications. These authors recommend modified RPEM method to circumvent the draw back of non-unique solution of the original RPEM and to increase the computational efficiency in modeling.

3.3.2.1 Rosenblueth PE method (RPEM)

Rosenblueth (1975) first proposed the RPEM to deal with problems involving symmetric, correlated and uncertain input factors. Later the original method was extended to work with asymmetric random variables (Rosenblueth, 1981). The approach can be better described by

putting calculation for function of random variable with two variables. let $Z = g(X) = g(X_1, X_2)$ be a reliability function where X_1 and X_2 are two random variables, say $X_1 = S$ and $X_2 = L$. We intend to obtain the mean (μ_Z) and variance (ν_Z) of the reliability (Z). The RPEM calculation steps are:

1. Calculate the output value of Z using the reliability evaluated with the values of mean plus one standard deviation for each of the two random variables.

$$Z_{++} = g\left[(\mu_{X1} + \sigma_{X1}),(\mu_{X2} + \sigma_{X2})\right]$$

Repeat for other combinations, as follows:

$$Z_{--} = g\left[(\mu_{X1} - \sigma_{X1}),(\mu_{X2} - \sigma_{X2})\right]$$

$$Z_{+-} = g\left[(\mu_{X1} + \sigma_{X1}),(\mu_{X2} - \sigma_{X2})\right]$$

$$Z_{-+} = g\left[(\mu_{X1} - \sigma_{X1}),(\mu_{X2} + \sigma_{X2})\right]$$

2. Calculate the point-mass "weights" (Rosenblueth 1975).

$$P_{++} = P_{--} = \frac{1}{4}(1 + \rho_{12})$$

$$P_{+-} = P_{-+} = \frac{1}{4}(1 - \rho_{12})$$

Where ρ_{12} is the correlation coefficient of the random variables X_1 and X_2. The sign of ρ_{12} is determined by the sign of the product of 1 and 2; that is 1 = -, 2 = + yields (-)(+) = (-). Each weighing function P is a probability and, hence, must satisfy the axiomatic necessities that it ranges between 0 and 1 and that their sum is unity. In concept, the information (expected values, standard deviations, and correlation coefficients) relative to the random variable X_i produces four estimates $Z_{++}, Z_{--}, Z_{+-}, Z_{-+}$ of the function $Z = g(X)$ (see Figure 3-3).

3. Calculate the expectation (mean, μ_Z) of Z (Rosenblueth 1975).

$$E[g] = \mu_Z = P_{++} \cdot Z_{++} + P_{--} \cdot Z_{--} + P_{+-} \cdot Z_{+-} + P_{-+} \cdot Z_{-+} \qquad (3\text{-}27)$$

4. Calculate the variance (ν_Z) of Z, follows from the well known relationship.

$$\nu_Z = E\left[Z^2\right] - E[Z]^2 \qquad (3\text{-}28)$$

Where $E[Z^2]$ is calculated from Eq.(3-27) with the Z^2 terms substituted for the Z terms.

For FRV with three variables, say $Z = g(X) = g(X_1, X_2, X_3)$, there are eight calculations in Step 1, and the point-mass weights in Step 2 are given by (Rosenblueth 1975):

$$P_{+++} = P_{---} = \frac{1}{8}(1 + \rho_{12} + \rho_{23} + \rho_{31})$$

$$P_{-++} = P_{+--} = \frac{1}{8}(1 - \rho_{12} - \rho_{23} + \rho_{31})$$

$$P_{+--} = P_{-++} = \frac{1}{8}(1 - \rho_{12} + \rho_{23} - \rho_{31})$$

$$P_{++-} = P_{--+} = \frac{1}{8}(1 + \rho_{12} - \rho_{23} - \rho_{31})$$

The sign of ρ_{12} is determined by the sign of the product of *i* and *j*. And Eq.(3-27) extend from a summation of four terms to a summation of eight terms for this case. The variance (ν_Z) in this case can be calculated from Eq.(3-28), after first using eight Z^2 in the extended Eq.(3-27) to calculate $E[Z^2]$. From this observation a generalization can be drawn, that is if there are *n* random variables, the terms in the summation of Eq.(3-27) will be 2^n, which correspond to the total number of combinations of + and – for all *n* random variables. Therefore, a model which involves a large number of random variables would result in a very large number of combinations of function evaluations. $\mu_{X1} - \sigma_{X1}$ $\mu_{X1} + \sigma_{X1}$

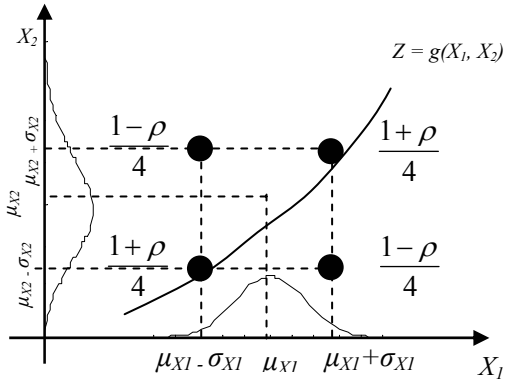


Figure 3-3. PE method schematization when $Z = g(X_1, X_2)$.

The RPEM is a computationally straight forward technique for the uncertainty analysis of engineering problems. It is capable of estimating a statistical moment of any order of a model output involving several stochastic variables that are correlated or uncorrelated, symmetric, or asymmetric. However, in multivariate problems with more than two stochastic variables involved, the RPEM is not able to provide a unique solution, rather than an approximate solution

to indeterminate problems. This is attributed to the fact that the number of unknowns to be solved is larger than the number of governing equations (Tsai et al., 2005).

3.3.2.2 Harrs PE method (HPEM)

Harr (1987) produced an alternative method (HPEM), using orthogonal axis transform, to circumvent the computational burden of RPEM when there are many random variables. In HPEM for a FRV involving *n* random variables only 2n model evaluations are required. Details of this method are presented in (Harr, 1987).

The claimed advantage of PE methods is that they overcome the difficulty with FOSM and SOSM by avoiding Taylor series expansion of FRV about the expectation of the random variables, which imposes restrictions on the FRV (existence and continuity at the first or second derivatives) and the requirement of computation of derivatives. Their limitation is accuracy because they base their calculation on only two points located based on standard deviation.

RPEM and HPEM methods have been applied for example, among others, in (Kuo et al., 2007), (Manache and Charles, 2004), (Lin and Yen, 2003), (Melching, 2001).

3.3.3 The Latin Hypercube sampling method

The Latin Hypercube sampling method (LHS) is a statistical sampling method more or less similar to the conventional MCSM. However, LHS utilizes a stratified sampling scheme instead of using random samples. Stratification divides the cumulative distribution curves of random variables into equal intervals on the cumulative scale. The LHS method was first described by (McKay et al., 1979) and further elaborated by (Iman et al., 1981). McKay (1988) proved that the LHS can achieve a convergence in performance of system more quickly than the MCSM. However, MCSM produces assumptions with the most randomness and hence will simulate real life situation best. LHS generated values more evenly and consistently across random variables distributions as a result it produces more accurate statistics (especially mean) given the same number of trials as MCSM. The computer program Crystal Ball 7.2 that is used in this research has an option for selection sampling methods to be either as MCSM or LHS. However, MCSM is exclusively used in this thesis. In terms of system resource requirement LHS requires more memory than MCSM but when MCSM requires more processor speed than LHS.

The basics of LHS can briefly be described as follows: In the context of statistical sampling, a square grid containing sample positions is a Latin square if (and only if) there is only one sample in each row and each column. A Latin hypercube is the generalization of this concept to an arbitrary number of dimensions, whereby each sample is the only one in each axis-aligned hyperplane containing it.

When sampling a function of random variable of *n* variables, the range of each variable is divided into *m* equally probable intervals. *m* sample points are then placed to satisfy the Latin hypercube requirements; note that this forces the number of divisions, *m*, to be equal for each variable. One of the advantages of this sampling method is that random samples can be taken one at a time, remembering which samples were taken so far. Figure 3-4 illustrates the difference between LH and random sampling methods.

Lin and Yen (2003) applied this in a comparative study of accuracy and efficiency of various reliability analysis methods through an example culvert. Manache and Melching (2004) applied LHS method, in combination with regression and correlation analyses, to the DUFLOW model developed for the Dender River Belgium. Kuo et al. (2007) applied LHS method together with suite of other uncertainty analysis methods in assessing dam overtopping risk.

Latin Hypercube sampling Random sampling

Figure 3-4. Comparison of LH and random sampling.

4 AN ANALYTICAL METHOD FOR TRANSFORMING pdfs

From the discussions thus far it is apparent that there is an increasing need for transparent and accurate method of quantifying uncertainty in outputs of multivariate Functions of Random Variables (FRV), such as those used in design and safety analysis of dams. This chapter discusses an analytical procedure for determining derived pdfs of RVs for quantifying randomness in outputs of those FRVs used in dam design and safety analysis. First basic theories and proofs surrounding the method are presented. Then, implementation architecture is developed and implications of the method on dam risk and safety analysis are discussed.

4.1 Definitions and Notations

Some probability theory definitions and notations are recapitulated with intent of indicating the context in this chapter. For detailed coverage on theory of probability the reader is referred to (Harris, 1966) and (Golberg, 1984).

Random Variable (X) *and Random Vector* ($\boldsymbol{X}$): X is a random variable if the probability of the event $X \leq x$ is defined for all real number x. $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ is a random vector if the probability of the event $X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n$ is defined for all real n-tuples $x_1, \ldots, x_n$. Random vectors are designated by upper case bold letters, random variables by upper case regular letters and their realizations by corresponding small letters. Section 9.1.1.3 demonstrates methods of modeling random variables using probability.

Function of Random Variables (FRV) *and Function of Random Vectors (***FRV**)*:* if X is random variable and $h(X)$ is a continuous function, then $Y = h(X)$ is a FRV. And if $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ is a random vector with n components and there is a mapping over a real coordinate space $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ yielding vector valued function $\boldsymbol{Y} = h_i(\boldsymbol{X})$, where $\boldsymbol{Y}=(Y_1, Y_2, \ldots, Y_m)$ and each component $Y_i = h_i(X_1, X_2, \ldots, X_n)$, $i = 1, 2, \ldots m$, then $\boldsymbol{Y} = h(\boldsymbol{X})$ is **FRV**.

Cumulative Distribution Function (cdf): if X is a random variable the function $F_X(x) = P\{X \leq x\}$ is a cdf (uni-variate case). The function $f_X(x) = P\{X = x\}$ is probability density function (pdf). A pdf is the first derivative of the cdf (see appendix 9.1.1.3 and Table 9.4).

Joint Cumulative Distribution Function (JCDF): if $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, for $n \geq 2$, the function $F_{\boldsymbol{X}}(x_1, x_2, \ldots x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\}$ is JCDF.

$$F_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \ldots \int_{-\infty}^{x_1} f_X(t_1, t_2, \ldots, t_n) dt_1\, dt_2 \ldots dt_n \qquad (4\text{-}1)$$

Where t_i is a set in the domain of x_i ($t_i = \{x: X_i \leq x_i\}$). The total derivative of $F_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n) = f_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n)$ is a Joint Probability Density Function (JPDF). For independent components of the random vector $\boldsymbol{X}$, $f_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$.

Marginal Cumulative Distribution Function (MCDF): if $f_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n)$ *is JPDF of random vector* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$. Marginal JCDF of k components from the n components of $\boldsymbol{X}$ can be calculated as (see also Eq.(9-10)):

$$F_{i_1, i_2, \ldots, i_k}\left(x_{i_1}, x_{i_2}, \ldots, x_{i_k}\right) = \int \ldots \int f_X\left(t_1, t_2, \ldots t_n\right) \prod_{i=1}^{n} dt \qquad (4\text{-}2)$$

Where $I = \{i_1, i_2, \ldots, i_k\}$, $0 < k < n$, be subset of $\{1, 2, \ldots, n\}$. Upper limits of integration are ∞, if $i \in I^c$; are x_i, if $i \in I$. All lower limits of integrations are $-\infty$ (I^c stands for I complement). The *Marginal Probability Density Function* (MPDF) is (see also Eq.(9-10)):

$$f_{i_1, i_2, \ldots, i_k}\left(x_{i_1}, x_{i_2}, \ldots, x_{i_k}\right) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f_X\left(x_1, x_2, \ldots x_n\right) \prod_{j \in I^c} dx_j \qquad (4\text{-}3)$$

4.2 Problem statement

In earlier sections of the dissertation it is discussed that design and safety assessment of hydraulic systems often deal with FRVs used for characterizing engineering performance. Variables forming these functions (material, hydrologic, hydraulic, structural, geophysical, environmental variables) are mostly random. Consequently, outputs of the FRVs (performance etc values) are random too. Risk-based design requires the specification of performance with an associated failure probability (P_f). It demands a comprehensive, transparent and accurate way of accounting uncertainty to the extent that makes calculation of performance exceedence probabilities (failure probabilities) possible. However, the estimation of P_f is a complex undertaking. It is complicated mainly because of the uncertainty associated with interacting design variables. It is difficult to estimate performance uncertainty arising from individual design parameters uncertainty although the individual design parameters uncertainty can be predicted with a relative ease.

The question to be resolved can be formulated mathematically as follows: if $\boldsymbol{X}$ is a random vector with n components $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and has a joint cumulative distribution $F_X(\boldsymbol{X})$; and $\boldsymbol{Y} = h(X) = (Y_1, Y_2, \ldots, Y_m)$ is a function of $\boldsymbol{X}$, whose range is a subset of m-dimensional Euclidean space for some integer m, $1 \le m \le n$ then $\boldsymbol{Y}$ is a random vector and has a cdf $F_{\boldsymbol{Y}}(\boldsymbol{y})$. We intend to compute $F_{\boldsymbol{Y}}(\boldsymbol{y})$, where both the FRV $h(\boldsymbol{X})$ and cdf $F_{\boldsymbol{X}}(\boldsymbol{x})$, are known. This problem is referred as distribution problem. In this formulation the random vector $\boldsymbol{X}$ can be compared to n random design parameters in a hydraulic systems performance function or in a design equation represented by $\boldsymbol{Y} = h(\boldsymbol{X})$. Thus, the goal is to determine randomness in performance value $\boldsymbol{Y}$ (i.e., randomness or uncertainty in outputs of the design equation $\boldsymbol{Y} = h(\boldsymbol{X})$). This is possible by determining the cdf or pdf of $\boldsymbol{Y}$ ($F_{\boldsymbol{Y}}(\boldsymbol{y})$ or $f_{\boldsymbol{Y}}(\boldsymbol{y})$).

As mentioned earlier in hydraulic design endeavors, if at all probabilistic design and safety analysis is done, distribution problems are solved using approximate methods such as MCSM, FOSM and SOSM (Hartford et al., 2004) (see also chapter 1). The use of one or more of these classical methods is demonstrated, for example, in (Joos et al., 2005), (Satoh H. et al., 2005), (Melih et al., 2004), (Lian et al., 2003), and (Negede, 2002). Such approximate methods

could be extremely valuable although they stop short of accurate analytical quantification of performance randomness. Consequently, it is important to look for a more precise, theoretically founded and practical analytical method of solving distribution problems. Such analytical procedure of evaluating performance randomness hasn't been used in dam design. This is because the profession hasn't been presented with proved and demonstrated analytical methods. This chapter tries to fill part of this gap by presenting analytical method of solving distribution problems. Subsequent chapters (chapters 6 and 7) will demonstrate applications of this method and compare its results with results from MCSM, FOSM and SOSM.

4.3 The analytical method

4.3.1 The uni-variate FRV

For continuous monotonic function h let X be a random variable with cdf $F_X(x)$ and $Y = h(X)$. We intend to drive a relation for cdf of Y, i.e. $F_Y(y)$ based on known cdf $F_X(x)$ and functional relation between $Y = h(X)$. Let A be any set in a sample space S taken from X. By definition for any set $A \in S$:

$$P\{X \in A\} = \int_A dF_X(x) \qquad (4\text{-}4)$$

Consequently, for any sub set T in the range of Y:

$$P\{Y \in T\} = P\left\{X \in h^{-1}(T)\right\} = \int_{h^{-1}(T)} dF_X(x) \qquad (4\text{-}5)$$

Eq.(4-5) can be interpreted as: the proportion of times that Y is in the set T clearly coincides with the proportion of times that X will assume any value x such that $h(x) \in T$. i.e. $P\{Y \le y\} = P\{X \le x\}$ (Figure 4-1a).

Replacing $T_Y = \{Y \le y\}$ in place of T in Eq.(4-5) yields the cdf of Y, $F_Y(y) = P\{Y \le y\}$ is:

$$F_Y(y) = \int_{h^{-1}(T_y)} dF_X(x) \qquad (4\text{-}6)$$

Apparently, it follows that, when $h(x)$ is monotonically increasing function (see Figure 4-1a), $F_Y(y) = F_X(h^{-1}(y))$ and if $h(x)$ is a monotonically decreasing function (see Figure 4-1b), then obviously $F_Y(y) = 1 - F_X(h^{-1}(y))$. Where $h^{-1}(y)$ represents the inverse of h.

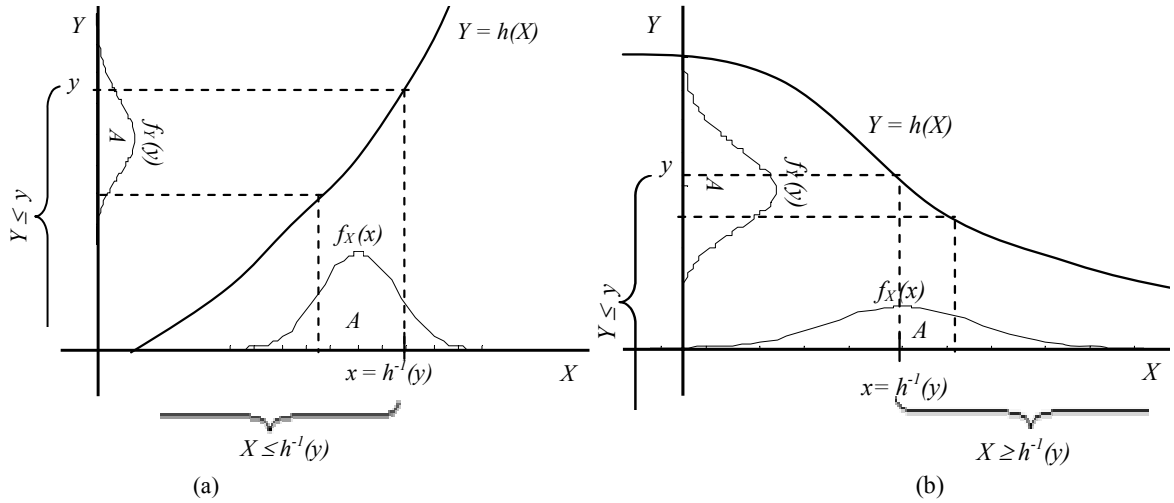


Figure 4-1. uni-variate-monotonic-FRV (a) increasing (b) decreasing.

Studding Figure 4-2 strikes thought towards a more general solution that is valid for both monotonic and non-monotonic function *h*. Let us take an arbitrary value of *h(X) = y* (horizontal line at *Y = y*). We see that there are *j* points where *h*(*X*) = *y*; i.e. *j* roots of *Y – y = 0*, *j = {1, 2, ...}*. Thus, the *X* sample space is divided into *j* disjoint intervals between successive roots ($x_{j1} = h_{j1}^{-1}(y)$, $x_{j2} = h_{j2}^{-1}(y)$). These intervals can be grouped into two: (a) Those above the line *Y = y* representing $h(x) > y \ \forall \ x \in (x_{j1}, x_{j2})$. This group contains intervals where *Y > y*, which is the complement of $F_Y(y)$, i.e. $P\{Y > y\} = 1.0\text{-}F_Y(y)$. Thus, we are not interested in this group. (b) Those below the line *Y = y* where $h(x) < y \forall x \in (x_{j1}, x_{j2})$. For intervals in this group we can write:

$$F_{Yj}(y) = F_X(x_{j2}) - F_X(x_{j1}) = F_X\left(h_{j2}^{-1}(y)\right) - F_X\left(h_{j1}^{-1}(y)\right) \qquad (4\text{-}7)$$

This group is composed of the intervals *Y < y*. The sum of probabilities of intervals in this group make up $F_Y(y) = P\{Y < y\}$. With straightforward extension of this observation a more general equation that is valid for both monotonic and non-monotonic function *h* can be written as:

$$F_Y(y) = \sum_j \left\{F_X\left(h_{j2}^{-1}(y)\right) - F_X\left(h_{j1}^{-1}(y)\right)\right\} \qquad (4\text{-}8)$$

Where $h_{ji}^{-1}(y)$, *i =1, 2* and *j = 1, 2 ...*, are roots of *Y – y = 0*. Note intervals are open since at any interval end point $h(x) = y$.

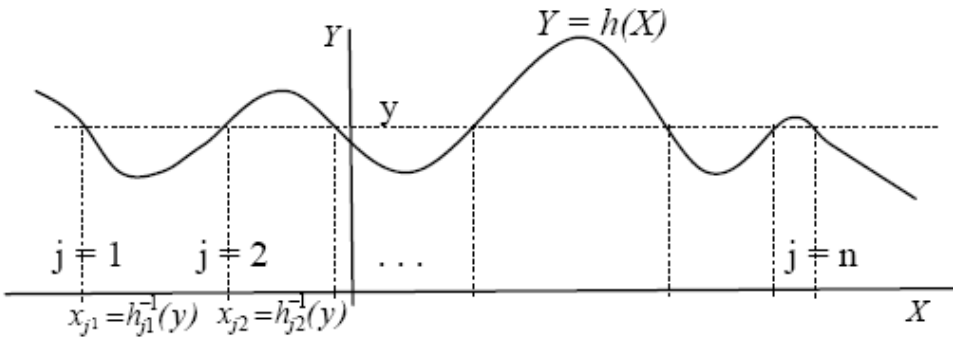


Figure 4-2. Uni-variate non-monotonic FRV

If X has a pdf $f_X(X)$, and $h_j^{-1}(y)$ has a continuous derivative for every j, then noting that $f_Y(y)=\frac{dF_Y(y)}{dy}$ and applying linearity and chain rules[1] of differentiation to Eq.(4-8) the pdf of Y can be given as:

$$f_Y(y)=\Sigma_j f_X\left(h_j^{-1}(y)\right)\cdot\left|dh_j^{-1}(y)\right| \qquad (4\text{-}9)$$

Note, while differentiating Eq.(4-8), for every j, the first term $dh_{j2}^{-1}(y)$ is always positive (because it represents a positive slope) and $dh_{j1}^{-1}(y)$ is clearly always negative. Thus, $dh_j^{-1}(y)$ can be factored out and given in absolute value.

Eq.(4-9) implies in case of non-monotonic multivariate function of random variables the function $Y=h(X)$ has to be broken up in to j segments over the intervals $\lfloor x_{2j}, x_{1j}\rfloor$, where x_{2j} and x_{1j} are roots of $h(X)$, and inverse transformations $x_j=h_j^{-1}(y)$ has to be calculated for each interval. This inverse transformations has to be replaced in place of the random variable X in its probability density function to get $f_X\left(h_j^{-1}(y)\right)$. Thus the product of this pdf of X written in terms of Y and the derivative of the inverse transformation $dh_j^{-1}(y)$ has to be computed for each interval. Adding this product of $f_X\left(h_j^{-1}(y)\right)$ and $dh_j^{-1}(y)$ gives the required $f_Y(y)$.

4.3.2 The case of multivariate FRV

In more complicated models, like those typical in design multiple random variables are involved. Such functions are made up of random vectors than random variables. Therefore, to make Eq.(4-9) applicable for multivariate FRV more generalizations are needed.

[1] Chain rule: if $f(x)=h(g(x))$ then $f'(x)=h'(g(x))\cdot g'(x)$. Diffrentiation is liniar: $(a\cdot f+b\cdot f)'=a\cdot f'+b\cdot f'$

Multivariate transformation involves calculation of determinant of a Jacobian matrix, which is part of the solution for $F_Y(y)$. And the calculation of determinants requires a full dimension (square) matrix. As a result, multivariate transformation has a requirement. The requirement is that for random vector with n components the same n number of multivariate functions should be available, i.e. we should have $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_n)$, where $Y_i = h_i(X_1, X_2, ..., X_n)$, $i = 1, 2, ..., n$. However, in most practical problems there is single function $Y = h(\boldsymbol{X})$ made up of multiple random variables. So, to fulfill the requirement, roundabout approach is needed. In the following paragraph, solution procedures are presented assuming that the requirement is satisfied. Later in this section a more general case, including those where this requirement is not satisfied, will be presented.

For multivariate functions the analogy for Eq.(4-5) is given as follows (note also Eq.(4-1) and Eq.(9-10)):

$$P\{\boldsymbol{Y} \epsilon T\} = P\{\boldsymbol{X} \epsilon h^{-1}(T)\} = \int_{h^{-1}(T_y)} ... \int dF_X(x_1, x_2, ..., x_n) \qquad (4\text{-}10)$$

Where: $h^{-1}(T) = \{\boldsymbol{X}: h(X) \epsilon T\}$. Replacing $T_{\boldsymbol{Y}} = \{\boldsymbol{Y}: Y_1 \le y_1, Y_2 \le y_2, ..., Y_n \le y_n\}$ in place of T. Let $\boldsymbol{X}$ be a random vector with cdf $F_{\boldsymbol{X}}(x_1, x_2, ..., x_n)$ and let $\boldsymbol{Y} = h(\boldsymbol{X})$ be an n-dimensional FRV then cdf of $\boldsymbol{Y}$ is:

$$F_Y(y_1, y_2, ... y_n) = \int_{h^{-1}(T_y)} ... \int dF_X(x_1, x_2, ..., x_n) \qquad (4\text{-}11)$$

Therefore, Eq.(4-6) can be restated to include all continuous functions:

$$F_Y(y_1, y_2, ..., y_n) = \sum_j \int_{h_j^{-1}(T_y)} ... \int dF_X(x_1, x_2, ..., x_n) \qquad (4\text{-}12)$$

Where $h_j^{-1}(T_y)$ refers to inverse transformation families $X_1 = h_{1j}^{-1}(Y_1, Y_2, ..., Y_n)$, $X_2 = h_{2j}^{-1}(Y_1, Y_2, .., Y_n)$, …, $X_n = h_{nj}^{-1}(Y_1, Y_2, ..., Y_n)$. The analogy for Eq.(4-9) can also be stated as follows: for $\boldsymbol{X}$ be a continuous random vector with pdf $f_{\boldsymbol{X}}(\boldsymbol{X})$ and for $x_i = h_{ij}^{-1}(y_1, y_2, ..., y_n)$, $\partial h_{ij}^{-1} / \partial y_k$ is continuous for all i, j, k and that each of the Jacobians J_j of the inverse transformations do not vanish identically:

$$f_Y(y) = \sum_j f_X\left(h_{1j}^{-1}(y_1, y_2, ..., y_n), h_{2j}^{-1}(y_1, y_2, ..., y_n), ..., h_{nj}^{-1}(y_1, y_2, ..., y_n)\right) \cdot |J_j| \qquad (4\text{-}13)$$

The absolute value of the Jacobians J_j is given in terms of determinant of families of inverse transformations ($h_{ij}^{-1} = X_i = h_{ij}^{-1}(Y_1, Y_2, ..., Y_n)$) as:

$$J_j = \begin{vmatrix} \frac{\partial h_{1j}^{-1}}{\partial y_1} & \frac{\partial h_{1j}^{-1}}{\partial y_2} & \cdots & \frac{\partial h_{1j}^{-1}}{\partial y_n} \\ \frac{\partial h_{2j}^{-1}}{\partial y_1} & \frac{\partial h_{2j}^{-1}}{\partial y_2} & \cdots & \frac{\partial h_{2j}^{-1}}{\partial y_n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial h_{nj}^{-1}}{\partial y_1} & \frac{\partial h_{nj}^{-1}}{\partial y_2} & \cdots & \frac{\partial h_{nj}^{-1}}{\partial y_n} \end{vmatrix}$$

Proof for Eq.(4-13) is provided in the appendix 9.2. For details on multivariate probability theory the reader is referred to (Harris 1966; Golberg 1984).

In risk analysis frequently real-valued function of random vector ***X***, or possibly vector-valued function whose range is in Euclidean *m*-space, $m < n$ are encountered. Such cases do not satisfy the requirement for equal number of variables and functions (they do not give a full dimension Jacobian matrix). But, introducing $n - m$ dummy functions Eq.(4-13) allows solving this type of problems as well.

For such cases of $m < n$ Eq.(4-12) applies in similar fashion with no changes except for the dimensionality of the range of transformation. In this case there are only *m* inverse transformations $Y_i = h_i(X_1, X_2, ..., X_n)$, where $1 < i < m < n$. However, by introducing $n - m$ dummy functions an auxiliary transformation function $\boldsymbol{Y}^{\circ} = (Y_1, Y_2, ..., Y_m, Y_{m+1}, ..., Y_n)$ can be defined so that the requirement $m = n$ is true, i.e. so that the Jacobian matrix is of full dimension. Then the pdf $f_Y(y_1, y_2, ..., y_m)$ can be given as the marginal probability density function of the first *m* components of $\boldsymbol{Y}^{\circ}$; i.e. $f_{\boldsymbol{Y}}(y_1, y_2, ..., y_m)$ can be computed by integrating $f_{\boldsymbol{Y}}(y_1, y_2, ..., y_n)$ over entire ranges of the introduced *n-m* dummy functions. Therefore, for the pdf $f_{\boldsymbol{Y}}(y)$ can be written as:

$$f_Y(y) = \sum_j \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(f_X h_{1j}^{-1}(y_1, y_2, ..., y_n), h_{2j}^{-1}(y_1, y_2, ..., y_n), \ldots, h_{nj}^{-1}(y_1, y_2, ..., y_n) \right) \cdot |J_j| \, dy_{m+1} dy_{m+2} \ldots dy_n \qquad (4\text{-}14)$$

The procedure for solving Eq.(4-13) and Eq.(4-14) is clearer when seeing the implementation architecture given in Figure 4-3. In addition, the solution procedures are demonstrated step wise using a case study in chapter 6.

4.4 Application procedures and requirements

In Figure 4-3 the analytical solution implementation architecture is provided in a self explanatory manner. For the application of ASDD method three requirements has to be fulfilled. First we will require that *h(**X**)* be a continuous differentiable function. In more advanced discussions, however, this condition will be replaced by 'measurability', but the discussion of such conditions is beyond the scope of this research. For further details on such conditions the reader is referred to (Harris, 1966). However, this is not a limitation for most engineering design equations with well bound domains for parameters. Second, a good understanding of the physics

behind performance of dams is required for expressing the system with FRV (mathematical model) indicating load strength interactions. This is also possible for equations dealing with most of the dominant failure mechanisms. When there are doubts a provision could be given to account for model uncertainty. Third, it is required that each random variable be specified with standard pdfs. This involves choosing the best-fit distribution based on goodness-of-fit test and judgment conducted using limited engineering data, which is usually the case in dam engineering. Goodness-of-fit test topic is well treated in many classical statistics and reliability engineering books, like in (D'Agostino et al., 1986), (Benjamin et al., 1970, (Crowder et al., 1991), and (Lawless, 1982). One may use formal goodness-of-fit tests; for example Chi-square (χ^2) test or graphical methods (see also appendix 9.1.3). In this research χ^2 and Kolmogorov-Smirnov methods are used.

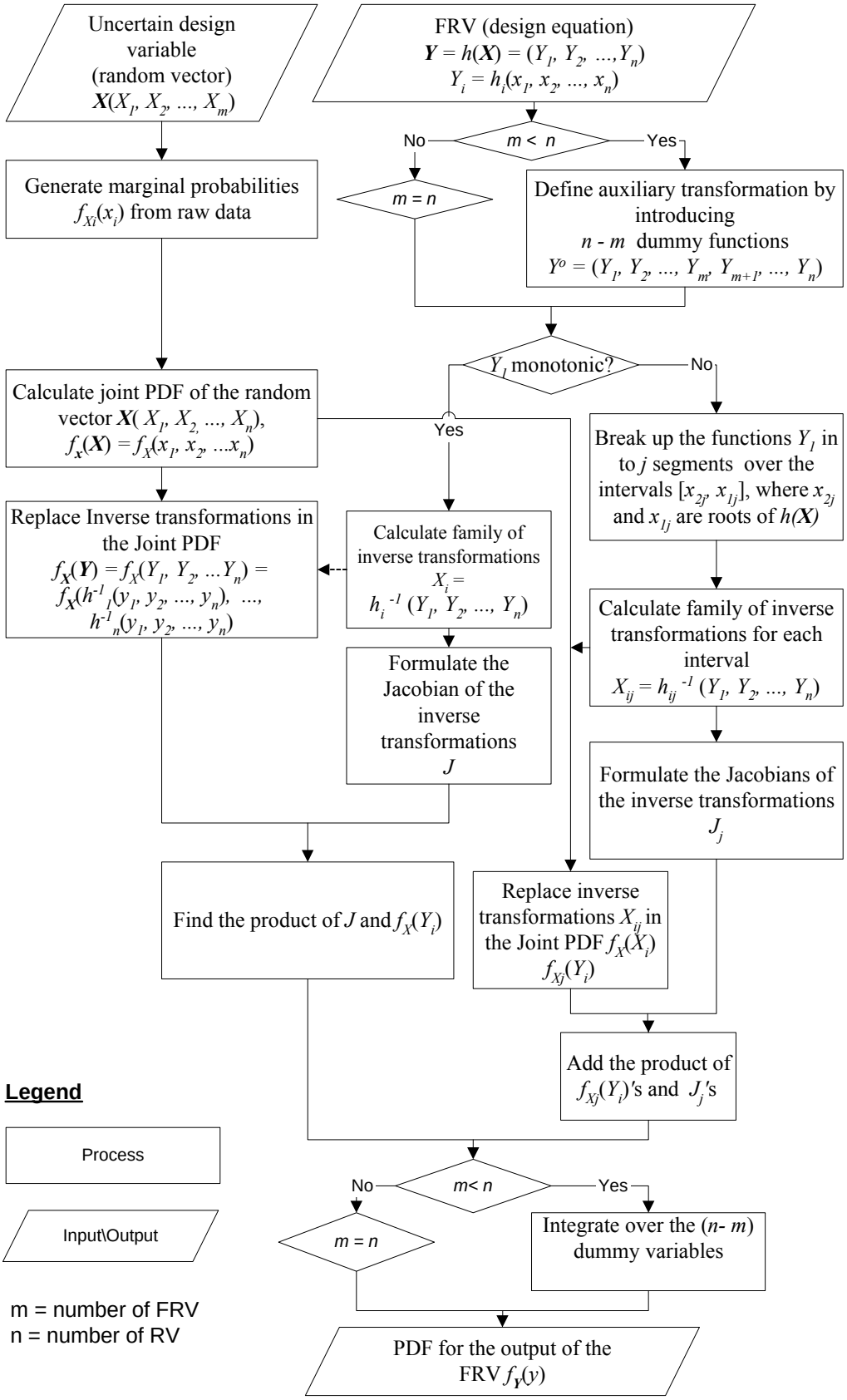


Figure 4-3. ASDD method implementation architecture.

4.5 Chapter conclusions

The analytical method is applicable for probabilistic design, risk and safety analysis of dams. These are problems that in most cases involve multivariate FRV. The analytical method is an exact method provided selected distributions are cable of modeling the random variables efficiently (data adequacy) and provided the FRV used to describe the physics of the system is adequate (model adequacy). It is suited to being programmed using symbolic programming language and it lends itself for easy adoption in analytically defined reliability functions, safety analysis and many other practical applications. The alleged difficulty of this method is its cumbersome computational requirements. This alleged difficulty could be reduced significantly through the use of software like Mathematica 5.2. If data and model adequacy are assured this method out perform the classical methods discussed in chapter 3 (MCSM, FOSM and SOSM methods) with its accuracy, transparency and reproducibility. However, one has to note that problem of data and model adequacy is equally shared by the other methods too. It is not a unique problem associated with the ASDD method. The topic presented in this chapter is assumed to bridge the gap between the available analytical probability theories and the suffering distribution problem in risk and safety analysis of dams.

5 DAM ENGINEERING AND THE CASE STUDY DAM: PROBABILISTIC PERSPECTIVES

5.1 Review on major causes of dam failure and their statistics

5.1.1 Types of dams

Dams are numerous in type classified usually in terms of materials used for their construction and their form. Common types are homogeneous or zoned earthfills; rockfills with earth core or concrete face; and concrete dams that depend on gravity, arch, or buttress resistance. Some dams are composites of various materials, including earthfill, rockfill, masonry, and concrete. A few have timber, asphaltic, or synthetic members. Topography, geology and availability of construction materials and technology are primary factors in weighing the comparative merits of dam types. Novak et al. (2003) give an initial broad classification of dams into two generic groups based on the principal construction material employed.

1. *Embankment dams*: are constructed of earthfill and/or rockfill; upstream and downstream face slopes are similar and of moderate angels, giving wide section and high construction volume relative to height.

2. *Concrete dams*: are constructed of mass concrete; face slopes are dissimilar, generally steep downstream and near vertical upstream, and dams have relatively slender profile. This group can be considered to include also older dams of appropriate structural type constructed in masonry.

Novak et al. (2003) also identifies the principal types within the two generic groups (see Table 5.1). The dam for the case study in this dissertation is a zoned earth fill dam.

Table 5.1: Types of dams and register statistics (ICOLD, 1988a, in Novak et al., 2003).

Group	Type	Percent from total number of constructed large dams
Embankment dams	earth fill rockfill	82.9
Concrete dams (including masonry)	gravity arch buttress multiple arch	11.3 4.4 1.0 0.4
Total large dams[2] (ICOLD, 1988a)		36235

[2] Large dams are defined by ICOLD as dams exceeding 15 m in height or, storage volume in excess of 10^6 m^3 or a flood discharge capacity of over 2000 m^3 s^{-1}. Based on this definition the case study dam in this research is categorized as a large dam with its height of 53 m and storage capacity of $1.86 \cdot 10^9$ m^3.

5.1.2 Major causes of dam failure and their statistics

There are varying statistics on causes of dam failure, for example statistics given by International Commission on Large Dams (ICOLD), US Army Corps of Engineers (USACE) and Novak et al. (2003) are given in Table 5.2. Many attempts have been made at compiling and assessing statistics on dam failure. Main attempts on worldwide scale have been made by International Commission on Large Dams (ICOLD) in 1974, 1983 and 1995. ICOLD (1995) states that foundation problems are the most common causes of failure in concrete dams, with internal erosion and foundation shear strength each contributing for 21%. In case of earth and rockfill dams, the most common cause of failure is overtopping (31% as primary cause and 18% as secondary cause) followed by internal erosion in the body of the dam (15% as primary cause 13% as secondary cause), and in the foundation (12% as primary cause and 5% as secondary).

Table 5.2: Statistics on causes of dam failure.

Source	Overtopping	Foundation defects*	Internal erosion**	Others
USACE (2006)	34%	30%	20%	6%
Novak et al. (2003)	30-35%	No data	30-35 %	
ICOLD (1995)	31% primary cause 18 % secondary cause	No data	27% primary cause 18 % secondary cause	

* Slope instability, differential settlement, high uplift pressure, foundation seepage
** Piping and seepage

5.2 Embankment dams failure mechanisms and design practices

It is comprehensible that the degree of importance of different causes of dam failure varies with dam type. Dam design principles and considerations evolved with the identification of major causes of dam failure and the progressive understanding of their mechanisms. In addition, knowledge on causes of dam failure is crucial for dam safety evaluation, dam monitoring and rehabilitation decisions. The following paragraphs provide a summary on the major causes of dam failure and their mechanisms.

Earthfill embankments may be damaged by distortions at critical points. Differential settlement may be severe at steep abutments and at structural interfaces where effective compaction is difficult to obtain. At these locations, deformation of the fill may open dangerous paths of seepage. For this reason, there have been many failures along outlet conduits. Although properly constructed embankments are able to accommodate substantial movement, they have relatively poor resistance to overflow; so their freeboard and associated spillway capacity must be determined conservatively.

In contrast, most concrete dams can withstand overtopping for at least several hours. The key to their safety may be the resistance of the foundation to impact of the spill. Essential criteria governing the structural competence of concrete dams are the margin of safety against overall structural stability (this includes safety against rotation and tipping of the dam; and translation and sliding of the dam body and natural rock foundation) in relation to all probable conditions of loading including empty reservoir condition. Moreover, there should not be over stress and material failure in the dam concrete and the rock foundation.

Arc dams can carry large loads, but their integrity depends inherently on strength of the abutments. Failure may be caused by rock deteriorations or by shearing under water pressures. Weakening of arch support also may be triggered by foundation erosion. Gravity dams are noted for durability because of their large masses, they can survive considerable weathering and site deficiencies. However, sometimes some have failed where foundation elements were susceptible to sliding. A few buttressed dams also have shown this tendency.

Novak et al. (2003) identifies the following principal defect mechanisms and failure modes for embankment dams:

1. *Overtopping leading to washout*: spillway and outlet capacity must be sufficient to prevent overtopping. Also there should be sufficient freeboard to prevent overtopping by wave action. The freeboard must also include an allowance for the predicted long-term settlement of the embankment, foundation compressibility and sedimentation. Overtopping has risk of serious erosion and possible washout of embankment.

2. *Internal erosion and piping with migration of fines from core* and *foundation*: regression of 'pipe' and formation of internal cavities, may initiate by internal cracking or by seepage along culvert perimeter. Seepage within and under the embankment must be controlled to prevent concealed internal erosion and migration of materials. Hydraulic gradients, seepage pressures and seepage velocities within and under the dams must, therefore, be contained at levels acceptable for the materials concerned. Care must be taken to ensure that outlet or other facilities constructed through the dam do not permit unobstructed passage of seepage water along their perimeters with risk of soil migration and piping.

3. *Embankment and foundation settlement (deformation and internal cracking)*: care must be taken with soft compressible foundations and proper compaction has to be done during construction of dams.

4. *Instability*: the embankment, including its foundation, must be stable under construction and under all conditions of reservoir operation. Instability might occur when downstream slope too high and/or too steep in relation to shear strength of the shoulder material or when there is rapid drawdown of water level or because of failure of downstream foundation due to overstress of soft, weak horizons. Face slopes must, therefore, be sufficiently flat to ensure that internal and foundation stress remains within acceptable limits under different conditions of loading. In this regard, the following loading and critical conditions must be analyzed:

    a. End of construction (both slopes critical);
    b. Steady state, reservoir full (downstream slope critical);
    c. Rapid drawdown (upstream slope critical);
    d. Seismic loading condition additional to 1, 2 and 3, if appropriate to the location

Major failure modes of different types of dams and appurtenant structures are often known and they shall be included in risk analysis of dams. For instance, Table 5.3 provides the common categories of failure modes for the different types of dams and appurtenant structures. Section 5.2.1 and 5.2.2 present brief summary of dam design practices against the two major causes of dam failure (sliding and overtopping). In these two sections a brief account of the

design practices is given. This puts the subsequent two chapters (chapters 6 and 7) in context and facilitates smooth understanding. Further, detail coverage on these design practices can be found from classical hydraulic structures or dam design text books, for instance (Novak, P. et al., 2003), (Senturk, 1994), (Jansen,1988), (Sherard et al., 1963).

In this dissertation two case studies are devoted for demonstrating the suggested probabilistic methods: one on flood-waves and reservoir safety evaluation (failure due to overtopping) and another on safety against sliding evaluation (failure due to sliding). These two failure mechanisms are selected due to their world wide significance in recorded dam failure statistics, each accounting to about a third of recorded dam failures world wide (refer section 5.1).

Table 5.3: Common causes of dam and appurtenant structures failure.

Earthfill dams	Rockfill dams	Concrete Dams	Spillways
– seepage and piping (foundation and dam body), – slop instability (sliding), – breach due to overtopping, – upstream face erosion due to waves, – cracking, – settlement	– leakage, – eroding and cracking of upstream concrete face/membrane, – settlement and translation, – piping through the core zone, – cracking of the core zone, – loss of freeboard due to excessive settlement, – slope instability (sliding).	– Leakage through foundation, – overtopping and downstream foundation erosion, – translation and sliding of the dam body and natural rock foundation, – cracking, weathering, concrete deterioration.	– deficient capacity (hydrologic-underestimation of peak flood) – hydraulic (failure to accommodate high energy condition), – structural (deterioration of flow surface, inadequate structural capacity, deficiencies in surface tolerance to preclude cavitation), – cavitation and abrasion, – excessive uplift pressure, – gate malfunctioning, – inadequate capacity of downstream and approach channel.

In the following two sub sections the widely adopted (deterministic) design practices and standards with regard to the two failure mechanisms are presented. This will help to later compare results from the adopted probabilistic analysis with conventional deterministic standards and design out puts.

5.2.1 Classical design practices and standards for slope stability

In earthfill construction it is necessary to consider the load-bearing characteristics of the compacted fill and also the behavior of the soil as construction proceeds. Problems related to the response of soils to specific loading conditions are generally grouped into two: *problems of deformation and problems of stability*. The problem of deformation deals with settlement and consolidation. A soil mass may undergo deformation as a result of changes in external loading – called settlement, and/or due to own weight of a compressible soil and changes in drainage conditions – called consolidation. A limited amount of deformation occurs with no net volume change, and is thus comparable with elastoplastic behavior of many non-particulate materials. The most significant soil deformation, however, usually involve volume changes arising from

alterations in the geometric configuration of the soil particles assemblage, e.g. loosely packed arrangement of soil particles will on loading adopt a more compact and denser structure. Such change occurs almost immediately on load application where the soil structure is relatively coarse, as with sands. In saturated clay soils, however, volume changes and settlement due to external loading will take place slowly through complex hydrodynamic process known as consolidation. Problems of settlement and deformation analysis are not part of the case study in this thesis and therefore no further discussion is given on this topic.

Problems of stability concern the equilibrium between forces and moments and the mobilized soil strength. When the forces and moments arising from loading (or from the removal of support as in a trench excavation), exceed the shear resistance which the soil can mobilize, failure will occur. Such a failure is generally manifested by progressive and, in the final phase, large and relatively rapid mass displacements (Novak et al., 2003). Therefore, stability problems involve concepts of soil shear strength and stress-strain response.

The strength and stress aspects related are discussed in the following paragraphs. Uncertainty associated to soil strength and stress parameters is discussed in 5.5. Chapter 6 presents a case study on probabilistic analysis of stability.

5.2.1.1 Total and effective shear strength

Stability depends on the balance between the resistance to shearing which can be mobilized, i.e. the shearing strength of the soil, and the shearing stress resulting from the principal loads. The shear strength of a soil is defined as the maximum resistance to shearing stress which can be mobilized, when this is exceeded failure occurs, usually along identifiable failure surfaces. Soil shear strength is commonly quantified through two component parameters:

a. *Apparent cohesion (c)*: essentially arising from the complex electrical forces binding clay-size particles together;

b. *Apparent Angel of shear resistance* (ϕ): developed by interparticle frictional resistance and particle interlocking.

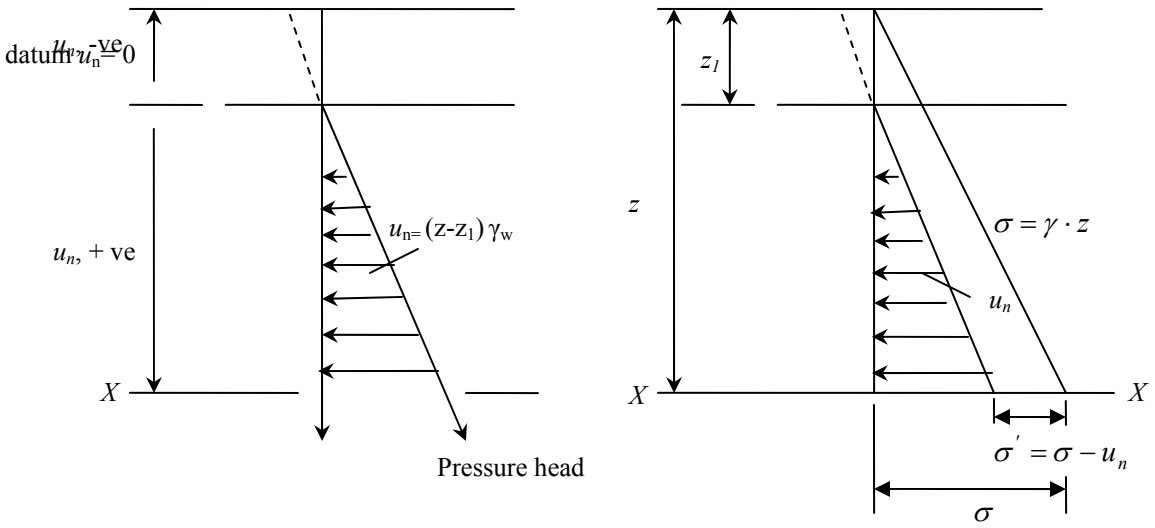
The shear strength (the maximum resistance to shearing) of a soil at a point on a particular plane can be expressed using the Mohr-Coulomb failure criteria as a linear function of the normal stress (σ) at that same point:

$$\bar{\tau} = c + \sigma \cdot \tan\phi \qquad (5\text{-}1)$$

Where $\bar{\tau}$ (kN/m^2) is shear strength at failure, σ (kN/m^2) is normal stress, c (kN/m^2) is apparent cohesion and ϕ (degrees) apparent angle of shear resistance.

A soil in embankments may constitute a two- or three-phase system comprising solid soil matrix and fluid, either air or water or both. Provided that water is present in the soil pores as a continuous liquid phase, Bernoulli's laws apply. That do mean hydrostatic pressures (porewater pressures) exists and it varies with the moisture content and boundary conditions, and most

importantly stress characteristics are also influenced with the level of this pressure. Silty soils and clays frequently employed in embankment fills are generally non-saturated when first compacted, i.e. some pore space is filled with compressible pore air. But, as reservoir is filled with water and the seepage front advances through the embankment the fill will progress to saturated state. In dam engineering much of soil mechanics practice is prescribed assuming this ultimate saturation condition. Therefore, stress and strength at a point in a body of the earth fill is determined from the combined effects of the solid soil matrix and the porewater. This is illustrated in Figure 5-1. Figure 5-1 shows a vertical section through a soil mass generating a vertical total stress (σ) and static porewater pressure (u_n) on the horizontal plane *X-X* at depth *z*.



Static ground water profile Vertical stress profile

Figure 5-1. Porewater pressures and vertical geostatic stress assuming static groundwater (after Novak et al., 2003).

Positive porewater pressure below the water table decreases interparticle contact pressure and the thus it decreases intergranular (effective) stress (σ'), transmitted through the soil particles. Effective stress (σ') is less than the total stress (σ) by an amount equivalent to the porewater pressure u_n, as given in Eq. (5-2).

$$\sigma' = \sigma - u_n \qquad (5\text{-}2)$$

The effective stress relationship is at the core of much of geotechnical practices because it determines the shearing resistance that a soil can mobilize and its compressibility.

In embankment dams high porewater pressures are generated by changes in external loading conditions, including construction and operations, and are very slow to dissipate. This is especially significant with fine cohesive soils in embankment dam zones. This characteristic is attributed to the nature of the clay particles, which have low permeability and strong affinity for water. Natural clay soils are usually in a saturated or nearly-saturated state.

As shown in Eq. (5-2) and Figure 5-1 there is a clear relationship between shear strength and increasing water content: at high water contents the cohesive forces between clay particles rapidly weaken, resulting in very much reduced shear strength. Therefore, in view of the importance of effective stress in controlling the soil behavior under saturated conditions, shearing resistance shall be determined from effective stress (σ'), i.e. interparticulate, rather than total stress level (σ). As a result a more appropriate form of Eq.(5-1) can be given as follows:

$$\bar{\tau}' = c' + \sigma' \cdot \tan\phi' \qquad (5\text{-}3)$$

Where $\bar{\tau}'$ (kN/m^2) is effective shear strength at failure, σ' (kN/m^2) is effective normal stress, c' (kN/m^2) is effective cohesion and ϕ' (degrees) is effective angle of shear resistance.

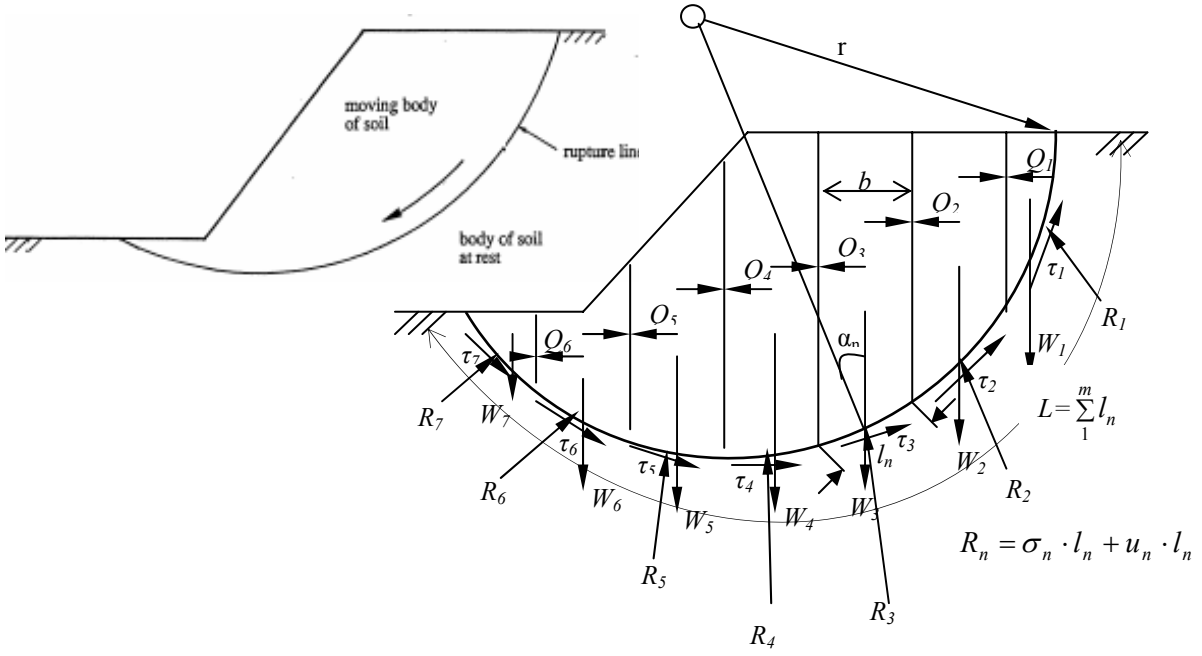


Figure 5-2. Embankment sliding and stability analysis using method of slices - formulation.

5.2.1.2 Slope stability and factor of safety against sliding

When a soil structure such as dam slope and foundation fails, there is a distinct movement of body of soil relative to another body of soil. The zone between these two bodies is a failure zone, which can be approximated by a curve without thickness. This curve is usually assumed to be a circle (see Figure 5-2). Stability analysis based on such an assumption is called slip circle method. This method takes stability analysis as a two-dimensional limit-equilibrium problem, which can be solved based on consideration of static equilibrium of potentially unstable 'active' mass of soil overlaying a circular conjectural failure surface. It can be solved by analytical or graphical methods. In both methods essentially the analysis applies the same principles of equilibrium though the techniques are different. In the graphical method the total sum of forces is considered; there is equilibrium (no sliding failure) if this sum equals zero. In the analytical

methods the overturning moment is compared with resisting moment. For that purpose the slice method is applied (see Figure 5-2 and Figure 5-3). There is equilibrium if the overturning and resisting moments are equal. The analytical methods are more convenient for use in probabilistic design and risk analysis as they lend themselves to applications using computer program. In Figure 5-2 and Figure 5-3 the elements needed for application of analytical slope stability analysis method are given.

When using analytic slope stability analysis methods in risk analysis the overturning – resisting moment equations have to be formulated as a reliability equation. In reliability terms the overturning moment can be taken as load and the resisting moment as resistance. The system fails only when the latter is less than the former. Based on this concept in the classical deterministic approach a global factor of safety against sliding ($F_{s,\,s}$) is defined as a ratio of the effective unit shear resistance which can be mobilized to the unit shear stress:

$$F_{s,s} = \frac{\sum \bar{\tau}'}{\sum \bar{G}'} \qquad (5\text{-}4)$$

Where $\bar{\tau}'$ and $\bar{G}'$ are, respectively, the effective unit shear resistance which can be mobilized and unit shear stress generated on the failure surface. The effective stress (gravitational driving force) ($\bar{G}'$, kN/m^2) for an active soil mass of unit width above a failure surface of length L (m) can be given as:

$$\bar{G}' = W \cdot sin\alpha \qquad (5\text{-}5)$$

Where, W is the weight of the active mass of unit width above the failure surface length L (m) and α is the angle of inclination of the slip surface to the horizontal. The situation is schematized in Figure 5-3 assuming the active mass is divided in to slices to facilitate the equilibrium analysis.

The expression for $F_{s,s}$ that correspond to the most commonly employed analytical methods is given as either the Swedish Circle (Fellenius) method Eq.(5-6) or as Alan Bishop semi-rigorous solution Eq.(5-7). The difference between these two methods is the assumption made with regard to the interslice geostatic and porewater forces, which are represented by Q_i, $Q_{i+1}, \ldots, Q_n$ in Figure 5-2, required for static equilibrium. In Fellenius method it is assume that interslice forces are horizontal (normal forces) at either side of a slice having equal magnitude and opposite direction ($Q_n' = Q_{n+1}'$). This means they cancel out each other. This assumption is equivalent to ignoring all (normal and shear) interslice forces. However, Bishop's method assumes that Q_n' and Q_{n+1}' are both horizontal forces having different magnitude. Both methods ignore existence of interslice shear forces and satisfy moment equilibrium.

$$F_{s,s} = \frac{\sum_1^m c' \cdot l_n + tan\phi' \cdot \sum_1^m \left(W_n \cdot cos\alpha_n - u_n \cdot l_n\right)}{\sum_1^n W_n \cdot sin\alpha_n} \qquad (5\text{-}6)$$

Where W_n and l_n are, respectively, the weight and base length of the slices into which the active mass is subdivided for analysis, α_n is the angle of inclination of the slice base to the horizontal, u_n is the porewater pressure at the slice base and $L = \sum_1^m l_n$ is the overall length of the failure surface. Other, variables as defined earlier.

$$F_{s,s} = \frac{\sum_1^m \left(c' \cdot b_n + \left(W_n - u_n \cdot b_n \right) \cdot tan\phi' \right) \cdot \frac{1}{m_\alpha}}{\sum_1^n W_n \cdot sin\alpha_n} \qquad (5\text{-}7)$$

In which,

$$m_\alpha = \left(1 + \frac{\tan\alpha_n \cdot \tan\phi'}{F_s} \right) \cdot \cos\alpha_n$$

Thus, using either Eq. (5-6) or Eq.(5-7) stability analysis (computation of factor of safety) is applied to all conceivable failure surfaces, and the supposed minimum factor of safety F_{min} is sought. In deterministic design the calculated minimum factor of safety has to always be compared with the required factor of safety to see if the conditions of design are satisfactory. Table 5.4 gives the commonly used deterministic guidelines on acceptable factor of safety against sliding.

In Bishop's method the factor of safety $F_{s,\,s}$ appears on both sides of the equation and iterative solution is required. To avoid this iterative procedure the probabilistic stability analysis in chapter 6 employed Fellenius method, i.e. Eq.(5-6).

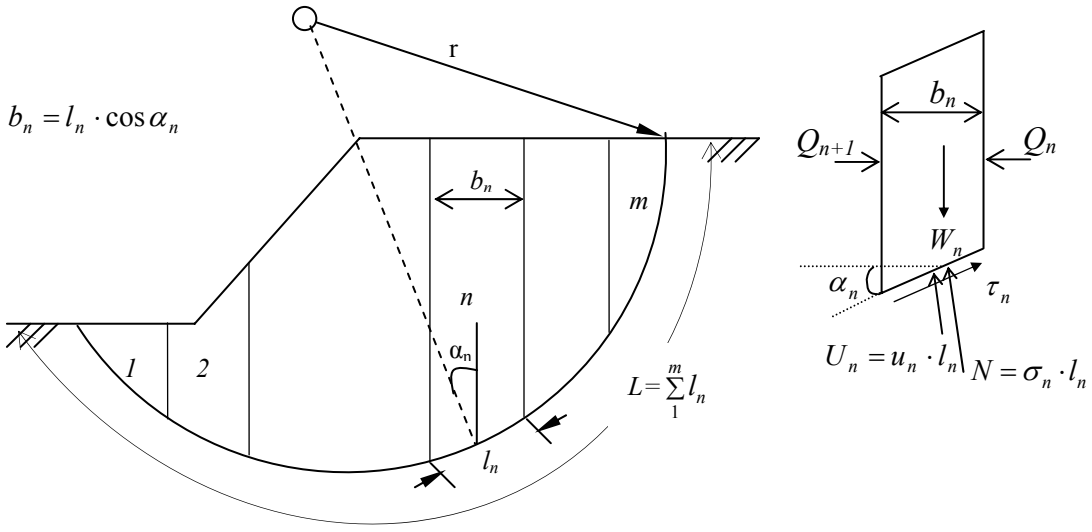


Figure 5-3. Slope stability analysis using analytical method (method of slices) – elements.

For additional reference on Fellenius and Bishop methods the following books and articles could be referred Novak et al. (2003), Bishop (1955), Bishop and Bjerrum (1960), Bishop and Morgenstern (1960), Jansen et al. (1988).

Table 5.4: Deterministic effective stress stability analysis $F_{s,\,s\,min.}$ guidelines after Jansen R. et al. 1988 (values in bracket from Novak et al. 2003 giving different guidelines for $F_{s,\,s\,min}$).

Design loading condition	Minimum Factor of safety, $F_{s,\,\min}$	
	Downstream slope	Upstream slope
1. under construction; end of construction	1.25	1.25
with earthquake loading in addition (pseudo static)	1.0 (1.1)	1.0 (1.1)
2. Steady seepage long-term operational at partial pull, upstream slope	1.5	1.5
with earthquake loading in addition (pseudo static)	1.25	1.25
3. Rapid drawdown	-	1.25 (1.2)
with earthquake loading in addition (pseudo static)	-	1.0 (1.1)
4. Seismic loading with 1,2, or 3 above	1.1	1.1

The expressions for determining $F_{s,\,s}$ are of varying rigor and inexact, a reflection of the complexity of the stability problem, where measured shear strength parameters can be subject to a variance of up to 30 - 40% (Novak et al., 2003). This already tells that an attempt to solve stability problem using classical deterministic approaches is simply unrealistic. Classically deterministic stability designs require to accept relatively high values of $F_{s,\,s}$ and to assume 'absolute safety' in a naïve way; whilst the level of safety provided remains uninnumerable. The entire stability analysis precision, among others, depends on:

- The accuracy of measured shear strength parameters. There is an element of in accuracy in the way ϕ' and c' are determined. It is impossible to take fully undisturbed samples and also during handling and testing of samples disturbance might take place. Apart from taking every precaution to ensure a reliable test result, it will be necessary to take a sufficient number of samples and treat the results in a responsible statistical and probabilistic way;
- The inaccuracy that occurs in the calculation methods as approximations are not flawless, as is amply demonstrated by the fact that various methods give different factors of safety for the same profile and soil conditions. It should be born in mind that the stability calculation methods are empirical although they are not void of certain theoretical elements;
- Limitations of limit-equilibrium method. The non-linearity at lower pressures is not well represented in this method. As Mohr-Columbs criteria considers only the maximum shear strength-stress relations;
- The induced geometric errors in calculating slip surfaces shape and area, volume of slices, and direction of action of forces. The form of the critical failure surface is controlled by many factors, including soil type and presence of discontinuities or interfaces;

- The element of uncertainty about the loading conditions such as more intensive earthquakes than taken into account in the design (earthquake is inherently uncertain phenomenon), excessive rainfall in combination with failures in the drainage system, unexpected loading etc.

Moreover, as discussed above stability is very sensitive to porewater pressure. Thus dam stability must be assessed in relation to the changing conditions of loading and seepage regimes which develop from construction through the first impoundment into operational services, including reservoir drawdown. Figure 5-4 shows the variation of embankment stability parameters during construction and operation. Thus, in deterministic design to partly capture variability it is customary to analyze the following critical conditions:

- End of construction (both slopes: upstream and downstream),
- Steady state, reservoir full (downstream slope critical),
- Rapid drawdown (upstream slope critical, usually caused by a failure of secondary dam),
- Seismic loading (both slopes)

While the stress at a point may relatively be readily determined, the local porewater pressure is more complex variable. Fine grained clay-type soils value of u_n for applied increments of total stress will depend upon properties of soil mineral skeleton and is strongly time dependent (Novak et al. 2003). Among many other factors the phreatic level fluctuation, which intern is a parameter that depends on the free water surfaces levels upstream and down stream of a dam and the soil permeability, influence the pertinent porewater pressure.

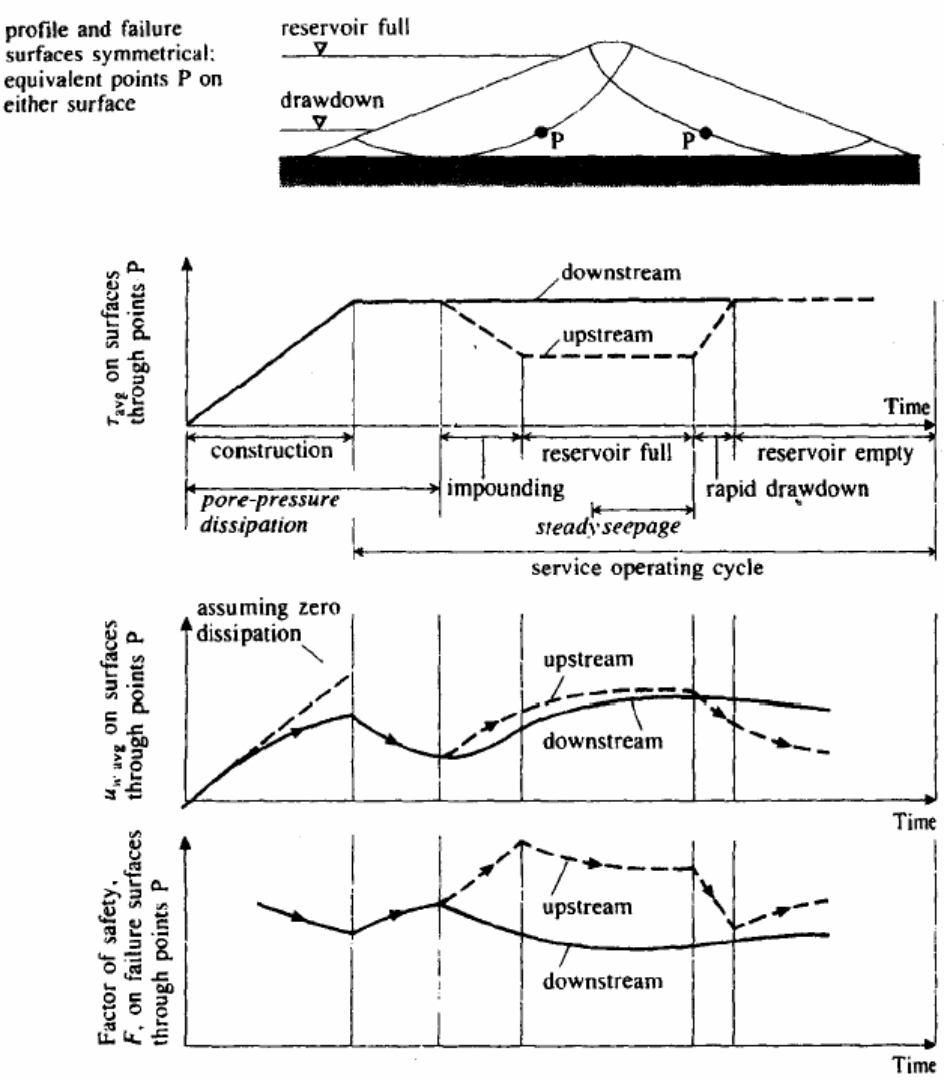


Figure 5-4. Variation of embankment stability parameters during dam construction and operation (after Bishop and Bjerrum (1960) also in Novak P. et al. (2003), pp 80).

From the discussions above it is very apparent that it is difficult to catch all these uncertain elements in one deterministic factor of safety. Consideration of soil properties, pore pressure and loading conditions, execution of works and construction activities in their full randomness allows seeing the real randomness of the safety (reliability) of dam against failure due to sliding. This fact is abundantly made clear using a case study presented in chapter 6. In chapter 5.5 typical randomness in values of physical parameters used in the process of stability analysis is presented using data from the case study dam.

5.2.2 Classical design practices and standards for flood and wave protection

Dam overtopping failure is defined as the case where the water level or individual waves exceed the highest water tight level of a dam. There are several standards that are in use for design of dams against flood and wave overtopping. Much of these standards are mostly

concerned with setting criteria for selecting peak spillway design flood magnitude (Q_P). Peak design flood represents the flood inflow which must be discharged under normal conditions with a safety margin provided by an accepted freeboard limit. The acceptable freeboard limit is set through iterative reservoir routing computations made using the selected design flood and a spillway capacity.

The selection of 'safe' design flood is based on estimation of annual exceedence probability that corresponds to an acceptable level of flood risk with respect to human and economic consequences of failure. For example, where an area is heavily populated and/or developed industrially and the failure of the dam will result in loss of life and great property damage, a design flood of very low probability calculated by different methods (such as flood frequency analysis on historical records of maximum observed floods, flood envelop curves, empirical and regional formulae, modern methods of rainfall-runoff analysis) will usually be justified. The highest standard uses the probable maximum flood (PMF), i.e. the extreme flood that is physically possible, which result from severe most combinations of meteorological, like the probable maximum precipitation (PMP) and hydrological factors. However, in agricultural areas where failure would result only in flooding of crops, a design for a much smaller degree of protection could be reasonable. When conditions lie between these two extremes, varying design flood of certain probability (return period) (e.g. Q_{150}, Q_{1000}, $Q_{10,000}$) will apply for the relevant level of dam safety.

In practice, the selection of the presumably 'safe' design flood is done following recommendations of exceedence probability for design floods given in standards. Most of the standards base for their recommendations on subjective judgment of potential effects of dam breach (no ridged monetary or quantified otherwise limit is given). ICOLD (1992) gives summary on this standards and approaches. ICOLD (1992), whilst not providing fixed rules for selection of 'safe' design floods, it reports criteria, which can be used to guide selection of suitable design floods, that are developed for selecting a spillway design flood by different bodies, amongst other government agencies, by Britain Institute of Civil Engineers (ICE) and United Sates Army of Corps of Engineers (USACE), Australian Committee on Large Dams (ANCOLD). These three guidelines are reproduced in Table 5.5 to Table 5.7 below. For their recommendation these standards generally capitalize on dam size, risk involved to life and property in case of failure and/or type of dam. Based on subjective consideration of these factors, a 'safe' design flood is recommended, which is given in terms of return period or it could be the PMF.

For designing dams to resist overtopping to a level recommended by either of the above mentioned standards a certain 'safe' design flood has to be selected. In classical design practices this design flood is taken to be a deterministic flood magnitude that is assumed to have a certain return period (i.e. a flood expected to come on average once in certain years) or it is taken to be the extreme physically possible flood (i.e. the PMF). However, flood is exceedingly complex natural hydrologic process which is a results of interaction between a number of component inherently random parameters; such as catchment characteristics, geomorphologic, geological, rainfall and other meteorological variables, antecedent moisture content, upstream development influencing water and land use/cover etc. Each of these component parameters further depend on multitude of constituent uncertain parameters. Therefore, analytically modeling of floods is difficult. Moreover, frequency analysis using measured flood data is also not capable of giving an

accurate deterministic peak flood magnitude because it involves larger extrapolation of return period well beyond available measured flood records. This extrapolation is inevitable in dam design because mostly flood with return periods of 1000 year or more are applicable. No river on earth has that long record which can enable flood frequency analysis to render deterministic value. This makes problem of estimating flood peaks a very complex and uncertain task. For example, Figure 5-15 and Figure 5-16 presents the randomness in measured flood peaks and hydrographs at the case study dam site. Additional discussion on this issue is given in section 7.3.

Table 5.5: ICE recommended spillway design flood (ICE 1996, from ICOLD 1992).

Potential effect of a dam breach	Initial reservoir condition	General Reservoir design flood (values in bracket refer to design flood if overtopping is tolerable)	Concurrent wind speed and minimum wave surcharge allowance
Category A Where a breach could endanger lives in a community[#]	Spilling long term average flow	PMF (10,000 years flood)	Mean annual maximum hourly wind speed
Category B Where a breach: i. could endanger lives not in a community[##] ii. could result in extensive damage	Just full (i.e. no spill)	10,000 year return period flood or 0.5 PMF, which ever is greater. (1000 year flood)	Mean annual maximum hourly wind speed. Wave surcharge allowance not less than 0.6.
Category C Where a breach would pose negligible risk to life and cause limited damage	Just full (i.e. no spill)	1,000 year return period flood (150 year flood)	Mean annual maximum hourly wind speed. Wave surcharge allowance not less than 0.4.
Category D Special cases where no loss of life can be foreseen as a result of a breach and very limited additional flood damage would be caused	Spilling long term average flow	150 year return period flood (Not applicable)	Mean annual maximum hourly wind speed. Wave surcharge allowance not less than 0.3

[#] refers to areas where people live and congregate, especially homes.
[##] refers to people within the flooded area who may be at risk whilst passing through the area, washing etc.

Table 5.6: ANCOLD recommended spillway design flood (reproduced from ICOLD 1992).

Potential effect of a dam breach	Reservoir Design Flood
High Incremental Flood Hazard Where the loss of human life or extreme economic damage would be expected as the result of a dam break.[#]	Probable Maximum Flood
Significant Incremental Flood Hazard Where significant economic damage would be expected as the result of a dam break or where there is potential for (but unlikely) loss of life.[##]	Design flood should be in the range of the 1:1,000-1:10,000 return period flood, but not less than half the estimated PMF
Low Incremental Flood Hazard Where only minor damage/economic loss and no loss of life can be expected as the result of a dam break.	<1:1,000

[#] Permanent population towns/dwellings exposed to hazardous flow conditions, where evacuation may not be possible due to factors such as limited warning times.
[##] Loss of life not expected but is possible. No significant urban development. Only a small number of habitable structures or people (travelers, farmers etc.) exposed to hazardous flow conditions.

Table 5.7: US Army Corps of Engineers recommended spillway design flood (ICOLD 1992).

Hazard Category	Hazard Definition	Reservoir design flood
Low	No loss life Minimal economic loss	Small Dam[1]: 50-100 year flood Medium Dam[2]: 100 year flood - 0.5 PMF Large Dam[3]: 0.5-1.0 PMF
Significant	Few lives lost-small number of habitable structures Appreciable economic loss	Small Dam: 100 year flood - 0.5 PMF Medium Dam: 0.5-1.0 PMF Large Dam: PMF
High	More than a few lives lost Extensive economic loss (community, industry, agriculture)	Small Dam: 0.5-1.0 PMF Medium Dam: PMF Large Dam: PMF

[1]Storage capacity 0.62-1.23 Mm^3 Height 7.6- 12.2 m
[2]Storage capacity 1.23-61.5 Mm^3 Height 12.2 - 30.5 m
[3]Storage capacity > 61.5 Mm^3 Height > 30.5 m

Besides, for the selected peak flood of certain return period or PMF magnitude an inflow hydrograph has to be constructed. This shall be done after consideration of all pertinent meteorological, hydrological and catchments data, including the extent and reliability of rainfall and stream flow records, which all are highly uncertain parameters.

In addition, to determine the spillway design discharge and estimate the required margin of freeboard, the design flood (in flow) hydrograph has to be converted into outflow hydrograph by using one of the established flood routing methods which, in turn, is a function of spillway type, hydraulics, size, and its operation, sedimentation and of the reservoir area. This whole process is a complex task involving a host of additional uncertain hydraulic, operational, hydrologic, metrological, and geophysical parameters.

However, from the discussion in the preceding paragraphs it is apparent that the whole notion of selecting a deterministic 'safe' flood magnitude, the use of deterministic hydrograph and routing procedure is wrong. The ultimate false sense of absolute safety deduced from such analysis is also unrealistic. There is no assurance that the magnitude of the selected design flood of certain return period or PMF is a deterministic value. The process of assigning return period to flood magnitude can have induced errors due to the use of limited sample data for extrapolating to extreme floods. The passage of this flood with time through the dam (the hydrograph shape) is also dependent on host of uncertain catchment and rainfall parameters, which can hardly be represented with a static or deterministic shape. The hydraulics of spillways, waves and reservoir is also far more complex and involves lots of uncertainty. Therefore, conventional design procedures are disguising the natural uncertainty in parameters and estimated design floods.

Increasing attention is being given to using probabilistic design, risk and safety analysis as an alternate useful guide in the selection of design flood and flood design of dams. This approach not only deals with uncertainties in a natural and transparent way but it also allows for economic optimization of designs and it gives clear criterion for selecting safe design flood. The safe design flood is chosen in such a way that it minimizes (on probability basis) the sum of the spillway plus dam and damage costs.

As the size of selected design flood increases, the capital cost of the structure will also increase. At the same time, the probability of experiencing damage as result of dam failure will decrease. By taking both of these elements and adding the two, a total cost curve will result. The most economical spillway-freeboard design can be found from the minimum of this curve. Such an approach has been previously limited to dams with low-significant risk (like category b, c, and d dams in ICE standards) and only when human lives are not treated (Novak, 2003), (ICOLD, 1992). The limitation in application of this method is put mainly because (1) the probabilistic methods were not well developed and theoretically justified this raised doubts in results from such methods. Mostly the MCSM is used with only one random design parameter, i.e. mostly only the design flood peak (Q_P), taken as random variable. All the other uncertainties in hydrographs, routing, spillway hydraulics, operation etc are ignored. (2) use of alternative, theoretically founded probabilistic and analytical methods was not possible either because the methods were non-existent in ready to use form or the computational difficulty they pose does not make them attractive. (3) the question if the value of human lives lost should be included in the calculation or not, if so what value should be assigned, is debatable. (4) techniques, tools and standards that support the implementation of the approach are undeveloped. Therefore, the tendency of design practices was to go for deterministic approaches which are wrongly supposed to give "absolute safety". Such a method usually compound a series of safety factors to over design the system but with still uncertain and unquantified reliability. Nevertheless, availability of more and more refined probabilistic methods, computational software and developed change of society concept to risk necessitated the change of this practice.

This research presents more comprehensive probabilistic methods of evaluating dam overtopping probability (see application details in chapter 7). The proposed approaches take in to account most of the significant uncertainties discussed above. Moreover, in chapter 7 results found from comprehensive probabilistic analysis are compared with results from classical methods and standards. In addition, application of available computational software for the purpose is demonstrated.

It is found convenient to discuss details of the analytical procedures involved in selecting a 'safe' design flood, construction inflow flood hydrograph, doing stochastic reservoir routing, wave heights computation, selection of safe freeboard limit, and in general, the procedures for a design of dams against overtopping in chapter 7, together with the application of proposed probabilistic methods. Therefore, its repetition here in this section is found unnecessary.

5.3 The case study dam (Tendaho Dam, Ethiopia) and Awash River basin

5.3.1 General

The probabilistic techniques discussed in earlier chapters are illustrated using the cases of seepage flux randomness, sliding failure probability, and overtopping failure probability computations using data and facts from Tendaho Dam, North Eastern Ethiopia (see Figure 5-5). The Tendaho Dam is a dam constructed for harnessing flows of Awash River at Tendaho, Ethiopia for the purpose of irrigating nearly 60,000 ha sugar cane plantation. The project is aimed to set up a sugar factories having target production of 500,000 tones of sugar per annum. Tendaho Dam is a zoned earth dam (see Figure 5-6).

5.3.2 The Awash River

The Awash drains the northerly part of the Rift Valley in Ethiopia from approximately 8.5 ºN to 12 ºN with total drainage area of 112 211 km^2 (Halcrow, 1989). the source of the Awash River is near the southern edge of the Ethiopian Highlands, some 150km west of the capital Addis Ababa, at an altitude of about 3,000m above sea level (see Figure 5-6). Available mapping and the physical geography and geomorphology of the Awash River basin are described in detail in Halcrow (2006). The outline characteristics include:

- The river flows south east for about 250km before entering the Rift valley, which it follows for the rest of its course. The river is endorreic (no mouth to the sea), and ends at Lake Abe, on the border with the Republic of Djibouti, at an altitude of some 250m; and
- The total length of the river is some 1,200 km, and it drains a catchment area of 112,211 km².

The basin may be divided into seven sub-basins, as follows (see Figure 5-7):

i. *Upland basin or Becho plain*: includes the headwaters and extends down to Koka Dam (9.6% of catchment area);
ii. *Western Highlands*: left bank tributaries from Kesem River to Logiya River (22.2% of catchment area);
iii. *Upper valley*: broadly from Koka dam to Awash Arba (11.2% of catchment area);
iv. *Middle valley*: between Awash Arba to Adaitu (14.0% of catchment area);
v. *Lower valley* between Adaitu and Logiya (3.5% of catchment area);
vi. *Lower plains*: downstream of Logiya, where the Awash River meanders and terminates at Lake Abe (5.6% of catchment area); and
vii. *Eastern catchment*: area on the right bank downstream of Koka Dam (33.9% of catchment area).

Halcrow (1989) estimate of the mean annual volume of runoff measured at gauges on the rivers entering Koka reservoir is 1520 Mm^3. At Awash station, estimated annual flow is 1546 Mm^3, only a small increase because the river flow is being depleted by losses from Koka reservoir and by diversions for irrigation. The mean annual runoff at Hertale, before Awash enters into the swamp, reaches to 2692 Mm^3. But after passing the swamp, the flow reduced to 2424 Mm^3 at Tendaho although tributaries originated from Wollo joins the Awash River.

Annual maximum and mean flows of Awash river at Tendaho dam site (drainage area 62,088 km^2) for thirty-five years (1965-2002 except 1994 and 1995), has been extracted from a daily flow data series obtained from Ethiopian Ministry of Water Resources (MoWR), (Halcrow 2006, 1989), (WWDSE, 2005a), and (Gibb, 1975). The annual maximum flow shows tremendous randomness (see Figure 5-8).

To show the temporal and spatial variability of precipitation on the basin mean monthly rainfall distribution at key sample stations in the basin is given in Figure 5-9. The uplands in the basin receive an annual rainfall of over 800 mm where as the lower plains receive an annual rainfall of less than 200 mm. From Figure 5-9 it can be seen that in both upland and lower land of Awash basin March, April, July and August receives more rainfall as compared to the other months. Accordingly, in the lower Awash at Tendaho, two flood seasons are experienced. The first is spring (Mid-Feb-April) and the second is during summer (July- September). In both

seasons equally high peak flows were observed, but the flood peak sustain longer during main rainy season than the spring flood. The hydrology regime of Awash is complicated by major losses in Gedessa swamp complex, abstractions by a number of large scale irrigation schemes along its reach, losses and operation at Koka reservoir, hydro powers at Koka, at Awash II and Awash III, huge elevation variations extending from around 3000 to less than 300 m a.m.s.l. Additional information on different data for Awash River basin used in this study is given in section 5.4.

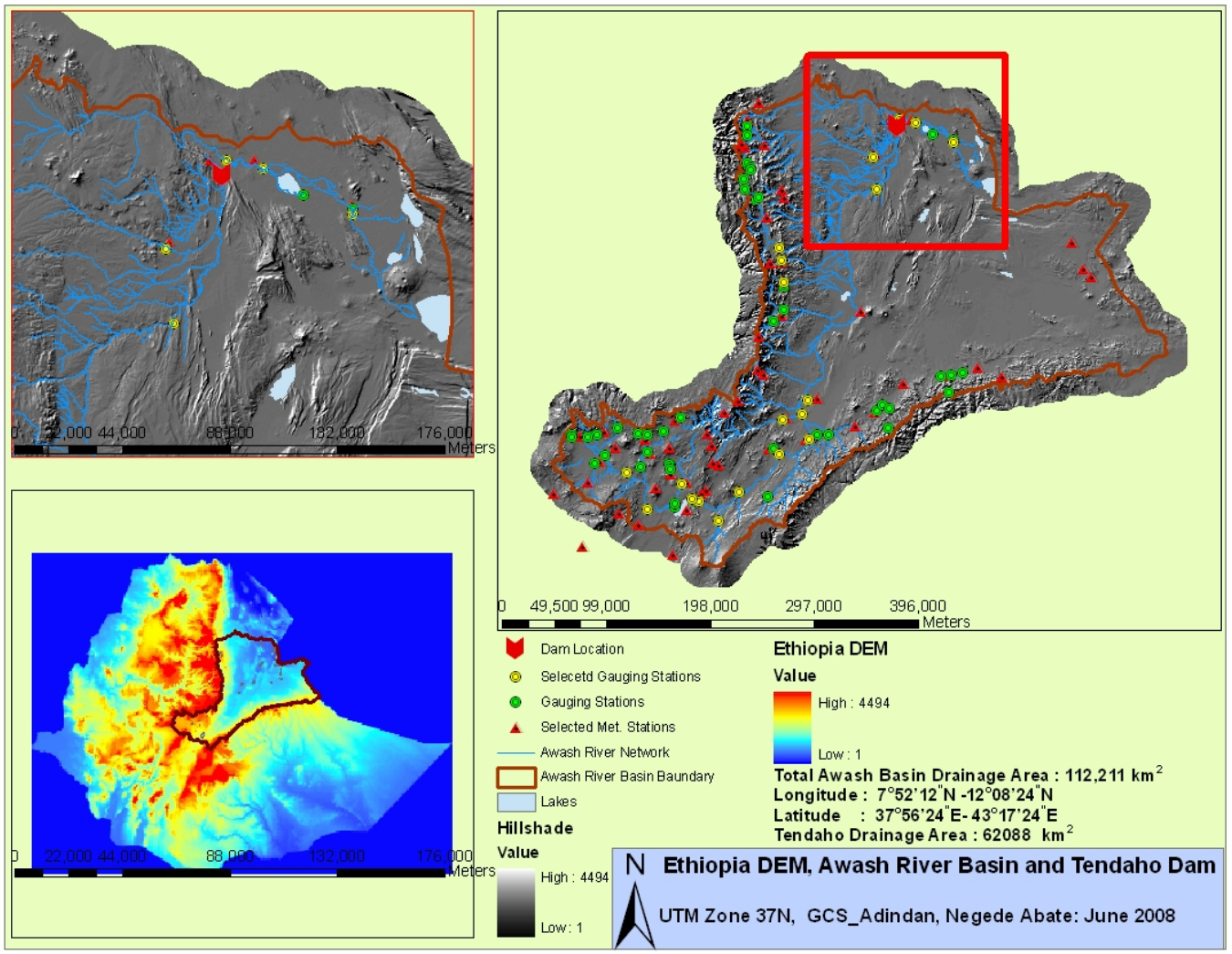
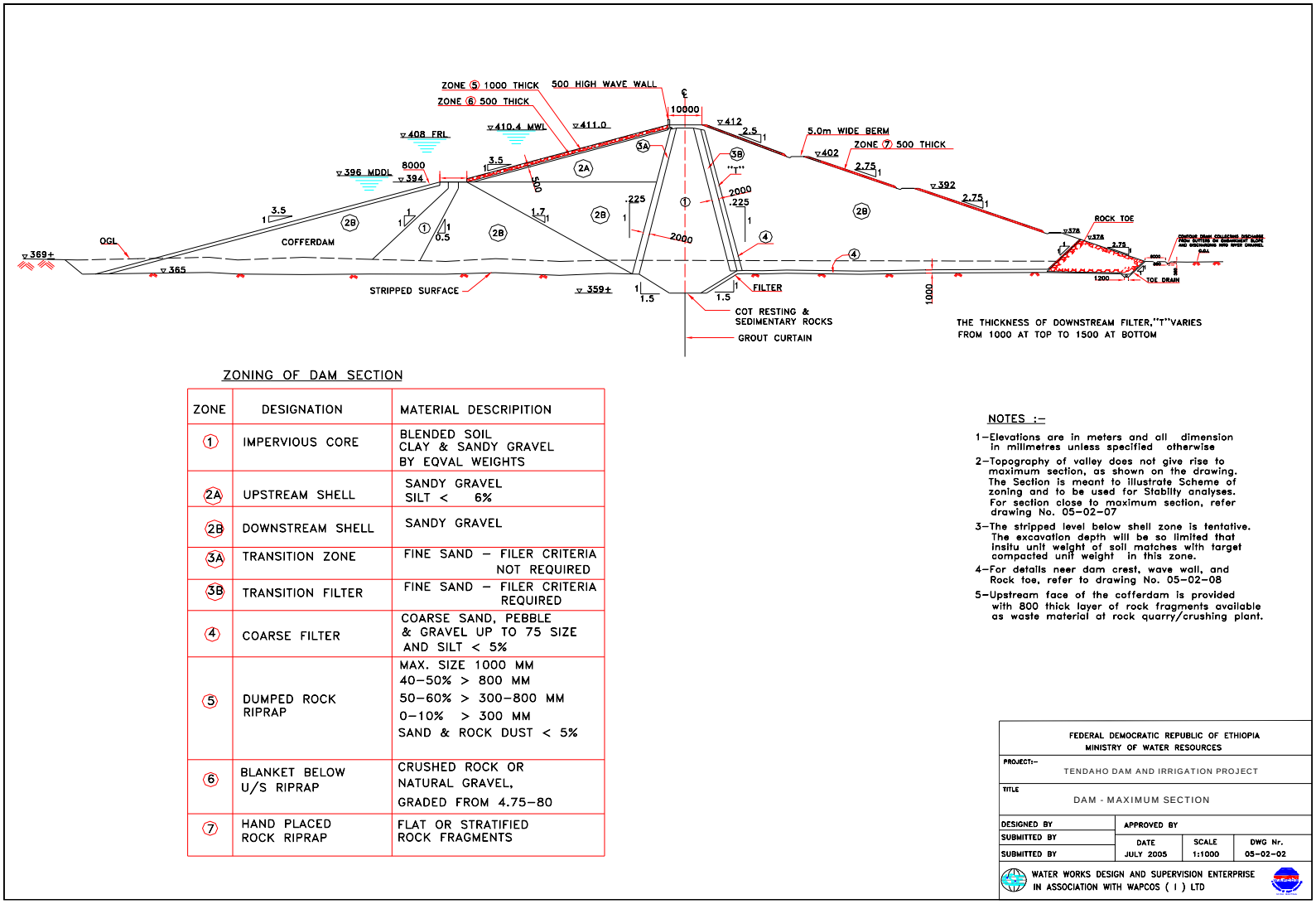


Figure 5-5. Tendaho Dam site and River Awash basin.



ZONING OF DAM SECTION

ZONE	DESIGNATION	MATERIAL DESCRIPITION
①	IMPERVIOUS CORE	BLENDED SOIL CLAY & SANDY GRAVEL BY EQVAL WEIGHTS
2A	UPSTREAM SHELL	SANDY GRAVEL SILT < 6%
2B	DOWNSTREAM SHELL	SANDY GRAVEL
3A	TRANSITION ZONE	FINE SAND – FILER CRITERIA NOT REQUIRED
3B	TRANSITION FILTER	FINE SAND – FILER CRITERIA REQUIRED
④	COARSE FILTER	COARSE SAND, PEBBLE & GRAVEL UP TO 75 SIZE AND SILT < 5%
⑤	DUMPED ROCK RIPRAP	MAX. SIZE 1000 MM 40–50% > 800 MM 50–60% > 300–800 MM 0–10% > 300 MM SAND & ROCK DUST < 5%
⑥	BLANKET BELOW U/S RIPRAP	CRUSHED ROCK OR NATURAL GRAVEL, GRADED FROM 4.75–80
⑦	HAND PLACED ROCK RIPRAP	FLAT OR STRATIFIED ROCK FRAGMENTS

NOTES :–

1–Elevations are in meters and all dimension in millimetres unless specified otherwise

2–Topography of valley does not give rise to maximum section, as shown on the drawing. The Section is meant to illustrate Scheme of zoning and to be used for Stability analyses. For section close to maximum section, refer drawing No. 05–02–07

3–The stripped level below shell zone is tentative. The excavation depth will be so limited that Insitu unit weight of soil matches with target compacted unit weight in this zone.

4–For details neer dam crest, wave wall, and Rock toe, refer to drawing No. 05–02–08

5–Upstream face of the cofferdam is provided with 800 thick layer of rock fragments available as waste material at rock quarry/crushing plant.

FEDERAL DEMOCRATIC REPUBLIC OF ETHIOPIA
MINISTRY OF WATER RESOURCES

PROJECT:– TENDAHO DAM AND IRRIGATION PROJECT

TITLE DAM - MAXIMUM SECTION

DESIGNED BY	APPROVED BY		
SUBMITTED BY			
SUBMITTED BY	DATE JULY 2005	SCALE 1:1000	DWG Nr. 05–02–02

WATER WORKS DESIGN AND SUPERVISION ENTERPRISE
IN ASSOCIATION WITH WAPCOS (I) LTD

Figure 5-6. Geometry and scheme of zoning of Tendaho Dam (from WWDSE, 2005b).

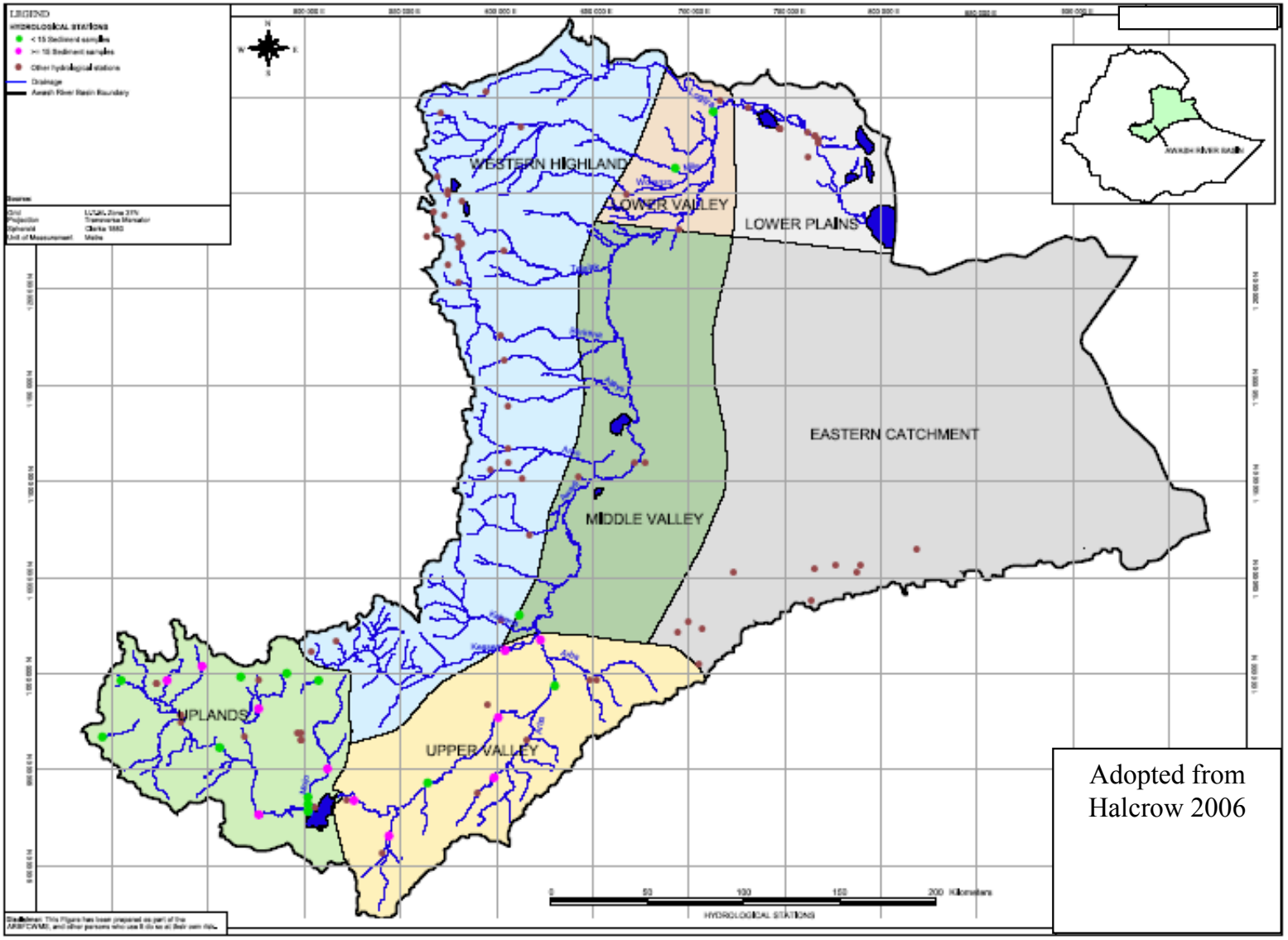


Figure 5-7. Awash River sub-basins and hydrological stations (from Halcrow, 2006).

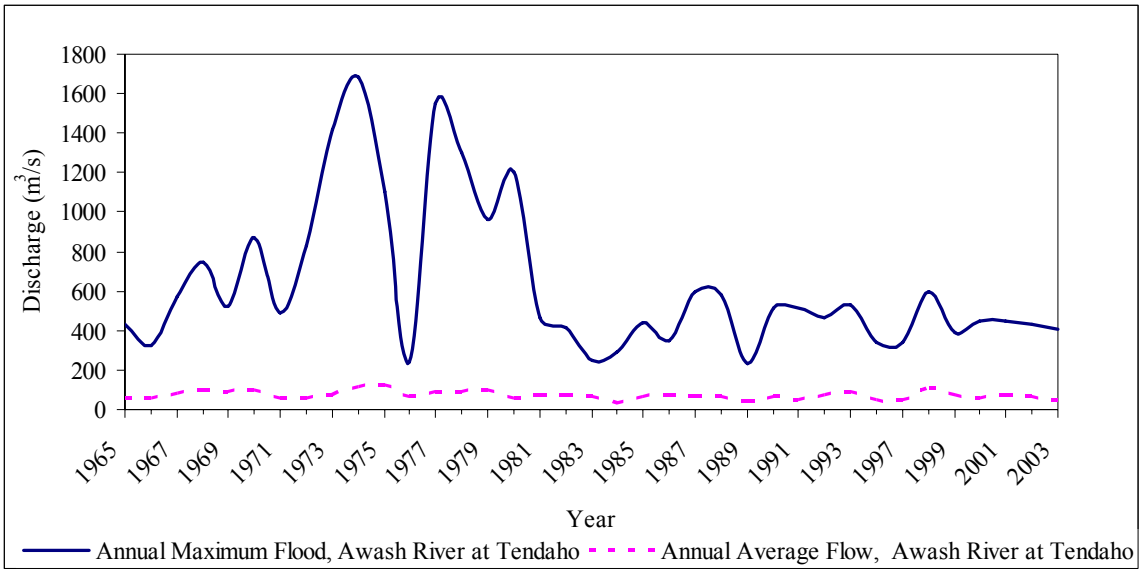


Figure 5-8. Annual maximum and mean flows of Awash River at Tendaho station.

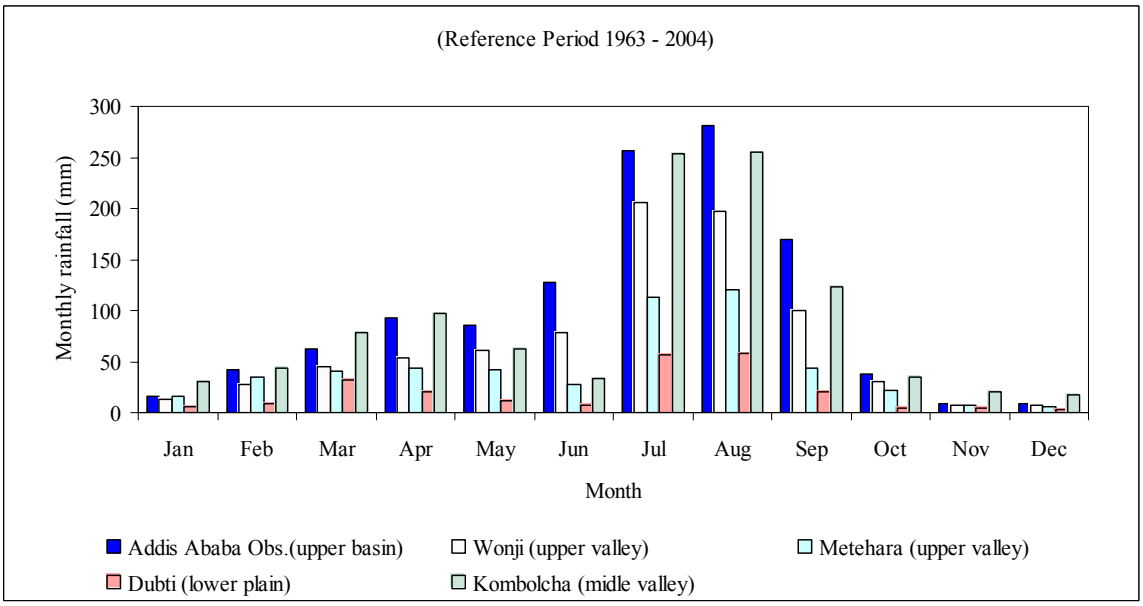


Figure 5-9. Mean monthly rainfall at key stations in Awash River basin.

5.3.3 Salient features of Tendaho Dam project

Appendix 9.3 gives the salient features of Tendaho Dam project in a summarized form.

5.3.4 Tendaho Dam basic design considerations and deterministic compliance of design requirements

In the original design of Tendaho Dam the basic design consideration has been to achieve "safety" consistent with economy. The original design document of the dam states that it is

essential for the dam to satisfy the following basic requirements so that it will be "safe" and stable under various conditions of operation of the reservoir. The basic requirements listed are:

1. The slopes of the embankment must be stable during construction and under all conditions of reservoir operation;
2. Seepage flow through the embankment, foundation, and abutments must be controlled so that no internal erosion takes place. The amount of water lost through seepage must be controlled and exit gradient at downstream toe of the dam should not be high enough to cause sloughing or piping action;
3. The embankment must be safe against overtopping during occurrence of the inflow design flood;
4. Adequate free board must be provided above FRL and MWL to guard against overtopping by wave action. Free board must also take into account the settlement of embankment under the effect of dynamic loading;
5. The upstream slope must be protected against erosion by wave action, and the crest and downstream slope must be protected against erosion due to wind and rain.

The original design of Tendaho Dam, like any classical dam design endeavors, assumed that standard deterministic design procedure will ensure the compliance of the above design requirements. It is assumed that:

1. The deterministic stability analysis of the designed section of dam ensures the compliance of the design requirement at serial 1 above;
2. The provision of chimney drain, horizontal filter and toe drain in the designed section of Tendaho Dam takes care of the requirement in respect of seepage control within embankment while the scheme of consolidation and curtain grouting below dam foundation ensures the compliance of basic design requirement at serial 2 above in regard to seepage control below the dam foundation;
3. Provision of chute spillway designed for probable maximum flood inflows using classical deterministic methods ensures the safety against overtopping during occurrence of the inflow design flood;
4. The provision of freeboard, ensure the compliance of design requirement at serial 4 above;
5. The provision of dumped rock riprap on upstream slope and hand placed rock riprap on downstream slope ensures the compliance of design requirement at serial 5 above.

However, it has been discussed time and again in preceding sections of this dissertation that an attempt to solve most of dam design problems using classical deterministic approaches is simply unrealistic. It obscures lots of uncertainty involved in design parameters and models. It simply requires to accept relatively high values for factor of safeties and to blindly assume 'absolute safety', whereas the level of safety provided still remains uninnumerable. Nowadays, society and designers raised doubts on the appropriateness of this approach. How can one be sure about the compliance of design requirements, for example, referring to seepage, stability and overtopping as stated in 1, 2 and 3 above? The parameters influencing stability, seepage and overtopping are random. Consequently, the compliance to design requirements thereof is random. One can not precisely tell about the compliance. At most it can be expressed in probabilistic terms and that requires probabilistic design approach.

5.4 Data used

The basis for quality dam risk and safety analysis is the availability of accurate and adequate data on hydrological, meteorological, geophysical, geotechnical, material characteristic, construction practice, geometry, economy, damage prevalence etc. This is because the quality of the various probabilistic analyses depends on the quality, quantity and resolution of the available data on parameters that directly or indirectly influence load and strength parameters.

For the case study dam in this research plenty of secondary and primary data has been collected from various sources: authorities in Ethiopia, previous studies, design documents, laboratory tests and site visits. To collect data from these different sources a field work has been conducted in the period between 6 July 2006 and 25 September 2006. The research has benefited from the fieldwork timing because it was scheduled while the dam was under construction. This arrangement gave the researcher easy access to material samples, study documents, field engineers' consultations and the chance to observe the construction practice en situe.

Moreover, the long history of Tendaho Dam, which spanned for almost half a century before it got implemented, has permitted this research to get plenty of secondary data. Six important studies have been made for assessing the water resources of Awash River basin and for the study of Tendaho Dam and irrigation project. The first one was by the Sogreah-FAO (1965) fundamental survey of Awash basin, in which key hydrometeorological stations were installed. Construction of a storage dam at Tendaho site was first conceived under this pioneer study made by the French consultants Sogreah in year 1965, although the project was not implemented until forty years after in 2005. The report of this study was edited and brought out by Food and Agriculture Organization of United Nations (FAO) under United Nations special fund in year 1965 by the title "Report on Survey of the Awash River Basin" (Sogreah-FAO, 1965). The second study was done by Sir Alexander Gibb and Partners, London in association with Hunting Technical Ltd. in year 1975 (Gibb, 1975) for the feasibility study of the Lower Awash Valley which basically focused at the design of Tendaho Dam and irrigation schemes. The third study was done by Uzgiprovodkhoz of Uzbek State in 1985 (Uzbek, 1985). This study titled "Feasibility Study, Proposal and Estimate of Cotton Development on the area of 60 thousands hectare in the Lower Awash Valley, Ethiopia" relates specifically to the Lower Awash Plain. The report on the studies made by Uzgiprovodkhoz deals in details on subjects like crop water requirement, water balance studies and irrigation system but is found to be very sketchy in respect of engineering structures of dam, spillway and bottom outlet required for harnessing the water of river Awash for irrigation use. The forth study in the series is the "Master Plan for the Development of Surface Water Resources in the Awash Basin" prepared during 1989 by Halcrow for Ethiopian Valley Development Studies Authority (Halcrow, 1989). The report includes a brief discussion on recommendations made under feasibility reports by Alexander Gibb and Uzbek consultants in respect of dam, spillway and irrigation bottom outlet. It is found that the report, while highlighting the contradictions in the two feasibility reports, did not make definite recommendations of its own but left all major issues to be taken care of at final design stage. The fifth study is the one conducted by Ethiopian Water Works Design and Supervision Enterprise (WWDSE) in association with Water and Power Consultancy Service India Ltd in 2005 (WWDSE, 2005a, b). This is the study which actually gave the final design and that got implemented. In addition to the above five studies, Halcrow has done a study in year 2006 for

Awash River Basin Flood Control and Watershed Management Study (Halcrow, 2006). This study focuses on providing information for a better understanding of the development challenges in the Awash River Basin area and to formulate effective and sustainable flood control, watershed management programs and projects.

Whilst the above series of studies produced or collected adequate hydrometeorological, sedementlogical and other basic data on the dam, the available data on characteristics of materials used for constructing the dam were not adequate to permit an acceptable probabilistic analysis. Therefore, additional laboratory tests were conducted on disturbed and undisturbed samples taken from stockpiles and borrow areas used for producing materials for the construction of the shell and core zones of Tendaho Dam. The laboratory tests were conducted in the Ethiopian Construction Design Share Company (CDSco) material test laboratory. The tests were conducted with full involvement of the researcher. Some replications were also given to Arba Minch University (AMU) soil mechanics laboratory. The tests conducted on the shell and core construction materials includes triaxial (Consolidated Undrained (CU)) test for determination of c' and ϕ', one dimensional consolidation/odometer test for determination of, C_c, m_V and C_α, permeability test for determination of K, Atterburg limits test for determination of *PL*, *LL*, *PI*, grain size analysis, compaction for determination of MDD and optimal moisture content.

The data collected from the field work has been processed and stored in easily accessible and utilizable form. Appendix 9.3.2 gives the complete list of data collected during the field work with its form, amount and source. Table 5.8 gives the most important resistance and load related data collected either from historic records, secondary sources, simulation or laboratory tests. Because of space limitation it is not possible to include all the raw data collected in this dissertation. However, summary of those data which are deemed to be important for the case study are included in appendix 9.3.3 to 9.3.11 and brief discussion on their method of acquisition, source, utilized standards and record length is given below. The rest is provided on CD-ROM.

5.4.1 Resistance related data

5.4.1.1 Data on shell and core material characteristics

As mentioned above, considerable amount of secondary data on core and shell material characteristics of Tendaho Dam has been collected however it was short of providing sufficient data to permit an acceptable probabilistic analysis. Therefore, it was necessary to conduct additional laboratory tests. These tests were conducted mainly in Ethiopian Construction Design Share Company (CDSco) material test laboratory with some replications made in Minch University (AMU) soil mechanics laboratory. The laboratory tests conducted includes shear strength, permeability, consolidation, grain size analysis, Atterburg limits and compaction tests. The standard procedures followed to carry out the tests are provided in Table 5.9.

Table 5.8: Selected Tendaho dam load and resistance data sources and methods acquiring.

Data Type	Method of data acquisition
resistance related parameters	
effective shear strength parameter, c' and ϕ', for core and shell materials	secondary data and data from laboratory experiment
permeability, *K*, for core and shell materials	"
specific gravity for core and shell materials	"
consolidation parameters, *Mv*, and *Cc* for core materials	"
compaction parameters *MDD* and *OMC* for core and shell materials	"
grain size (gradation) and Atterburg limits for core and shell materials	"
elevation storage-area relationship	secondary data from earlier surveys
topography, geometry and dam dimensions	"
load related parameters	
flood records	historic records
wind velocity	historic records
rainfall and other climatic variables	historic records
pore water pressure	simulation
reservoir water level fluctuation and hydrostatic trust due to reservoir level	simulation
synthetic flow records and reservoir simulation results	simulation
sediment loading	historic record and regression

a) Compaction tests

The compaction tests were done using the standard proctor compaction procedure (BS 1377) for the clay samples and relative density for the granular samples.

b) Permeability tests

The permeability tests were done on compacted samples. The samples in the case of the clay soils were compacted to a maximum dry density (MDD) and Optimum Moisture Content (OMC) based on standard proctor compaction procedure mentioned above. In case of the granular samples, the permeability test was conducted on sample size less than 4.75 mm and a relative density of a maximum of 70%. The limitation of the size is made due to absence of apparatus to accommodate all the grain size for testing.

c) Shear strength tests

To represent different drainage conditions, the clay samples were tested in the Unconsolidated undrained (UU) and consolidated undrained (CU) conditions on remolded samples. In the UU test saturated and partially saturated samples were tested. Similar to the permeability tests the shear tests on granular samples were conducted on grain size less than 4.75 mm due to absence of apparatus to accommodate all sizes.

5.4.1.2 Dam and appurtenant structures dimensions, topography and elevation storage-area relationship

Data on dam and appurtenant structures dimensions, topography and elevation storage-area relationship is taken from Tendaho Dam final design documents (WWDSE, 2005b)-"Dam and appurtenant works" and (WWDSE, 2005a) Hydrology of Tendaho Dam.

Table 5.9: Standard procedures utilized for material testing.

Type of test	standard
grain size analysis - hydrometer	BS Test 7(D)
grain size analysis – dry sieve	BS
compaction- standard proctor compaction	BS 1377:1975
permeability	ASTM D2434
triaxial tests CU/UU	BS 1377:1975

5.4.2 Load related data

5.4.2.1 Discharge data

For the flood frequency analysis in this research annual maximum flood data of the Awash River at the Tendaho dam station over the period 1965-2002 (except 1994 and 1995) are extracted from the daily flow data obtained from the Hydrology Department of the Ministry of the Water Resources (MoWR), Ethiopia and Halcrow 1989. The daily data from MoWR, including the stage–discharge-rating curves have been examined. The hydrology Department utilized three differ rating curves due to the unstable nature of the river banks and aggrading/degrading river bed. The first rating curve is used over the period 1965-1972, the second one is over the period from 1973-1982 and the last one is from 1983 on ward. Halcrow (1989) has conducted an extensive study on the surface water resources of the Awash River basin and published annual maximum flow data. Reference period for Halcrow Analysis was 1963-1986. At some years the annual maximum flood magnitude published by MoWR and (Halcrow 1989) show slight differences. This difference came from adjustments made by (Halcrow, 1989) on the MoWR raw data. Halcrow (1989) argues that at most hydrological gauging stations in Ethiopia, including Tendaho station, only two water level observations are taken daily at fixed observer reading times of (08h00 and 18h00). For some other gauging stations continuous records are (or have been in the past) kept through the use of chart (and now loggers in very limited instances) recorders. However, even where continuous records have been kept using charts, the recorded hydrographs have not normally been digitized or even fully analyzed manually. Instead the two water levels as noted by the on-site observer have been used and the charts archived. The result of this practice is that "MoWR published" peak flood figures usually correspond to the annual maximum daily average value. The real peak flood value is therefore inevitably underestimated. The margin of error will potentially be highest in the tributaries where the durations of peaks are limited. In this study when the MoWR hydrology department data are different from the Halcrow (1989) maximum daily flow data, the Halcrow maximum annual flow data are used for the period 1965-1986. The historic mean monthly, mean annual and annual maximum flow data for Awash River at Tendaho station is given in Appendix 9.3.4. Figure 5-8 shows the plot of mean and annual maximum flows of Awash River at Tendaho station.

The inflow into the Tendaho reservoir is mainly determined by three factors: (1) the runoff characteristics of the Awash basin upstream of the Tendaho dam, (2) the irrigation and water supply abstractions from the Awash river, (3) the magnitude of evaporation and seepage losses in the Gedebessa swamp complex. The historical flow data measured at the Tendaho dam site is thus the integrated effect of these abstractions and additions of runoffs in the basin. Figure 5-8 shows the time series of the annual flow of the Awash River at the Tendaho Dam site. It is seen that the lowest mean annual flow (30 m^3/s) was observed over the period 1982-1988 corresponding to high irrigation development upstream of the Tendaho Dam (Appendix 9.3.4). In the reservoir simulation, the effects of the current and the future irrigation-water supply abstractions and additions should be accounted. In the reservoir simulation of Tendaho Dam hydrologic study it is assumed that the 1982-2003 mean flow condition will repeat in the future and used as a basis of reservoir simulation. Therefore, the monthly flow data at the Tendaho dam site over the period 1982-2003 are used as base data, as these are corresponding to relatively high irrigation development in the basin, for generating long-term series based on the Thomas-Fiering synthetic monthly flow generation model.

5.4.2.2 Sediment data

Sediment transport rate data is useful in determining the useful life of dam, i.e. fixing the dead storage level at the end of the design period for deciding the location of sluices, and it is also required for fixing reservoir operation rules. The sedimentation process determines the available live storage at the end of any year. In overtopping probability analysis this influences the volume available for accommodating in coming flood. Therefore, it has to be accounted in doing reservoir simulations while running water balance models. The knowledge on the available storage volume decides on the spillage in a given time period. Sedimentation has a role in determining the likely water level in a reservoir because operators tend to keep water levels high to compensate for a storage lost due to sedimentation.

The sediment transport data available at Tendaho are data measured during the Sogreah (1965) study (measurements from mid 1961/62 to mid 1963/64), and eight additional measurements made by MoWR hydrology department in three years 1985, 1986, and 1987. The Sogreah (1965) report gives monthly flow and corresponding estimated suspended sediment load (data given in Appendix 9.3.5). The monthly data are derived from the daily flow-sediment transport rating equation developed for the Awash at Dubti station based on 74, 96 and 11 sediment samples measured in 1962, 1963 and 1964 respectively. The total number of the flow-sediment concentration data measured over the period 1962-1964 by Sogreah was 181. After 1964, however, only eight discharge-sediment concentration samples were taken over the period 1985-1987 (data given in Appendix 9.3.6).

The monthly discharge-sediment load relationship established based on the Sogreah data is given in Eq. (5-8) and Figure 5-10 shows the relationship graphically in log-log scale. This relationship is also mentioned in Gibb (1975) and WWDSE (2005b).

$$Q_{sediment} = 0.0101 \cdot Q^{2.1778} \qquad (5\text{-}8)$$

Where $Q_{sediment}$ is monthly sediment load in 10^3 tones and Q is monthly flow in 10^6 m^3, respectively.

In this relationship high and low flow sediment concentration variation with discharge are well represented. This relationship is used for the analyses in this study. The monthly discharge – sediment load rating equation Eq. (5-8) is applied to the monthly historical flow over the period 1962-2002 of the Awash River at Tendaho. The monthly historical flow data are given in Appendix 9.3.4. The resulting monthly sediment load is given in Appendix 9.3.7. The yearly variation of sediment load is shown in the last column of Appendix 9.3.7. It is seen that the annual sediment load varies from minimum 2 10^6 tones/year 1984 (drought year) to 90 10^6 tones 1975 (wet year). The mean value of the suspended sediment load is 26 million tones per year. The bed load is taken as 5% of the suspended load.

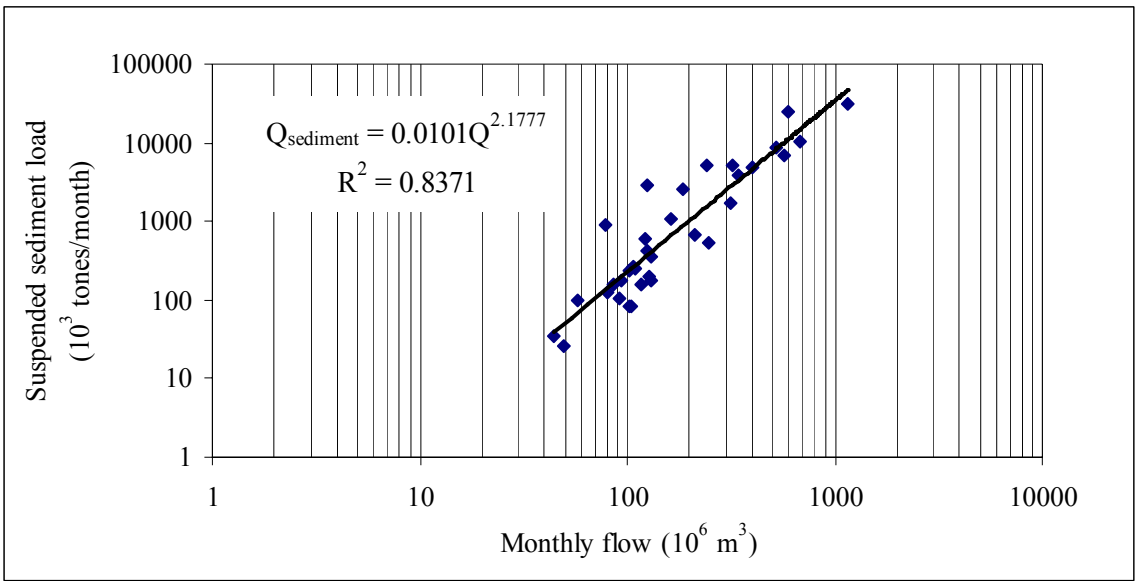


Figure 5-10. Monthly flow – sediment transport rate relationship of the Awash River based on Sogreah data.

It has also been tried to establish a sediment flow-sediment load relationship for the Awash River at Tendaho using the eight additional measurements made by MoWR hydrology department. The sediment rating equation based on this recent data is given as:

$$Q_s = 4.69 \cdot Q^{2.0361} \qquad (5\text{-}9)$$

Where Q_s is suspended sediment transport rate (tones/day) and Q is corresponding sediment and flow discharge rate (m^3/s).

The number of data, which is only eight, used to generate the later rating equation is very small for establishing reliable sediment rating curve. Therefore, the Sogreah relationship is assumed to be valid for estimating sediment transport at Tendaho and the succeeding reservoir sedimentation computations. This has also been the opinion in Tendaho Dam design document (WWDSE, 2005). For the sake of comparison the mean annual sediment load as estimated by the

two rating equations is given in Appendix 9.3.8. In addition, in Appendix 9.3.8 the estimated mean annual suspended sediment and bed load volume transported to Tendaho reservoir is provided. For estimation of reservoir sedimentation distribution at Tendaho Dam the Borland and Miller method (Area Reduction Method) is used. This has been the method employed in the original design document of the dam. This method is favored for use at Tendaho because it requires limited information, i.e. elevation-capacity-area curve and the total sediment load. The Tendaho Dam reservoir capacity elevation curve is illustrated in Figure 5-11 and the data is given in Appendix 9.3.9.

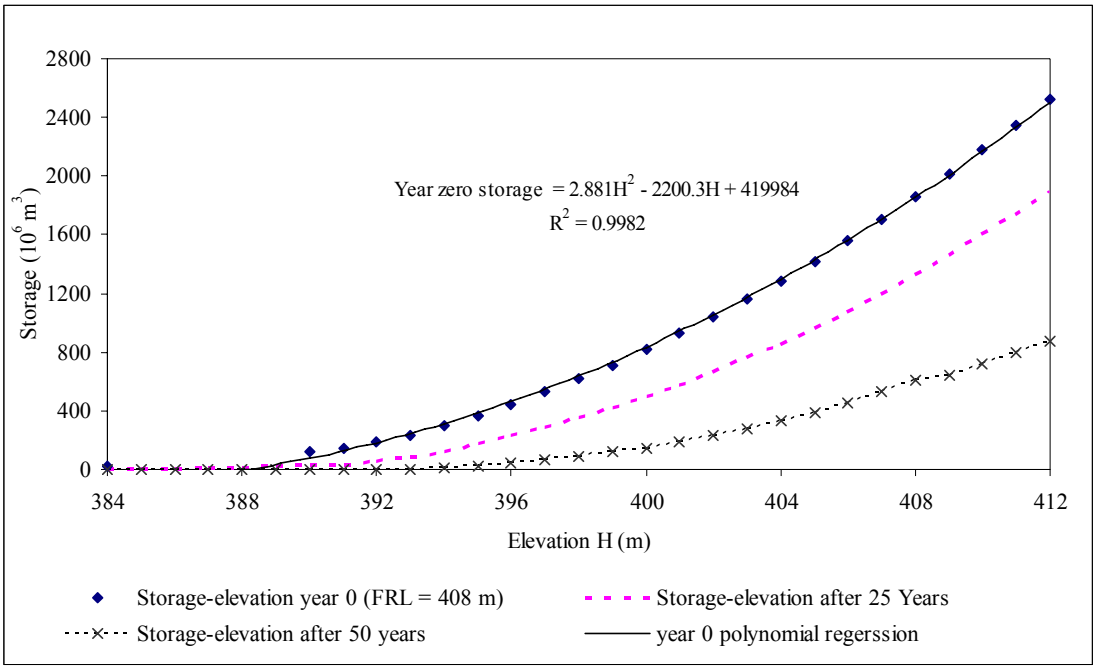


Figure 5-11. Area-capacity-elevation curve for the Tendaho reservoir at year 0, 25 and 50.

5.4.2.3 Wind speed data

Data on basic wind speed right at Tendaho dam site is not available. In this study above ground wind speed data from near by area (Melka Worer) is used (Appendix 9.3.11). This same data is used in the hydrologic design of Tendaho Dam (WWDSE, 2005a).

5.4.2.4 Simulated reservoir water levels and synthesis of hydrological sequences

For the overtopping risk analysis a water balance model has to be constructed to determine possible water levels at time of peak flows, i.e. to asses the initial water level randomness. This parameter is influenced by inflows and abstractions to and from reservoirs. The inflow into the Tendaho reservoir is mainly determined by three factors. The first is the runoff characteristics of the Awash watershed upstream of the Tendaho Dam, the second is the irrigation, and water supply abstractions from the Awash river, and the third is the magnitude of evaporation and seepage losses in the Gedebessa swamp complex. The historical flow data measured at the Tendaho Dam site is thus the integrated effect of these abstractions and additions

of runoffs in the basin. Figure 5-8 illustrates the time series of the annual flow of Awash River at Tendaho Dam site. In constructing water balance models and doing reservoir simulation the effects of the current and the future irrigation-water supply abstractions and additions should be accounted. This demands for synthetically generated inflow series, using stochastic methods, for future time all along the expected life of the dam. In the hydrologic studies of Tendaho Dam long-term series, based on the Thomas-Fiering synthetic monthly flow generation model (Thomas and Fiering, 1962), has been generated to run reservoir simulation for evaluating the chance of meeting the sugar cane plantation gross irrigation demand. WWDSE (2005a) has published the generated long term (50 years) flows for Tendaho Dam site (Appendix 9.3.10). In thus study this synthetically generated flow series is adopted.

5.4.2.5 Rainfall and other climatic data

There are more than 109 meteorological stations in and around the Awash River basin according to MoWR 2006 information. Their number varies from year to year (some are abandoned and some are added in the network). A catalog of hydrometric stations with 83 stations is given in Halcrow (1989). There area three stations in the Tendaho project area. These are Dubti, Ditbahari, and Assayita stations. Climatic elements such as precipitation, temperature, relative humidity, wind speed, and sunshine hours are primarily required in estimating potential evapotranspiration (ETo), probable maximum precipitation and reservoir evaporation in the Tendaho reservoir and irrigation command area. In this regard this study adopted analysis results with regard to irrigation requirements, evapotranspiration, and evaporation etc from the hydrologic studies of Tendaho Dam. Just for the sake of completeness the raw climatic data will be included in CD-ROM attached with this dissertation.

5.4.3 Porewater pressure data, geometry, scheme of meshing and FEM stability and seepage analysis of Tendaho Dam

In this research the deterministic analysis of stability, seepage and pore-water pressure distribution is done using the computer program GeoStudio 2004 (SLOPE/W and SEEP/W). Results from the probabilistic methods are thus compared amongst each other and with the results from the deterministic analysis. The geometry, scheme of zoning of Tendaho Dam is illustrated in Figure 5-6. The meshing pattern used for the finite element analysis of seepage and pore-pressures distribution in the dam is given in Figure 5-12.

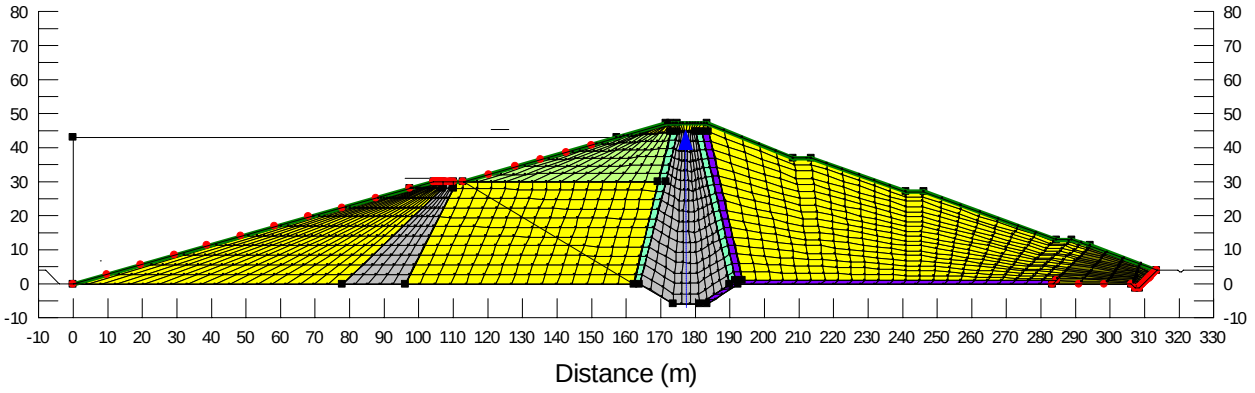


Figure 5-12. Meshing pattern for FEM seepage and pore-water pressure analysis.

As shown in Figure 5-12 for the deterministic seepage and pore-water distribution analysis a mixed-structured mesh is used. The analysis type employed is steady state. The seepage and pore-water distribution for assumed steady state conditions is determined at three critical water levels as used in the original design, i.e. for full retention level (FRL), which is the most likely case, and for the two extreme cases of maximum water level (MWL) and minimum draw down level (MDDL). The MWL corresponds to a reservoir level of 45.4 m above the upstream toe and the FRL and MDD correspond to a reservoir level of 43 m and 31 m respectively. The computed pore-water pressures at the critical slip surface, for a steady condition at the three water levels, determined the minimum, the most likely (mode) and the maximum pore-water pressures (forces) that are used to build a pdf for the pore-water pressure (pore water force). Further discussion on this is given in chapter 6.

The boundary conditions used for the seepage analysis are as follows:

- On upstream side total head boundary condition is taken, i.e. Dirichet boundary condition or type one boundary condition. The alternative for this was flux boundary condition or Neumann boundary condition. For nodes existing at the bottom of the reservoir and upstream dam face below MWL - the boundary total head is taken to be equal to the elevation head at the top of the reservoir, i.e. elevation head of MWL. The upstream boundary conditions for the three cases of MWL, FRL, and MDDL are therefore constant total heads of 45.4 m, 43 m and 31 m, respectively. The original deterministic design considered a static water level at the MWL for the design.

- At the downstream there is rock toe and under-drain. The granular material in the rock toe is highly permeable, making it reasonable to assume that the water level in the rock toe and under-drain is the same as the water table downstream of the dam. This implies that the permeability is high enough that there is no head loss in the drain relative to the small amount of seepage that will come through the low permeability embankment material. Downstream of the dam toe, the water table is at the ground surface; that is, the water pressure is zero at the ground surface. The water level in the drain is the same as the water table beyond the downstream toe as reasoned earlier. This implies the conditions around the perimeter of the rock toe and downstream of the toe are known. The total head around the rock toe perimeter and beyond is therefore assumed to be the elevation of the original ground surface (OGL), which is 4 m.

- Alternate way of assigning boundary condition for the nodes around the perimeter of the rock toe is to use constant pressure boundary condition. The water table is considered to be at the ground surface; i.e. 4 m head. From seepage analysis point of view, the water pressure is zero at the ground surface. The elevation of each node around the perimeter of the rock toe is different and therefore the total head is different at each node. SEEP/W has a special feature to handle this condition. The

head can be specified with a special condition that makes the pressure $(\frac{u}{\gamma_w})$[3] zero. This is the actual boundary condition used at the downstream toe side in this research. In draw boundary condition menu of SEEP/w if the option pressure head, P is selected and action 0 is set, SEEP/w find the z coordinate of each node and sets the total head equal to the z coordinate.

- When analyzing possible seepage face iterative process is used to determine the size of the seepage face.

- The maximum water level of reservoir is 45.4 m. The ground elevation at upstream toe of the dam is 0 m, i.e. elevation at the upstream toe is taken as datum. The original ground surface is at 4 m; that is, the dam rests on a relatively firm ground after stripping 4 m of the top soil.
- For Tendaho Dam the selected scheme of excavation below core seat provides for complete removal of rock debris, weathered and stained rock on left and right abutment so as to provide a relatively uniform surface on fresh impervious rock (WWDSE et al., 2005b). The valley floors on both left and right banks as well as the section in the river channel part are found to be covered by recent alluvium comprising of silt sand. The depth of this alluvium is found to be varying from 6 to 15 meter. The alluvium is assessed to be overlaying lake sediment deposits comprising mudstone, silt stone and conglomerates. The specified scheme of excavation in this part stipulates removal of the recent alluvium so as to expose the bed of lake sediment deposit. In addition, the foundation is treated by shotcrete soon after excavation to seal fissures and cracks in the bedrock. The core trench in the valley floors and river channel section is covered by 50 mm thick layer of shotcrete and the excavated surface of the core trench on left and right abutments is covered by 75 mm tick layer shotcrete soon after excavation. Moreover, based on borehole water loss tests conducted, extensive grouting, both consolidation and curtain, is applied on the dam abutments and below the core seat (WWDSE et al., 2005b). Therefore, one can fairly assume, in all the five parts of the foundation (right abutment, right bank, river channel, left bank and left abutment), that there is no significant seepage through the foundation below the core seat depth. The seepage analysis in this research has also taken this consideration and took the foundation as impermeable.

5.5 Evaluation of uncertainty of strength and load parameters-stability

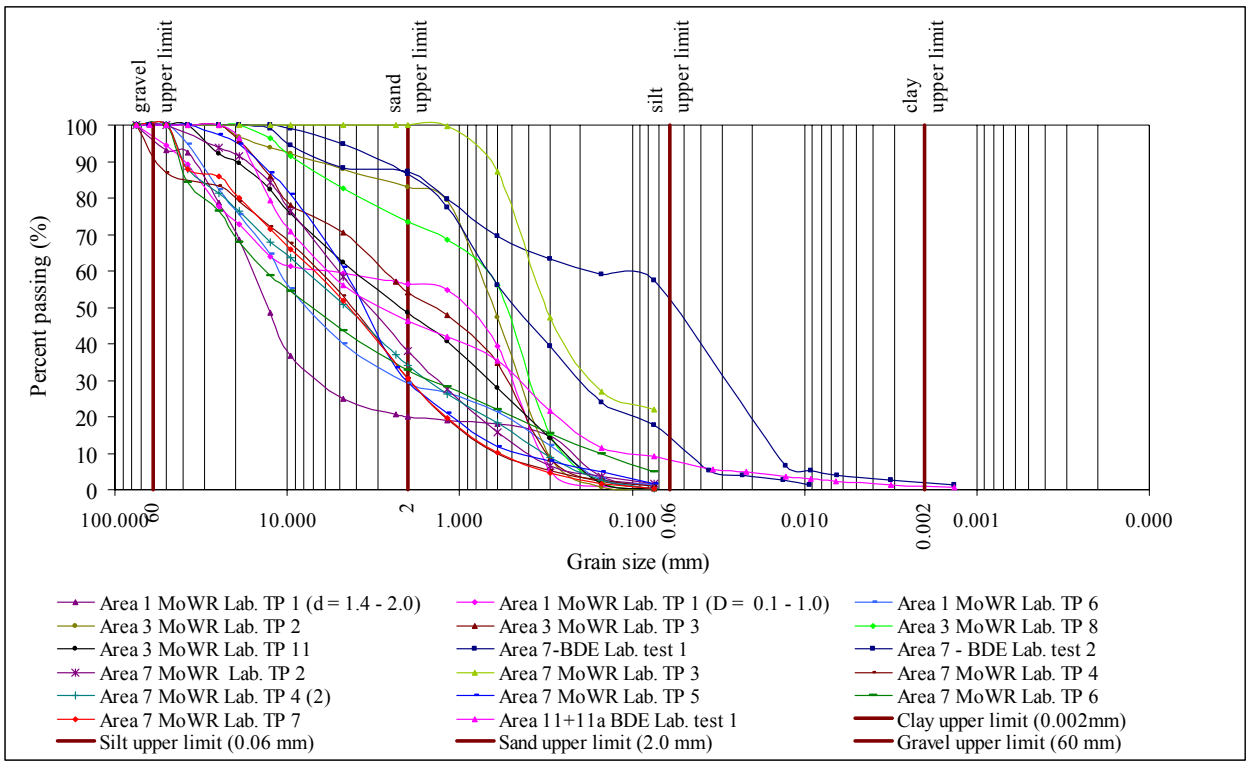
In risk and safety analyses understanding the nature of individual load and resistance parameters randomness is vital to finding methods of modeling their uncertainty. Such an understanding gives confidence to relying on the results of the deployed probabilistic design and reliability analysis. In this and the next sub-sections brief account of randomness in selected

[3] $H = \frac{u_n}{\gamma_w} + z$, where, H is the total head (m), u_n is the pore water pressure (kPa), γ_w is the unit weight of water (KN/m^3), z is the elevation (m)

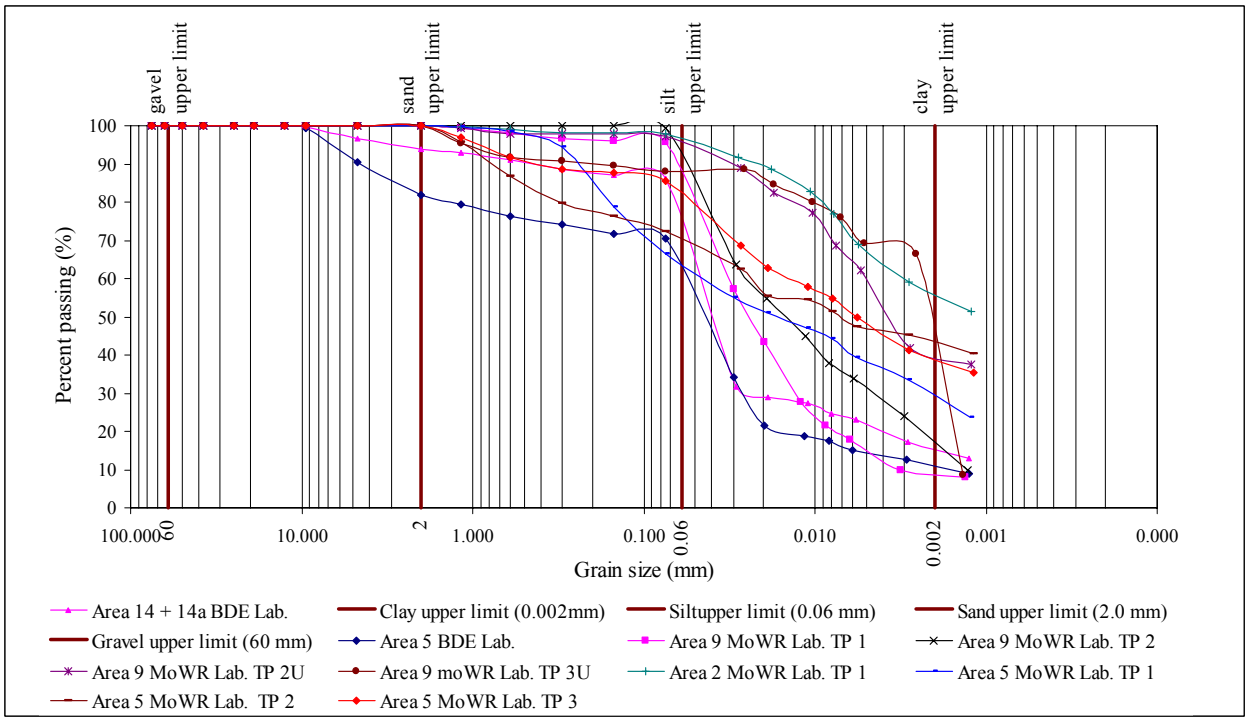
material properties, load and strength parameters are given. Further detailed coverage on randomness of all relevant design parameters, their uncertainty models and representation in risk analysis formulations is given in different sections of chapter 6 and 7.

5.5.1 Uncertainty in fill materials selected engineering properties

As mentioned in section 5.1.1 overwhelming majority of embankment dams are of earthfill dams constructed from range of natural soils, which are the least consistent of construction materials showing high variability. Soil particles vary in size from over 100 mm (cobbles) down through gravels, sands and silts to clays less than 0.002mm size. Naturally occurring soils commonly contain mixtures of particle sizes which mean that the different constituents could behave differently under different loadings. For illustration, grading curves of materials from different borrow areas used to produce the materials for constructing Tendaho Dam are given in Figure 5-13. The figure shows the range of variation in particle size of the materials used for the construction of the core and shell zones of the dam. In addition to the particle size variation, there are many other induced variations in other soil characteristics such as shear strength, permeability, plasticity, bulk density, porosity, etc. Such inherent soil property variations have significant influence on fill materials engineering characteristics influencing strength and loading and associated safety concerns. As an illustration, the randomness in some selected stability and seepage related engineering properties of materials used for construction of the shell zone of Tendaho Dam are presented in Figure 5-14. It is apparent that this randomness will influence the ultimate accuracy of achieving intended functionalities from any design effort.



(a)



(b)

Figure 5-13. Grading curves of materials from different borrow areas based on repeated samples (a) shell material (b) core material.

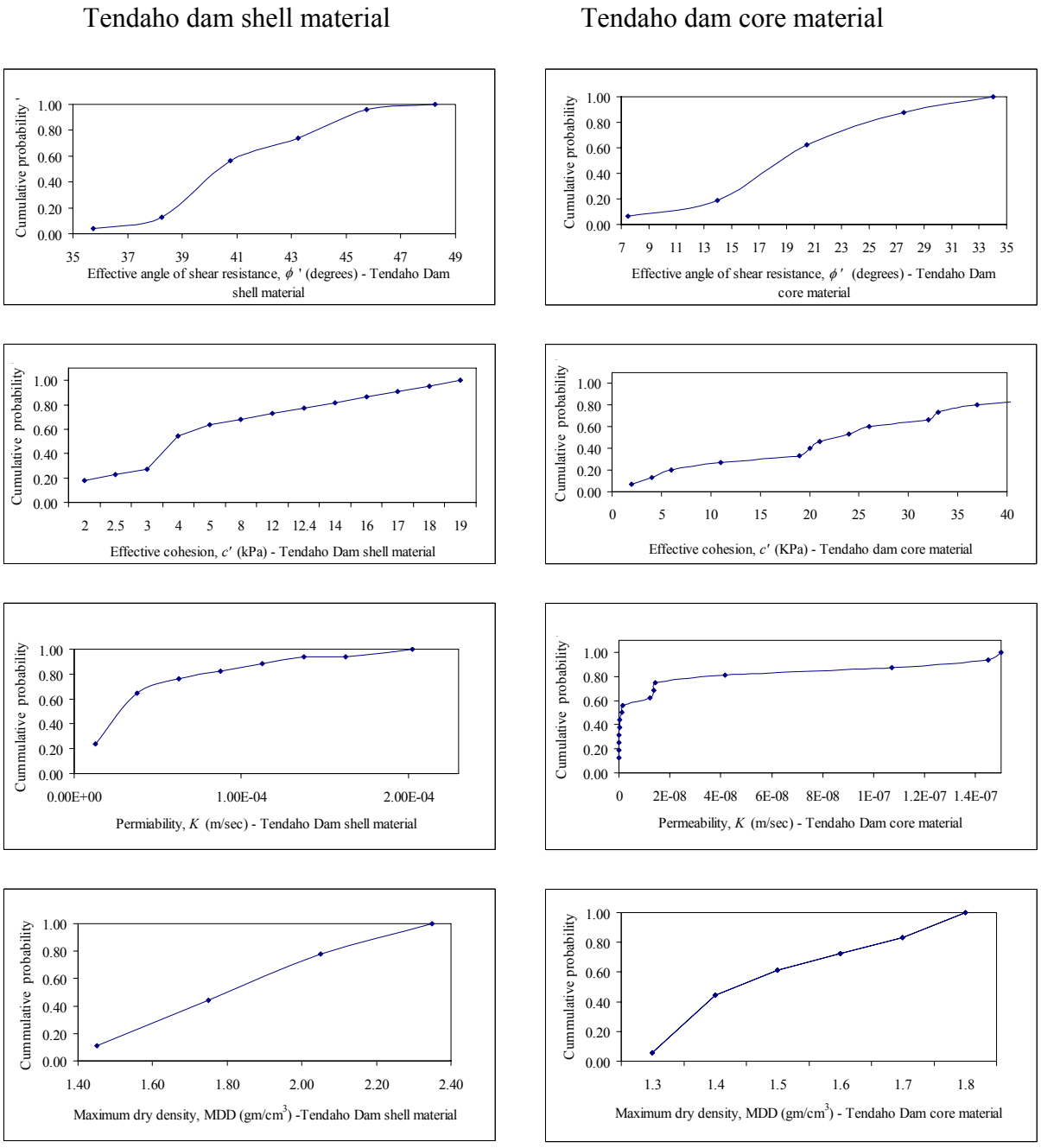


Figure 5-14. Uncertainty in selected engineering properties of Tendaho dam shell and core zone materials.

5.6 Evaluation of uncertainty of strength and load parameters-overtopping

Similarly, in dam overtopping risk and safety analysis there are plenty of uncertain load and strength parameters involved. Among the most important sources of uncertainty that have relevance to the evaluation of dam overtopping risk includes, but are not limited to, uncertainty in peak flood magnitude and its hydrograph, fluctuations in reservoir water levels as a function of

random inflows and outflows, changes in stage-discharge characteristic of the reservoir through time due to sedimentation, randomness in wind speed and waves etc. For instance Figure 5-15 provides the randomness in recorded annual maximum flood data of Awash River at Tendaho Dam site and Figure 7-2 illustrates the randomness in the maximum flood hydrographs using historic records. The randomness in estimated design flood representing extreme event will even further be magnified when trying to estimate the design flood of certain return period using theoretical probability distributions. Further detailed coverage on stochasticity of design parameters related to dam overtopping evaluation is given in section 7.3.

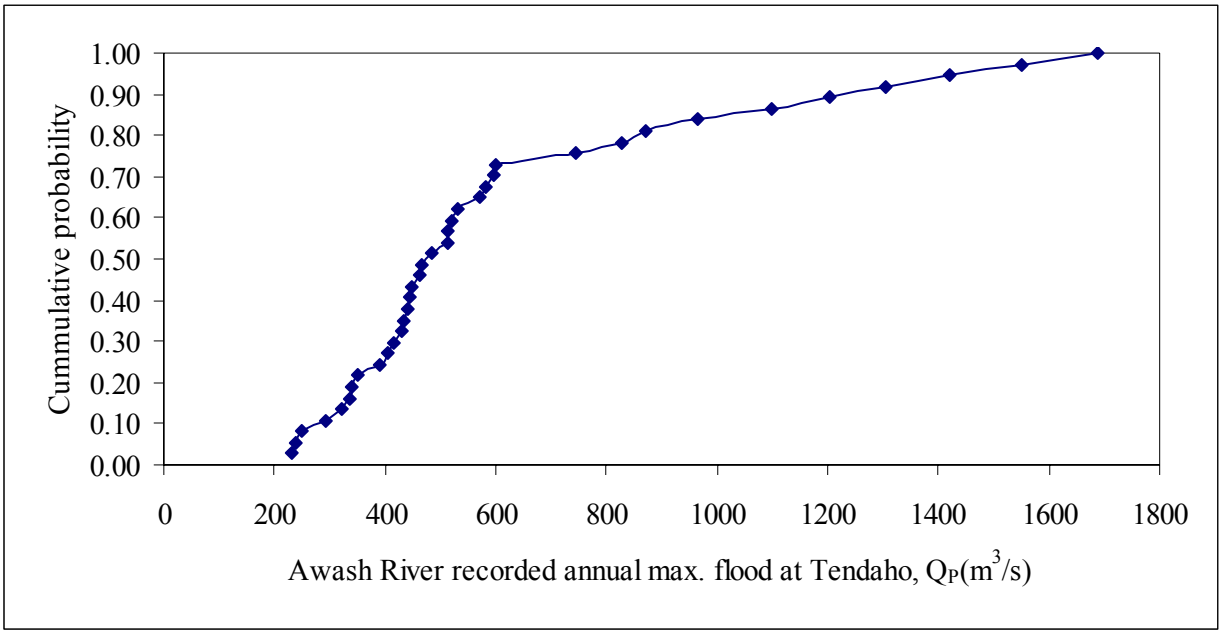


Figure 5-15. Uncertainty in recorded annual maximum flood of Awash River at Tendaho.

The chance of overtopping obviously depends on waves and waves are function of another highly random variable – wind speed. Figure 5-16 provides the uncertainty in the maximum daily wind speed at Tendaho based on historic records over thirty one years. Similarly, Figure 5-17 shows the uncertainty in the amount of sediment flowing into Tendaho dam per annum.

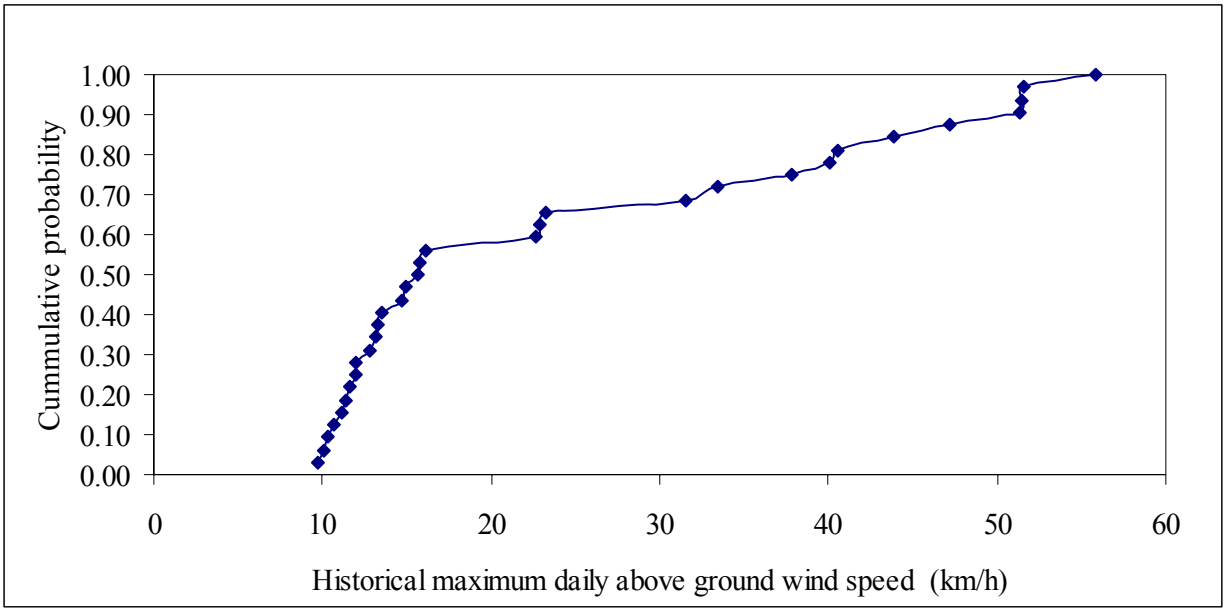


Figure 5-16. Uncertainty in historic records of maximum daily over land wind speed.

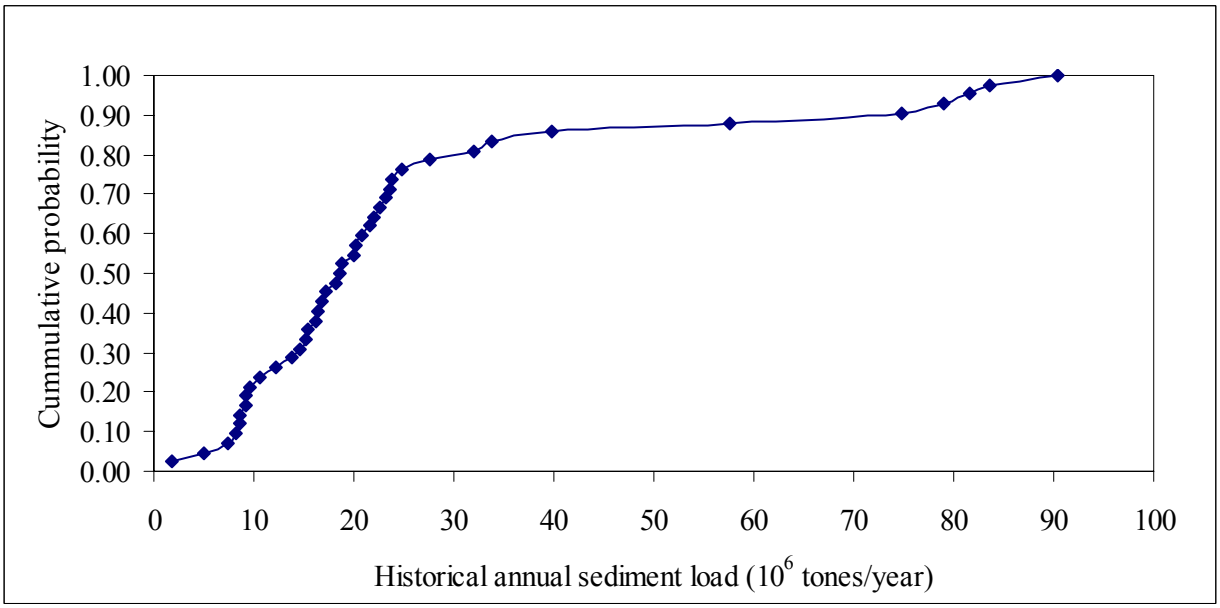


Figure 5-17. Uncertainty in annual sediment load in Awash River at Tendaho.

6 APPLICATIONS OF THE PROPOSED PROBABILISTIC AND ANALYTICAL METHODS - EMBANKMENT DAM STABILITY AND SEEPAGE PROBLEMS

In this chapter the Analytical Solution for determining Derived Distributions (ASDD) method, whose theory has been discussed in chapter 4, is implemented on dam stability and seepage analysis problem. Solution procedures are demonstrated. Results from the ASDD method are compared and discussed with results from deterministic method and the thee classical probabilistic methods discussed in chapters 3; i.e. MCSM, FOSM, and SOSM methods. This chapter aims mainly at demonstrating and comparing applicability of these methods (particularly the new ASDD method) for probabilistic design, and safety analysis of dams.

6.1 Introduction

It has been said that design of dams involves lots of parameters with huge unavoidable uncertainties. Data (material, hydrologic, hydraulic, structural, geophysical, environmental etc.) and models in dam design are more or less characterized by uncertainty. As a natural consequence of uncertain design input variables and models there results uncertain design output. Consequently, designed systems exhibit random performance. This might cause unaccounted risk and surprises owners with problems occurring during operation. Conventional design approaches are deterministic that ignore uncertainty in governing design input variables and that use real numbers or integers for quantifying physical parameters. In other words, deterministic data models are applied. To offset this limitation of the deterministic approach engineers attempt to balance the uncertain parameters with large element of engineering judgment and safety factors. This approach increases the cost of the structure and also lulls the engineer with wrongly perceived absolute safety. To circumvent this problem dams' performance has to be quantified together with specification of its uncertainty. This is possible by specifying the probability of failure (P_f) of a designed system. Determination of P_f requires transparent and accurate method of evaluating uncertainty in design parameters and equations.

In hydraulic design endeavors, if at all there is an attempt for quantifying uncertainty and conduct probabilistic design and safety analysis, it is usually done using approximate methods such as MCSM, FOSM and SOSM (Hartford et al., 2004). For example, (Joos et al., 2005), (Satoh H. et al., 2005), (Melih et al., 2004) and (Lian et al., 2003) has applied one or more of these approximate methods. Such methods could be extremely valuable although they stop short of accurate analytical quantification of performance randomness.

There is a growing understanding that design of dams requires thoughtfulness and more transparency in handling the highly random design input parameters. In recent years work is intensified towards formulating a comprehensive transparent and accurate ways of accounting uncertainty to an extent that makes calculation of performance exceedence probabilities (failure probabilities), risk and safety possible. In this regard, probabilistic methods (stochastic data models) are among the ones believed to provide more realistic solutions, particularly when input data exhibiting pure stochastic properties are considered. These methods help in understanding

and quantifying uncertainty and it allows engineers to get confidence in their engineering judgment.

This chapter primarily demonstrates the application of a new analytical method of computing P_f by using derived distributions. Results from this method are compared with results from MCSM, FOSM, SOSM and deterministic methods. Theories related to the method have been presented in a preceding chapter 3, 4 and in (Negede and Horlacher, 2008a) and (Horlacher and Negede, 2008). In the sections below an attempt is made to promote understanding and to guide through the computation procedures through the use of case studies. As mentioned earlier, for demonstrating and comparing the methods a case study employing stability and seepage performance evaluation of Tendaho Dam is used. Salient features and basic design considerations of Tendaho dam are provided in section 5.3. Laboratory data on the case study dam's fill material characteristics have been collected (see section 5.4).

6.2 Computational framework and case study synthesis

Uncertainty models are topic of controversial discussion in the scientific circle. Möller and Beer (2004) discuss types of uncertainty and different models that can be used for their description. Among the methods for modeling uncertainty in individual design parameters; probability distributions, fuzzy theory and fuzzy randomness are the commonly used. All types of uncertainty may not be stated with sufficient accuracy using only one of these data model. Probability distributions are suitable to describe stochastic uncertainty, which are characterized by pure randomness and containing only objective information that satisfy statistical laws, like the paradigm identically independently distributed and representation of data points as crisp. Fuzzy set theory is suitable to describe informal and lexical uncertainty, which is characterized by subjective influences, like when only small number of observations is available, when boundary conditions are subject to arbitrary fluctuation, when system overview is incomplete. Fuzzy random models are suitable to model uncertainty that exhibit partial stochastic nature with some objective information and some subjective influences. In dam engineering it can be assumed that uncertainty in major design parameters satisfies statistical laws and contain purely objective information, thus, stochastic data models can fairly be adopted. This is equivalent to assuming randomness underlay uncertainty in dam design parameters.

In order to describe randomness in individual design parameters using stochastic data models either historical records are taken or laboratory measurements are repeated (flood, meteorological, geophysical records, material laboratory test data, etc) and lumped together in a crisp data sample. Mathematical statistics offer methods for describing such data samples with aid of random variables (see section 9.1.1.3). A common approach is to specify a pdf in order to obtain stochastic data model (see sections 9.1.1.2 , 9.1.1.3 and appendix 9.1). Design parameters may then be viewed as random variables. Most physical variables used in dam design in fact are random variables. Since dam performance or design equations are mostly functions of one or more of these random design variables they can thus be taken as Function of Random Variable or Vector (FRV). A central tool in safety and risk analysis is, therefore, the ability to compute with pdfs and FRVs by analytically modeling reliability of systems through fault hunting process (failure modes and effect, event tree or fault tree analysis) (see section 2.2). For further detailed discussions on these fault hunting methods reader is referred to (Hartford and Gregory 2004 and

Frankel 1988). As discussed earlier, these methods often use reliability functions. This is a concept where every major performance (failure mechanism) is expressed as reliability function (Z) (see section 2.1). Reliability function (Z) is expressed as a difference between capacity (resistance or strength) (S) and load (L), as given in Eq. (2-1).

Reliability functions are constructed based on reasoning relaying on physics of systems. For example, embankment dam stability is analyzed using limiting-equilibrium models of soil mechanics dealing with stress and strain. Effective shear strength at failure ($\bar{\tau}'$, kN/m^2) can be given using Mohr-Coulomb equation, Eq. (5-3) and effective stress (gravitational driving stress) ($\bar{G}'$, kN/m^2) using Eq. (5-6). Eq. (5-3) and Eq. (5-6) can be rewritten as in Eq.(6-1) and (6-2) providing $\bar{\tau}'$ and $\bar{G}'$ for a defined slice of base length l_n ($l_n = b_n/cos\alpha$) that has a unit width along the dam axis direction (Figure 6-1); i.e. in *KN/m* per l_n *m* slice base which is represented by τ and G (here after in this dissertation τ represents the effective shear strength unless specified otherwise), respectively. This formulation is suitable for writing factor of safety with respect to force equilibrium along slice base, inline with Fellenius or ordinary method of stability analysis (method described in 5.2.1.2).

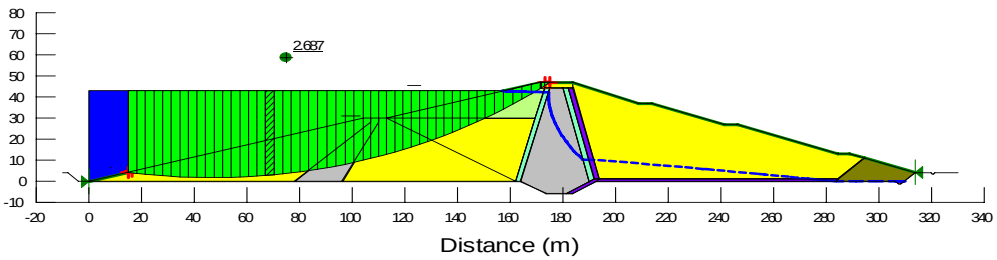
$$\tau = c' \cdot l_n + (\gamma \cdot b_n \cdot h_n \cdot cos\alpha_n - u_n \cdot l_n) tan\phi' \qquad (6\text{-}1)$$

$$G = \gamma \cdot b_n \cdot h_n \cdot sin\alpha_n \qquad (6\text{-}2)$$

Where τ (kN/m) is effective shear strength at the slice base of length l_n, G (kN/m) is shear stress, σ' (kN/m^2) is effective normal stress, c' (kN/m^2), ϕ' (degrees) and γ (kN/m^3) are effective cohesion, effective angle of shearing resistance and unit weight of soil, respectively. u_n (kN/m^2) pore pressure, W (kN) slice weight, b_n (m) slice width along a horizontal direction, z (m) depth of overburden, h_n (m) average height of soil slice and α_n (degrees) inclination of slice base (Figure 6-1). A summation of τ and G running over all slices provide the τ and G situation for entire slip surface.

Thus, sliding reliability (Z_s) can be defined as:

$$Z_s = \tau - G \qquad (6\text{-}3)$$



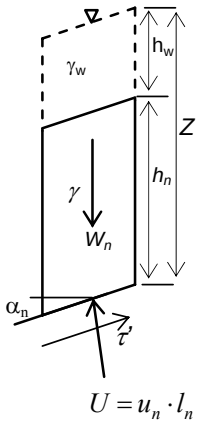


Figure 6-1. Embankment and slice geometry (Tendaho Dam).

And sliding factor of safety ($F_{s,s}$) is given by re-writing Eq. (5-4) as ratio of the effective shear strength to the gravitational driving force for a unit width of slice as:

$$F_{s,s} = \tau / G \qquad (6\text{-}4)$$

Failure occurs when load exceeds or equals strength of a designed system. The limit state situation is usually at $L = S$. Thus, for stability problem cases where $Z_s \leq 0$ or $F_{s,s} \leq 1.0$ define failure. Apparently, sliding failure probability ($P_{f,s}$) can, therefore, be given as:

$$P_{f,s} = P\{\tau - G \leq 0\} = P\{\tau / G \leq 1.0\} \qquad (6\text{-}5)$$

This implies, knowing the cumulative distribution function (cdf) or probability density function (pdf) of Z_s or $F_{s,s}$ makes the determination of $P_{f,s}$ possible. But, Z_s and $F_{s,s}$ are FRV since the load (G) and strength (τ) mechanisms are functions of uncertain design parameters. Hence, the task of determining $P_{f,s}$ deals with computations using FRVs for the purpose of defining cdf or pdf of Z_s or $F_{s,s}$.

Once failure probability (for any failure mood) is computed then the risk (R) is expressed as an expected damage, which is given by the product of failure probability (P_f) and damage (D) as given by Eq. (2-4).

Uncertainty in individual load or resistance design parameters can be modeled with relative ease using probability distributions derived from data sets. As a result, a logical approach for estimating performance uncertainty shall commence from quantifying uncertainty in design parameters and shall progress towards quantifying the extent of randomness in load and resistance mechanisms and their interactions. This is the approach employed in this dissertation. Alternative ways discussed by other authors, for example by (Hartford and Gregory, 2004), includes statistical estimates based on empirical data (past experience) and judgment (intuitive recognition of patterns and process of reasoning).

Table 6.1 and Figure 6-1 provide summary of the case study dam geometric values and scheme of zoning.

As mentioned above, this chapter is mainly devoted for demonstrating the application of a new to the method analytical method of computing P_f, particularly $P_{f,s}$, by using derived distributions. And to compare results from this method with results from three other customary probabilistic MCSM, FOSM, SOSM and from a deterministic method. For the illustration the following cases are considered: (1) Computation of shear strength (τ, kN/m) and shear stress G (kN/m), Eqs. (6-1) and (6-2); (2) Combined effect of τ and G in determining embankment dam stability performance ($F_{s,s}$ and $P_{f,s}$), Eqs. (6-4) and (6-5). (3) Probabilistic evaluation of seepage flux (q, m^3/day/m) through dam core and foundation.

Table 6.1 provides the best fit distributions used for describing random design variables. The last column in this table gives deterministic values of the design variables used during the original design of the case study dam. In this study uncertainty of the random design parameters is modeled using stochastic data models, which are fitted on measured and/or modeled data set.

Pdfs defining material characteristic randomness are selected based on distribution-fitting and χ^2 and Kolmogorov-Smirnov goodness-of-fit test conducted on measured data found from laboratory. The triangular pdf for pore water pressure (u_n) is determined from a deterministic steady state seepage analysis done for three boundary head conditions (reservoir levels); maximum water level (MWL) (45.4 m), full retention level (FRL) (43 m), and minimum drawdown level (MDDL) (31 m). This analysis determined the maximum, minimum and mode parameters for the pdf of u_n. For the deterministic stability and seepage analysis the software GeoStudio 2004 is used.

Table 6.1: Geometry and random variables pdfs.

Geometry		**Assigned pdfs for random variables**					
Variable	Value	Variable	pdf	pdf Parameter			Deterministic
c'(kN/m2)	0			μ	σ		35
h_w (m)	22.8	ϕ'(Degree)	Normal	42.12	3		
h (m)	17.2			Shape, α	Scale, β	Location, L	0.25
α_n (Degree)	6.16	K (m/day)	Gamma	$-1.4\text{x}10^{-5}$	0.009	0.2946	
b (m)	3.26	γ (kN/m^3)	Weibull	13.13	29.8	-11.12	18.5
l (m)	3.28	H (m)	Weibull	28.33	12.68	4.011	43
γ_w (kN/m^3)	9.81			Min.	Mode	Max.	1285.2
n_f/n_d*	11/4	u_n (kN/m^2)	Triangular	270.83	415.87	392.33	

* n_f and n_d are for number of flow channels and equipotential drops respectively. *nf / nd* is called a shape factor and it has implicit length dimension Eq.(6-23)

6.3 Analytical solution (ASDD method)

This is an analytical method for transforming pdfs of random variable(s) to a pdf that defines the randomness in the outputs of FRVs which they constitute.

As mentioned in section 4.2 the question to be resolved using the ASDD method can be mathematically formulated using a generic function and variables as follows: If $\boldsymbol{X}$ is a random vector with n components $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and has a joint cdf $F_X(\boldsymbol{X})$ and $Y = h(\boldsymbol{X})$ is a function of $\boldsymbol{X}$, whose range is a subset of m-dimensional Euclidean space, for some integer m, $1 \leq m \leq n$, then Y is a random vector and has a cdf $F_Y(y)$. We intend to compute $F_Y(y)$, where both the FRV $h(\boldsymbol{X})$ and $F_X(\boldsymbol{X})$, are known. Herein, this problem is referred as distribution problem. Some basic definition on random variables, FRV and multivariate probability terms are given in section 4.1 and complete discussion on the ASDD method is provided in section 4.3. Appendix 9.1.1.3 and 9.1.1.4 can be referred for additional reference on uni- and multi-variate probability theory and modeling of random variables using distribution functions.

In this formulation $\boldsymbol{X}$ can be considered as set of n random design parameters in performance function. And the function $Y = h(\boldsymbol{X})$ can be taken as a performance function. The aim is to determine performance randomness (outputs of $h(\boldsymbol{X})$). This can be achieved by determining either the cdf or pdf of Y ($F_Y(y)$ or $f_Y(y)$). For instance, in evaluating stability

performance, if parameters γ, ϕ' and u_n are taken to be random variables. Thus, Eq.(6-1) is a function of random vectors FRV with $n = 3$ random components $\boldsymbol{X} = (u_n, \phi', \gamma)$. This implies that τ is a random variable and will have a cdf $F_\tau(\tau)$ describing its randomness. The problem of determining $F_\tau(\tau)$ is a distribution problem, which can be solved using the analytical method (ASDD) or the classical approximate methods (MCSM, FOSM and SOSM). Similarly, Eq.(6-2) (G) is uni-variate ($n = 1$) FRV with the random variable $X = (\gamma)$. Here, the interest will be in getting $F_G(g)$.

Once, the randomness in τ and G are explicitly represented by $F_\tau(\tau)$ and $F_G(g)$, then Eq.(6-3) and (6-4) can be considered as FRV with $n = 2$ components $\boldsymbol{X} = (\tau, G)$ and the cdfs $F_{Zs}(Z_s)$ and $F_{Fs,\ s}(F_{s,\ s})$ can be determined. Afterwards, the sliding probability ($P_{f,\ s}$) can easily be extracted from $F_{Zs}(Z_s)$ and $F_{Fs,\ s}(F_{s,\ s})$. Thus, in this way, the entire stability problem can be formulated as a distribution problem.

6.3.1 Load-shear stress (G)

The approach for solving distribution problems can easily be perceived by taking simple uni-variate FRV, for example, the equation for shear stress (G), Eq.(6-2). The theory for distribution transformation involving a uni-variate FRV is given in section 4.3.1.

Before proceeding to the analysis, note that when soil slice is beneath a reservoir of height h_w the water column contributes to the slice weight thus $W = \gamma \cdot h_n \cdot b_n + \gamma_w \cdot h_w \cdot b_n$ (see Figure 6-1). Eq.(6-2) should, thus, be adjusted accordingly to give:

$$G = (\gamma \cdot h_n + \gamma_w \cdot h_w) b_n \cdot \sin\alpha_n \qquad (6\text{-}6)$$

Replacing unit weight of water and relevant geometric values yields:

$$G = h(\gamma) = (\gamma \cdot 17.2 + (9.81) \cdot (22.8)) \cdot 3.26 \cdot sin(6.16) = 0.349 \cdot (17.2\gamma + 223.6) = 6.0 \cdot \gamma + 78.03 \qquad (6\text{-}7)$$

In Eq.(6-7) γ is a random variable with a known cdf $F_\gamma(\gamma)$, which is Weibull distributed (Table 6.1) and G is obviously monotonically increasing uni-variate FRV (Figure 6-2). There is a unique value of G for each value of γ and vice versa. Intuitively, with reference to Figure 6-2, for any value γ_o of γ and for $G_o = h(\gamma_o)$ the probability that $G \le G_o$ equals the probability that $\gamma \le \gamma_o$. Thus we can write:

$$F_G(G) = F_\gamma(\gamma) = F_\gamma\left(h^{-1}(G)\right) \qquad (6\text{-}8)$$

Taking the parameters for the Weibull distribution of γ from Table 6.1 and using the general expression for a Weibull cdf (see Eq.(9-61) in appendix 9.1.2 for equation of Weibull cdf) , the cdf of γ $F_\gamma(\gamma)$ is given as:

$$F_\gamma(\gamma) = 1 - e^{-\left(\frac{\gamma - L}{\beta}\right)^{\alpha}} = 1 - e^{-\left(\frac{\gamma-(-11.12)}{29.8}\right)^{13.13}} = 1 - e^{-4.4\cdot10^{-20}(11.12+\gamma)^{13.13}} \quad (6\text{-}9)$$

From Eq.(6-7) an expression for γ in terms of G, i.e. the inverse of h, can be written as:

$$h^{-1}(G) = \gamma = -13 + 0.166 \cdot G \quad (6\text{-}10)$$

Replacing Eq.(6-10) in Eq.(6-9) gives $F_\gamma(\gamma)$ in terms of G:

$$F_\gamma(\gamma) = 1 - e^{-4.4\cdot10^{-20}(11.12+(-13+0.166\cdot G))^{13.13}} \quad (6\text{-}11)$$

From Eq.(6-8) and (6-11) transformed cdf of the load term G, i.e. $F_G(G)$ can given as:

$$F_G(G) = F_\gamma(\gamma) = F_\gamma\left(h^{-1}(G)\right) = 1 - e^{-4.4\cdot10^{-20}(11.12+(-13+0.166\cdot G))^{13.13}} \quad (6\text{-}12)$$

Noting that first derivatives of cdfs give pdf (see appendix 9.1.1.3 and Table 9.4) differentiating Eq.(6-12) shall yield the pdf of G, $f_G(G)$, which is given in Eq.(6-16). The steps of the differentiation are given as follows.

Note that Eq.(6-12) is a composite function, thus while differentiating Eq.(6-12) chain rule applies[4]. In Eq.(6-12) the expression at the power of e is a function of G itself. Let us replace this expression powering e by $q(G)$,

$$q(G) = -4.4 \cdot 10^{-20}(11.12 + (-13 + 0.166 \cdot G))^{13.13} \quad (6\text{-}13)$$

Replacing this in Eq.(6-16) in Eq.(6-12) allows to clearly seeing the composite function as follows:

$$F_G(G) = 1 - e^{q(G)} \quad (6\text{-}14)$$

Therefore, applying chain rule,

$$\frac{dF_G(G)}{dG} = \frac{d(1 - e^{q(G)})}{dG} \cdot \frac{dq(G)}{dG} \quad (6\text{-}15)$$

Note the following:

$\frac{dF_G(G)}{dG} = f_G(G)$, (first derivative of cdf is pdf, see 9.1.1.3 and Table 9.4)

[4] Chain rule: derivative of a composite function $h(x) = F(\mathrm{G}(x))$ with respect to x is $\frac{dF(G(x))}{dx} \cdot \frac{G(x)}{dx}$

$$\frac{d\left(1-e^{q(G)}\right)}{dG} = -e^{q(g)} = -e^{-4.4\cdot10^{-20}(11.12+(-13+0.166\cdot G))^{13.13}}$$ (derivative of exponential function[5]),

and

$$\frac{dq(G)}{dG} = 9.61\cdot10^{-20}\cdot(-1.88+0.166\cdot G)^{12.13}$$

Substituting the above results in Eq.(6-15) yields the equation for the pdf of G:

$$f_G(G) = 9.61\cdot10^{-20}(-1.88+0.166\cdot G)^{12.13}\cdot e^{-4.4\cdot10^{-20}\cdot(-1.88+0.166\cdot G)^{13.13}} \qquad (6\text{-}16)$$

$f_G(G)$ is plotted in Figure 6-2.

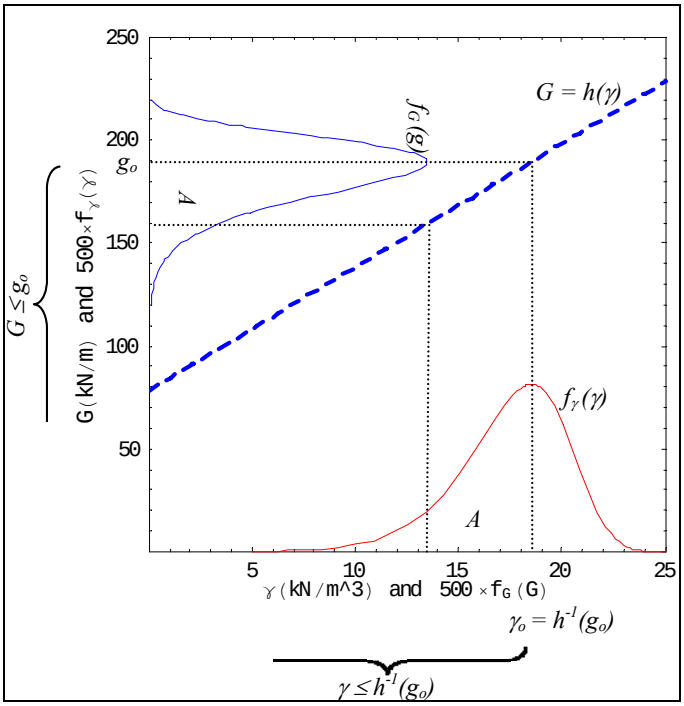


Figure 6-2. Uni-variate FRv of *G*.

The procedure of transforming $f_\gamma(\gamma)$ to $f_G(G)$ is straightforward because it is a uni-variate function of random variable. The generalized relationship is given by Eq.(4-9). Eq.(4-9) can be utilized for such pdf transformations involving continuous uni-variate FRV, without requiring going through all the steps shown above. Concerning basic theories behind uni-variate-FRV further reference can be made to section 4.3.1 and (Negede and Horlacher, 2008a).

[5] Derivative of exponential function $\frac{d(e^x)}{dx} = e^x$

The detailed presentation above is included here only for demonstrating the process and to show the underling logic clearly. Otherwise, straight from Eq.(4-9) and noting that G is a monotonically increasing function (thus $j = 1$ and the summation in Eq.(4-9) reduces to a single term) and one could write:

$$f_G(G) = f_\gamma\left(h^{-1}(G)\right) \cdot \left| dh^{-1}(G) \right| \qquad (6\text{-}17)$$

In here, the inverse transformation of $h(\gamma)$ is $h^{-1}(G) = \gamma = (G - 78.03)/6.0 = -13 + 0.166 \cdot G$ (see Eq.(6-7)). And its derivative $dh^{-1}(G)$ = 0.166. Replacing the inverse transformation $\gamma = -13 + 0.166 \cdot G$ in the Weibull pdf of γ (see Eq.(9-62) for general equation of a Weibull pdf and Table 6.1 for the parameters for γ pdf) and applying (6-17) yields the same $f_G(G)$ as in (6-16). $f_G(G)$ is plotted in Figure 6-2 together with $f_\gamma(\gamma)$.

6.3.2 Strength- shear strength (τ)

As mentioned above Eq.(6-1) is multivariate FRV with $n = 3$ components, $\boldsymbol{X} = (\phi', \gamma, u_n)$. The transformation of the pdf of the random vector $\boldsymbol{X}$ ($f_X(\mathbf{X})$) to pdf of τ ($f_\tau(\tau)$) is a multivariate distribution problem. Eq.(4-13) or Eq.(4-14) can be utilized equivalently for solving distribution problems involving multivariate FRV (see also section 4.3.2). Eq.(4-14) is suitable when there are different number of FRV and RV, which make non square matrix where as Eq.(4-13) is useful only when there are equal number of FRV and RV (i.e. $n = m$ case).

Thus, by using Eq.(4-13) or Eq.(4-14) the pdf bounding the randomness of τ, which is $f\tau(\tau)$, can be derived from $f_X(\boldsymbol{X})$ and the form of functional relation in the FRV for τ Eq.(6-1). Negede and Horlacher (2008a) present a comprehensive proof for Eq.(4-13) and Eq.(4-14) and provided a generic implementation architecture to facilitate understanding by practicing engineers (see also section 4.3.2 for the detailed proof).

As discussed in section 4.3.2 the solution of Eq.(4-13) demands the Jacobian matrix J be a square matrix; i.e. for its solution equal number of random variables and FRVs are needed. This is a requirement because computation of the Jacobian Matrix, which is part of the solution, requires solving determinant. However, in many practical cases there are less number (m) of design equations (FRVs) than random variables (n). Such cases do not satisfy the requirement. It is mentioned that this type of problems can be solved after defining an auxiliary transformation $Y^o = (Y_1, Y_2, ..., Y_m, Y_{m+1}, ..., Y_n)$ made by combining the m components and $n - m$ dummy components (see section 4.3.2). The $n - m$ dummy functions can be selected in such a ways that the necessary computations are as easy as possible and the requirement is satisfied. The dummy functions effect will later be extracted through integration. This process amounts to using Eq.(4-14). The procedure of introducing dummy functions to create an auxialry function that satisfy the $n = m$ requirement and the entire process of calculating an analytically transformed pdf is illustrated in the calculation of pdf for the strength parameter τ, Eq.(6-1).

Let us name Eq.(6-1) as $Y_1 = \tau = h(\phi', \gamma, u_n)$, this is the only FRV available. Thus, there is single FRV ($m = 1$) and three random variables ($n = 3$). Therefore, $m < n$ thus an auxiliary transformation function should be defined by introducing $n - m = 2$ dummy functions so that the requirement $n = m$ satisfied, which is a requirement to solve the determinant in Eq.(4-13) or Eq.(4-14). We can introduce two dummy functions $Y_2 = \phi'$ and $Y_3 = \gamma$. A judicious choice of Y_2 and Y_3 is made so that the Jacobian J is readily evaluated and the sum in Eq.(4-13) or Eq.(4-14) reduces to one term. Therefore, the auxiliary transformation function will be $Y^o = (Y_1, Y_2, Y_3) = (Y_1, \phi', \gamma)$.

Note, for a soil slice beneath a reservoir body of height h_w (see Figure 6-1), $\sigma' = \gamma \cdot h_n \cdot b_n + \gamma_w \cdot h_w \cdot b_n - u_n \cdot l_n$. Noting this and replacing relevant geometric values:

$$Y_1 = \tau = (724.037 - 3.28\, u_n + 55.695\, \gamma) \tan\phi' \qquad (6\text{-}18)$$

Y_1 (Eq.(6-18) is monotonic thus $j = 1$. After finding inverse transformations for $Y_i = h_{ij}(\boldsymbol{X})$ and doing replacement, the family of inverse transformations $X_i = h_{ij}^{-1}(Y_i)$, $i = 1, 2, 3$, that will be used to create a Jacobian matrix are:

$$\begin{gathered} X_1 = h_{1j}^{-1}(Y_i) \Rightarrow u_n = 221.025 + 17\, Y_3 - 0.305\, Y_1 \cdot cot(Y_2) \\ X_2 = h_{2j}^{-1}(Y_2) \Rightarrow \phi' = Y_2 \text{ and } X_3 = h_{3j}^{-1}(Y_3) \Rightarrow \gamma = Y_3 \end{gathered} \qquad (6\text{-}19)$$

Absolute value of the Jacobian of inverse transformations, i.e. Jacobian of $X_i = h_{ij}^{-1}(Y_i)$ is thus:

$$J_j = \begin{vmatrix} \frac{\partial u_u}{\partial y_1} & \frac{\partial u_u}{\partial y_2} & \frac{\partial u_u}{\partial y_3} \\ \frac{\partial \phi'}{\partial y_1} & \frac{\partial \phi'}{\partial y_2} & \frac{\partial \phi'}{\partial y_3} \\ \frac{\partial \gamma}{\partial y_1} & \frac{\partial \gamma}{\partial y_2} & \frac{\partial \gamma}{\partial y_3} \end{vmatrix} = \begin{vmatrix} \frac{0.305}{tan\,\phi'} & \frac{\partial u_u}{\partial y_2} & \frac{\partial u_u}{\partial y_3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0.305\, cot\,\phi' \qquad (6\text{-}20)$$

Substituting inverse transformations given Eq.(6-19) for the X_i's, i.e. substituting the expressions for ϕ', γ, and u_n written in terms of Y_1, Y_2 and Y_3 which is given in Eq.(6-19), in the pdfs of ϕ', γ, and u_n yield the pdf of the X_i's written in terms of the Y_i's as follows (note that the approperate pdfs ϕ', γ, and u_n are named in Table 6.1. and the generic density functions of these pdfs can be found in appendix 9.1.2):

$$f_\gamma(Y_3) = 5.79 \cdot 10^{-19}\; e^{-4.4 \cdot 10^{-20}(11.12 + Y_3)^{13}} (11.12 + Y_3)^{12.13}$$

$$f_\phi(Y_2) = \left(e^{-\frac{1}{18}(-42.12 + Y_2)^2} \right) / 3\sqrt{2\pi}\,, \text{ and}$$

$$fu_u(Y_1,Y_2,Y_3)=\begin{cases}-5.6\cdot10^{-3}+1.9\cdot10^{-3}\,Y_3-3.5\cdot10^{-7}\,Y_1\cdot cot(Y_2),\\ 270.85\le 221+17\,Y_3-0.31Y_1\cdot cot(Y_2)\le 392;\\ \\ 0.11-9.97\cdot10^{-3}\,Y_3+1.79\cdot10^{-4}\,Y_1\cdot cot(Y_2),\\ 392<221+17\,Y_3-0.31Y_1\cdot cot(Y_2)\le 415.87;\\ \\ 0,otherwise\end{cases}$$

Assuming ϕ', γ, and u_u are independent random variables their joint pdf $f_{\boldsymbol{X}}(\boldsymbol{X})$ can be given as $f_{\boldsymbol{X}}(\boldsymbol{X}) = = f_{u_u}(u_u)\cdot f_\gamma(\gamma)\cdot f_{\phi'}(\phi')$ (see Eq.(9-15)). Therefore, pdf of transformation function Y^o ($f_{Y^o}(\boldsymbol{y})$) follows as:

$$f_{Y}^{o}(\boldsymbol{y})= fu_u(Y_1,Y_2,Y_3)\cdot f_\phi(Y_2)\cdot f_\gamma(Y_3) \qquad (6\text{-}21)$$

This completes the steps required for applying Eq.(4-14). Applying Eq.(4-14), yields:

$$f_Y(y)=f_\tau(\tau)=\sum_j\int_0^\infty\int_0^\infty f_{Y^o}(Y)\cdot|J_j|\,dy_2dy_3 \qquad (6\text{-}22)$$

Where $j = 1$. Lower limits of integration are zero because minimum possible values for $Y_2 = \phi'$ and $Y_3 = \gamma$ are zero. All the computations for the analytical method are done using the computer program Mathematica 5.2. The expanded form of Eq.(6-22) is long, because of the step function in $fu_u(Y_1, Y_2, Y_3)$, and it is unpleasant for inclusion in here, it is therefore omitted. However, its plot is given in Figure 6-3 and the complete equation can be found in the appendix CD-ROM in the file (:\slope stability-seepage analysis\ASDD method\single slice ASDD τ-G calculation.nb). Figure 6-4 compares the pdf of τ computed using the analytical method with pdfs estimated by other approximate methods discussed in section 6.4 below. In Figure 6-5 $f_\tau(\tau)$ and $f_G(G)$ are plotted on one axis. Deterministic values of τ and G, which are also used in the original design of the dam, are shown by the vertical lines at τ = 325 kN/m and G = 189 kN/m.

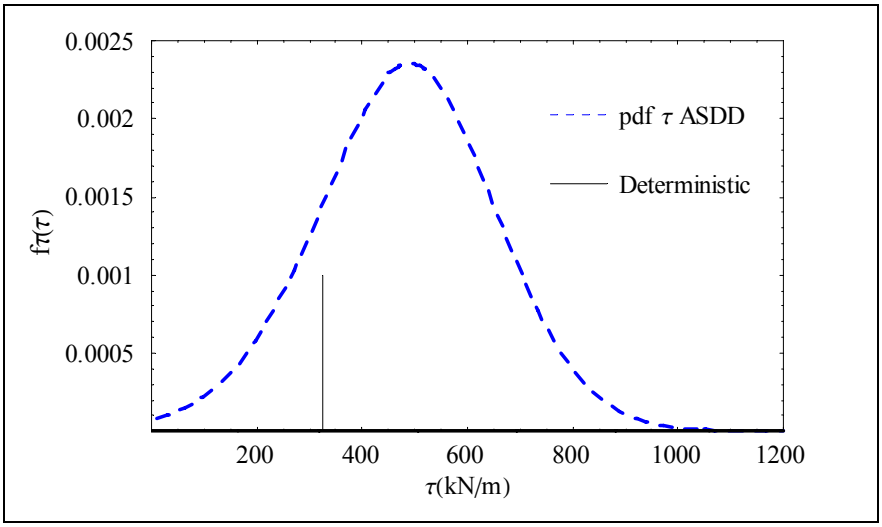


Figure 6-3. pdf of τ.

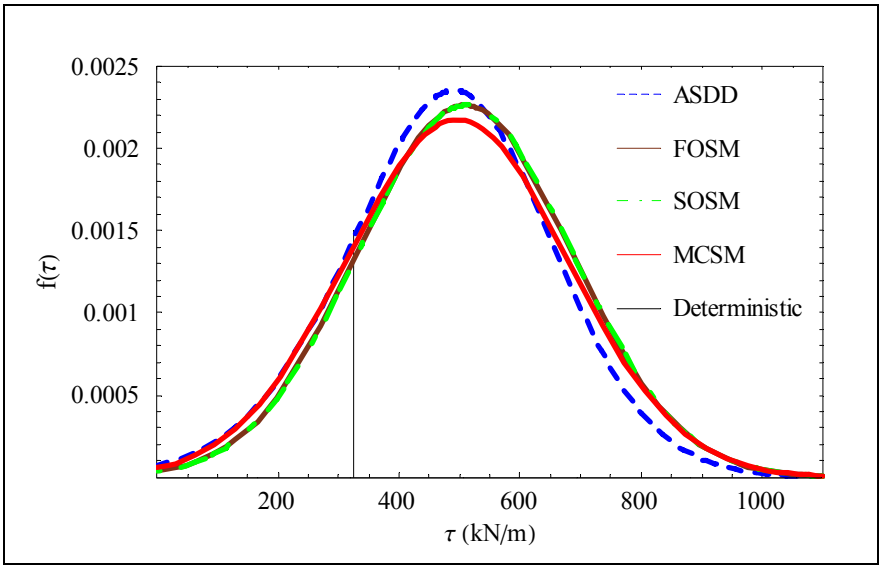


Figure 6-4. pdf of τ - all methods.

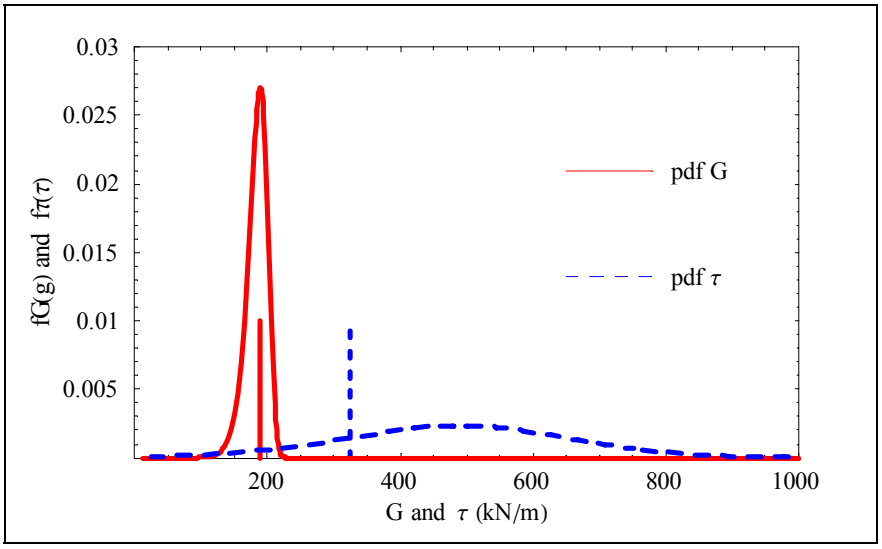


Figure 6-5. pdf of G and pdf of τ.

6.3.3 Stability performance $f_{Fs,\,s}(F_{s,\,s})$ and $P_{f,\,s}$

Once (τ) and (G) randomness are explicitly represented by $f_\tau(\tau)$ and $f_G(G)$. Then, the randomness in Z_s or $F_{s,\,s}$ can be evaluated by taking Eq.(6-3) or Eq.(6-4) as FRV with $n = 2$ components $\boldsymbol{X} = (\tau, G)$. $f_{Zs}(Z_s)$ or $f_{Fs,\,s}(F_{s,\,s})$ can be evaluated using Eq.(4-14), as illustrated using the case of τ. The detailed computation and codes for doing them are given in appendix CD-ROM in the Mathematica file (:\slope stability-seepage analysis\ASDD method\factor of safety sliding ASDD.nb).

The computed pdfs, either $f_{Zs}(Z_s)$ or $f_{Fs,s}(F_{s,s})$, allows for specification of stability performance (sliding failure probability $P_{f,s}$). Figure 6-6 provides computed pdfs for the factor of safety ($F_{s,s}$). The sliding failure probability is given by the value $P_{f,s} = P\{F_{s,s} < 1\} = 0.0435$. The deterministic factor of safety value is given by the line at $F_{s,s}$ = 1.72 and it corresponds to ≈14% quantile.

6.4 Classical probabilistic methods solutions (MCSM, FOSM, SOSM methods)

6.4.1 FOSM and SOSM solutions

General relations for FOSM and SOSM approximations of the mean value (μ_Y) and variance (ν_Y) of a multivariate FRV $Y = h(\boldsymbol{X})$ where X = (X_1, X_2, … , X_n) is continuous random vector, is given by Eq.(3-14), Eq.(3-21), Eq.(3-22) and Eq.(3-23), respectively. For prove of these equations refer section 3.2.4. Example calculations on applications of these methods in stability analysis is given in appendix 9.4. Mathematica codes for doing the complete computations are given in the appendix CD-ROM in the file (:\slope stability-seepage analysis\FOSM-SOSM methods\FOSM-SOSM sliding-seepage plus their plots together with ASDD-MCSM.nb)

For example, by replacing Z_s for h in Eq.(3-14) and Eq.(3-22) and using values of mean and variance of variables, SOSM and FOSM approximation to the mean and variance of Z_s can be computed (see appendix 9.4 for calculation examples). Mean and variance of random variables (ϕ', γ, u_u, c') can be calculated from their respective pdfs given in Table 6.1. FOSM and SOSM approximations of mean and variance of $F_{s,s}$ can be computed in a similar fashion (see appendix 9.4.4 for calculation details). Results for μ_{Fs} and $\sqrt{\nu_{Fs}}$ found from FOSM and SOSM analysis are given in Table 6.2. Figure 6-6 provides the pdf and cdf for $F_{s,s}$ constructed using the computed $\mu_{Fs,s}$ and $\sigma_{F,s}$ and assuming Central Limit Theorem (CLT). In the same figure results from other methods are also provided.

6.4.2 MCSM solution

To determine randomness in $F_{s,s}$ Monte Carlo Simulation has been run 10,000 times using the computer program Crystal Ball 7.2 on a computational framework developed on spreadsheet (see in the appendix CD-ROM under file (:\slope stability-seepage analysis\MCS methods\MCSM F_s sliding crystal ball.xls). This simulation suggested $F_{s,s}$ is a normal distributed with $\mu_{Fs,s}$ = 2.66 and $\sigma_{Fs,s}$ = 0.87 (Table 6.2 and Figure 6-6).

6.5 Seepage flux

Similarly, randomness in seepage flux (q, m^3/day/m) through the case study dam core and foundation has been evaluated using the same set of probabilistic methods (FOSM, SOSM, MCSM and ASDD). For this, the classical flow net representation of Darcy's seepage equation, for an isotropic condition, has been utilized Eq.(6-23). The upstream boundary head condition

(H, m) and core permeability (K, m/day) are taken to be uncertain variables. Figure 6-7 and Table 6.3 provide results from this analysis.

$$q = KH(n_f / n_d) \qquad (6\text{-}23)$$

The detailed Mathematica codes, spreadsheets computations and plots for seepage flux randomness according to the different stochastic methods are given in the CD-ROM in the files (:\slope stability-seepage analysis\ASDD method\seepage analytical method final (Q m^3 per day).nb); (:\slope stability-seepage analysis\FOSM-SOSM methods\FOSM-SOSM sliding-seepage plus their plots together ASDD-MCSM.nb); and (:\slope stability-seepage analysis\MCS methods\MCSM seepage flow crystal ball.xls).

6.6 Comparison of results

Results ($F_{s,\,s}$ moments and failure probabilities $P_{f,\,s}$) from all probabilistic methods are provided in Table 6.2. And $F_{s,\,s}$ distributions are plotted in Figure 6-6. The computations and equations of the plot given in Figure 6-6 are available in (:\slope stability-seepage analysis\ASDD method\plots.nb) and (:\slope stability-seepage analysis\FOSM-SOSM methods\FOSM-SOSM sliding seepage plus their plots together with ASDD-MCSM.nb).

Table 6.2 and Figure 6-6 show that there is a reasonable match among the probabilistic estimates for $F_{s,\,s}$. It can be clearly seen that the absolute safety at $F_{s,\,s} = 1.72$ estimated by the deterministic method is just naive. There actually is a sliding failure probability of $P_{f,\,s} = P\{F_{s,s} < 1\}$ that ranges between *2.4* and *4.35 %.*

Economic optimization, thus intelligent rehabilitation and upgrading decisions, and risk analysis are possible by using the computed $P_{f,\,s}$.

In the case of q, results from FOSM and SOSM showed divergence from that of ASDD (Figure 6-7). This is related to error induced by assumptions of CLT that states interaction of independent but not identically distributed parameters tend to be normally distributed (when no one distribution is dominant). This assumption is not plausible here, because K has highly skewed and dominant distribution. The appropriateness of this assumption can be evaluated by computing skew and higher order moments for checking symmetry and squatness. Here, MCSM also showed problem of convergence. In such circumstances ASDD gives the most reliable result. Computation cods and plots of Figure 6-7 are given in the appendix CD-ROM in the file (:\slope stability-seepage analysis\FOSM-SOSM methods\FOSM-SOSM sliding seepage plus their plots together with ASDD-MCSM.nb).

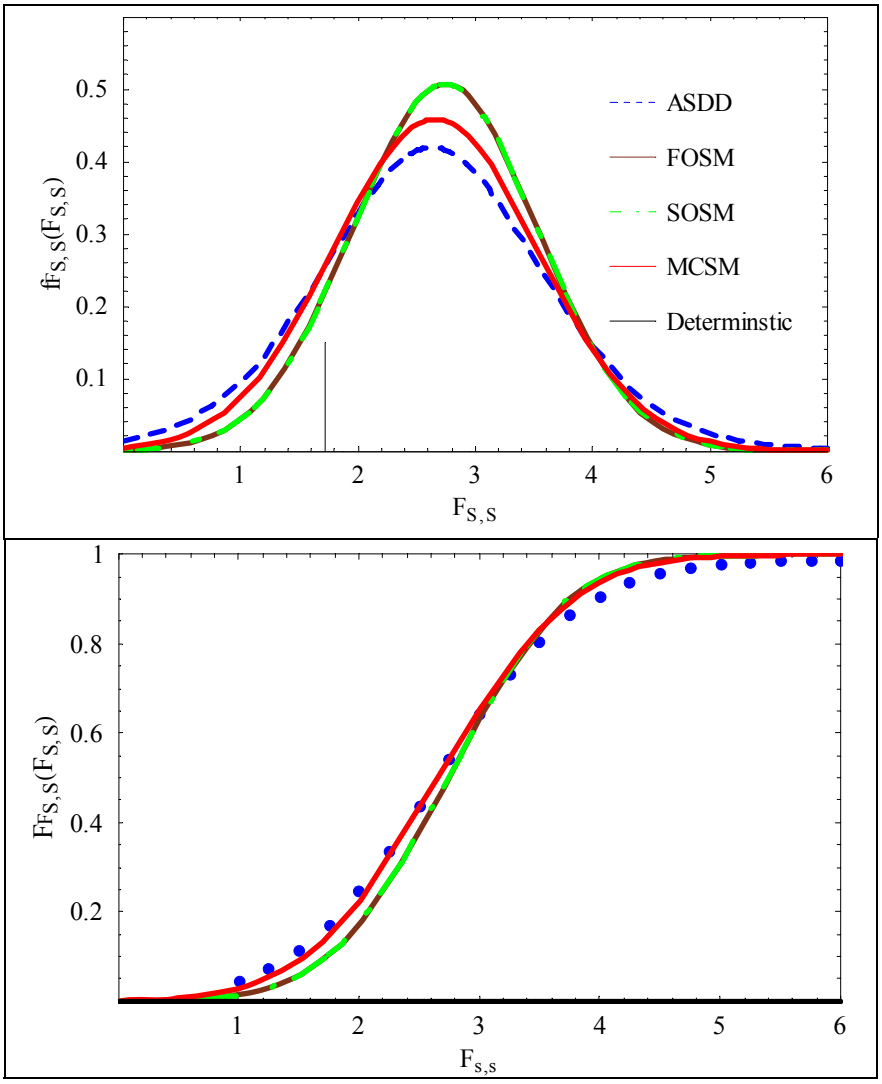


Figure 6-6. Computed pdf (top) and cdf (bottom) of $F_{s,s}$ – all methods.

Table 6.2: Computed moments of $F_{s,s}$ and $P_{f,s}$.

Methods	Factor of Safety		$P_{f,s}$
	$\mu_{Fs,s}$	$\sigma_{Fs,s}$	
Analytical	2.62	0.98	0.0435
SOSM	2.75	0.79	0.0243
FOSM	2.75	0.79	0.0248
MCSM	2.66	0.87	0.0285
Deterministic	1.72	0	0

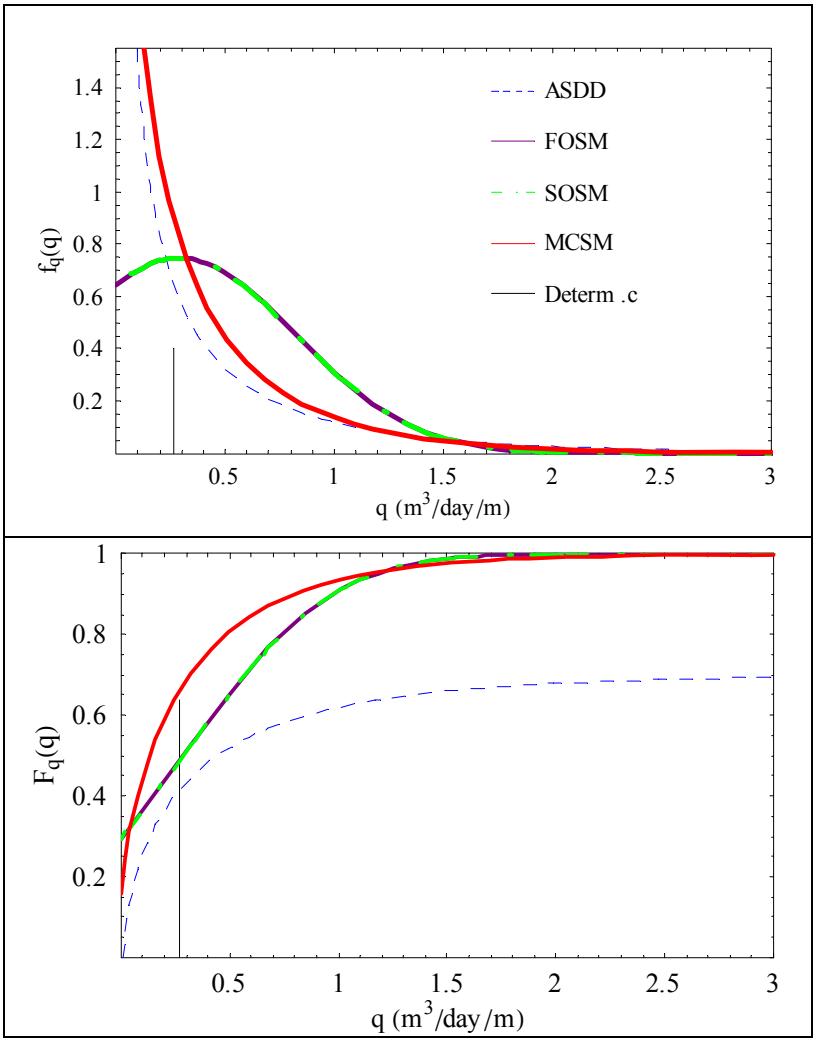


Figure 6-7. Computed pdf (top) and cdf (bottom) for q – all methods.

Table 6.3: Computed moments of q.

Methods	Seepage (q, m^3/day)	
	μ_q	σ_q
ASDD	0.2647	0.44
SOSM	0.2896	0.53
FOSM	0.2896	0.53
MCSM	0.3037	0.41
Deterministic	0.6784	0

6.7 Chapter conclusions

The discussed probabilistic and analytical methods are applicable for the computation of P_f with varying precisions. The methods help to avoid mistakes by providing realistic solutions. The analytical method out performs the classical methods by its theoretical foundation, accuracy

and reproducibility. It is a good practice to compare results from the different probabilistic methods. The lack of demonstrated analytical methodology to meet the growing demand for precision in risk and safety analysis of hydro-systems makes this work important. Most classical modeling tools in hydraulic engineering help to enhance engineering judgment but do not enhance predictive capacities in same degree. However, by combining modeling and probabilistic analysis both engineering judgment and prediction ability could be enhanced. The procedure presented can be adopted in many such practical applications.

7 APPLICATIONS OF THE PROPOSED PROBABILISTIC AND ANALYTICAL METHODS - DAM OVERTOPPING PROBLEM

7.1 Introduction

Overtopping accounts for about 30-35% of dam failures world wide (Novak et al., 2003), (ICOLD, 1995), (Webb, 2006), (ASDSO, 2009) and (Pohl, 1999). Consequently, there are several published attempts towards establishing methods for estimating $P_{f,\ OT}$. Many of these attempts used MCSM mostly with Q_P and sometimes h_o and wave generating wind too assumed as RVs characterized by probability distributions (pdf). For instance, (Phol, 1999) evaluated $P_{f,\ OT}$ using MCSM taking Q_P, wind and h_o as uncertain parameter. He assumed that Q_P is the only RV affecting FH. Kuo et al. (2007) applied five uncertainty analysis methods for assessing dam $P_{f,\ OT}$, namely Rosenblueth's Point Estimate, Harrs's Point Estimate, MCSM, Latin Hypercube Sampling and the Mean-value First-order Second-Moment methods. The most classical return period method has been extensively used for decades although it considers only flood frequency. This later method literally uses deterministic approach. Often, in similar attempts for estimating $P_{f,\ OT}$. influences of other flood hydrograph shape (FH) parameters than Q_p are not accounted. In the present study, FH is modeled using an existing gamma function. Effects of taking all or part of FH parameters as RV is evaluated. Besides, in the process of $P_{f,\ OT}$ evaluation wave parameters and h_o (as influenced by continuous steady river flow, abstractions, evaporation, environmental release and sedimentation) are considered as additional RVs.

The generally accepted and frequently practiced concept of hydraulic design, also summarized in (Pohl, 1999), is to select a deterministic value for design parameters and to do calculations using limit state equations. The calculations are made with intent of giving designed dimension that guarantee "sufficient factor of safety" against limit state situations. Iterations of this procedure may be made by changing selected deterministic values of design variables, so that structures are dimensioned in such a way that a subjective balance between safety and economy is set. In case of design against overtopping such an approach uses a deterministic Q_P of certain return period characterized by deterministic FH. However, because of host of uncertainties involved in hydrologic, meteorological, hydraulic, structural, geotechnical parameters and models such an approach does not allow to see the entire horizon of designed magnitudes and built-in uncertainties. This leave safety unenumerated and does not allow setting an objective balance between safety and economy.

To capture the entire range of a design output representing dams' performance, uncertainties in load and strength mechanisms has to be characterized using stochastic data models. This could be derived from measured uncertainties of relevant design parameters and functional relations in design equations. Such an undertaking is not a straight forward procedure, especially when multi-variate function of RVs has to be processed, like in the case of $P_{f,\ OT}$ evaluation. To solve such problems researchers mainly resort to the method of statistical trials-MCSM. In this study, besides MCSM; FOSM, SOSM and ASDD methods are utilized for the computation of $P_{f,\ OT}$. The ASDD method, which has been described in chapter 4 and in (Negede

and Horlacher 2008a), enables to drive a pdf characterizing performance randomness from design variables pdfs'.

The $P_{f,\,OT}$ calculation procedure using these set of stochastic methods procedure is applied on a case study dam (Tendaho dam) on Awash River, Ethiopia. The dam is designed to irrigate nearly 60,000ha of sugar cane plantation. Additional information on the case study dam and River Awash is given in section 5.3.

7.2 The framework

In this thesis dam overtopping failure is defined as the case where reservoir water level or individual waves exceed the highest water tight level regardless of whether damages to the integrity of the dam occurs or people are injured or die. It is assumed that there is a dam of constructed height (h_d) holding a steady initial water level (h_o). h_o depends on the continuous steady river inflow, abstractions, evaporation and environmental releases. It could be anywhere between the minimum drawdown (MDDL) and full retention levels (FRL). Sedimentation has a role in determining the likely h_o because operators tend to keep water levels high to compensate for the storage lost due to sedimentation. The freeboard, vertical distance between h_d and h_o, is assumed to accommodate the following loads: (1) flood surcharge (h_f), i.e. the rise in h_o due to flood routing. (2) wind set-up (h_s), tide due to shear induced by continuous wind in one direction. (3) wave height and run-up (h_r). (4) seiche effects (h_i), periodic undulation of reservoir caused by earthquake and pressure changes, disregarded in medium sized reservoirs like Tendaho. Based on these definitions overtopping phenomena can be expressed in terms of reliability function (Z); i.e. as a difference between strength (S) and load (L) (see section 2.1, Eq. (2-1)). In this formulation the water tight dam height h_d is the S and the sum of the heights h_o, h_f, h_s, h_r and h_i making up the frontal water level is the L.

$$Z = S - L = h_d - (h_o + h_f + h_s + h_r + h_i)) \qquad (7\text{-}1)$$

Hence, $L \geq S$ or $Z \leq 0$ defines failure. Overtopping failure probability ($P_{f,\,OT}$), given extreme events selected for design, can therefore, be given as:

$$P_{f,OT} = P\{Z \leq 0\} = P\{S - L \leq 0\} = P\{h_d - (h_o + h_f + h_s + h_r + h_i) \leq 0\} \qquad (7\text{-}2)$$

The return period of a design parameter is defined as the period of time (T, in years) in which the parameter occurs again on average. Therefore, assuming that each year with in a period of T years has equal chance and are independent, the annual exceedence probability of selected design extreme event of 1 in T years return period is $\frac{1}{T}$. This is valid because the sum of exceedence probability each year with in a period of the T years should add up to 1. Thus, non-exceedence probability of selected design extreme event is thus $1-\left(\frac{1}{T}\right)$. The non-exceedence probability in n years dam life is $\left(1-\left(\frac{1}{T}\right)\right)^n$. Hence, the probability (R) that extreme events occur

at least once in n years is $R = 1 - \left(1 - \left(\frac{1}{T}\right)\right)^n$. Therefore, the probability of overtopping, at least once in n years dam life, is given as:

$$R_{f,OT} = P_{f,OT} \cdot R \qquad (7\text{-}3)$$

We know that strictly speaking the load terms in Eq.(7-1) are functions of uncertain design parameters. The three main load components of freeboard are uncertain parameters: (1) h_f (influenced by Q_P, base flow Q_b, FH, and spillway feature), (2) h_s and h_r (influenced by reservoir depth, wind speed (U) and fetch, angel of wind to the fetch, geometry and material of the dam's upstream face), (3) h_o (influenced by prior abstractions, draught and flood conditions). Strictly speaking, strength h_d is uncertain parameter too. It is influenced by settlement and other geotechnical phenomenon although this is not considered here.

Consequently, Eq.(7-1) is a function of random variables (FRV). Hence, determining Z's pdf and $P_{f,\ OT}$ involves computation with FRV. The uncertainty in individual L and S parameters (U, Q_p, Q_b, and other FH variables) can be modeled with relative ease using pdfs derived from data sets. As a result, a logical approach to evaluate $P_{f,\ OT}$ shall commence from quantifying uncertainty in individual design variables and move towards estimating randomness in overtopping related L and S mechanisms and their interactions. Overtopping related RVs, their uncertainty, interactions, analysis and applied methods of modeling are discussed next.

7.3 Random variables, analysis and representation methods

7.3.1 Flood surcharge (h_f)

The value of flood surcharge (h_f) depends on Q_p magnitude, *its flow characteristic (FH shape), flood routing and spillway*-capacity, hydraulics, availability, operation. In this study, spillway related issues are treated deterministically. For more comprehensive probabilistic representation reduction in spillway effective discharge, due to gates mechanical failure, blockage and operational failures can be taken as RVs. Discussions and values on failure frequencies of gates and valves at dams can be found in (Pohl, 2000). He analyzed data on 536 gates and valves on dams and found that in about $5^{o}/_{oo}$ (p = 0.005) of all cases a regular gate opening was not possible during inspection and in about p = 0.0025 of all cases a regular opening was not possible during flood requiring spillway operation, including complete and partial failure to open, gate in revision and delayed opening. (Lecornu, 1998) estimated availability of gates and other dam measurement and operational instruments. According to him the non-availability of gates is $3.4^{o}/_{o}$ (p = 0.034). However, this includes the time to repair as well as revision and it is not exactly equivalent to failure-to-open probability in case of demand. In $P_{f,\ OT}$ analysis gate status can be given as discrete distribution (open/close/partially open) or as a continuous pdf of discharge reduction coefficient C_d. Thus, for instance, in each cycle of MCSM trials C_d will take a certain value between 0 (closed) and 1 (open). This approach is applied in (Pohl, 1999 and 2000). In this paper, Q_P and its FH shape are taken as RVs. The Q_P randomness is estimated from frequency and confidence interval (CI) analysis. FH shape is a function of Q_P, its shape

parameters-time to peak (t_p), time to centroid (t_g) value, and the base flow (Q_b). Its randomness is derived from the shape parameters. FH randomness is derived from these parameters uncertainty.

7.3.1.1 Randomness in Q_P

Flood is exceedingly complex hydrologic process resulting from interaction of number of component parameters: catchment characteristics- geomorphologic, geological; meteorological variables; antecedent moisture content; upstream water and land use/cover change, each further depending on multitude of constituent uncertain parameters. This makes problem of estimating flood peaks a difficult task leading to different approaches. Among methods used in moderate and large catchments are: empirical and regional formulae and unit hydrograph (UH) methods. Another alternative is the method of frequency analysis using measured flood data. The later method is applied in this study. Further references on this topic can be found in (ICOLD, 1992), (Chow, 1951), (Gringorten, 1963), (Gumbel 1958, 1943, 1942, 1941) and (NERC, 1975). 35 years , 1965-2002 except 1994 and 1995, of annual maximum flood data of Awash River at Tendaho (drainage area 62,088 km^2) has been extracted from daily flow data series obtained from Ethiopian Ministry of Water Resources(MoWR) and (Halcrow, 2006 and 1989), (WWDS, 2005a) (see Figure 5-8 and Appendix 9.3.4).

Flood frequency analysis involves large extrapolation of return periods (T) using theoretical pdfs. This is customary in dam design because applicable design recurrence periods are much longer than available length of records. For example, at Tendaho 35 years record is used to estimate the T = 10,000 years design flood ($Q_{10,000}$). Among the several theoretical pdfs used for such extrapolations Gumbel (EV-I) and Log-Pearson-III are widely used (Ponce, 1989). The annual maximum series at Tendaho is fitted to three pdfs: Log-Normal, Gumbel (EV-I) and Log-Pearson-III (see Figure 7-1). χ^2 and Kolmogorov-Smirnov data fit tests showed that EV-I gives the best fit. The estimated Q_P magnitude ($Q_{10,000}$ magnitude in this case) dramatically depends on the pdf chosen to describe the data. Variates estimated from theoretical pdfs fitted using limited sample data consists induced errors. Such estimates only give central value somewhere between possibilities. Hence, showing the confidence interval (CI) that give limits for Q_P between which its true value may lie with specific probability is desirable. In this regard (Chow 1951) has shown that pdfs applicable in hydrologic studies can be expressed by the general equation of hydrologic frequency analysis:

$$X_T = \bar{x} + K \cdot \sigma \qquad (7\text{-}4)$$

Where X_T, $\bar{x}$, σ are, the T years recurrence value, mean and standard deviation of the variate X. K is frequency factor that depends upon T and the assumed pdf. To describe the randomness of EV-I estimated Q_P base on the confidence interval analysis a normal distribution (natural law of errors) is assumed (see Figure 7-1). The normal pdf is fitted using the 95% CI upper and lower limit quantiles computed using Eq.(7-4). For this task Crystal Ball 7.2 software is used. This software allows fitting a normal distribution based on known quantiles.

Two Q_P magnitudes are selected for use in the stochastic overtopping analysis, these are the Q_P for T = 10,000 years, ($Q_{10,000}$) and probable maximum flood (Q_{PMF}). The selection is made based on standards in (ICOLD, 1992), (ICE, 1998) and assumptions in the original design. The

estimated $Q_{10,000}$ and Q_{PMF} means are 3,535 m^3/s and 6,084 m^3/s respectively. $Q_{10,000}$ is found from frequency analysis and Q_{PMF} from simulation done using HEC-1 package and the concept of Probable Maximum Precipitation (PMP). There are a number of techniques for estimating the PMP (WMO, 1973). For the Tendaho project the statistical technique is selected. The data, mainly the annual maximum daily rainfall, required for the PMF estimation are available for 19 stations. The statistical estimation techniques of the WMO (1973) guideline is used for the estimation of the 24 hour PMP based on annual daily rainfall of 19 stations in and nearby the modeled 13 tributaries. Results and data concerning the PMP and PMF analysis are adopted from the original design document. The computed PMF represents a flood of about T = 19 million years on the EV-I fitted line.

The uncertainty of $Q_{10,000}$ magnitude that is extrapolated for the design return period using flood frequency analysis and the uncertainty of the computed Q_{PMF} are estimated using CI analysis. The result is illustrated in Figure 7-1 and the pdf parameters are provided in Table 7.2.

In the original deterministic design of Tendaho Dam, providing a spillway hydraulic capacity sufficient to withstand the event of $Q_{10,000}$ is assumed adequate, when the $Q_{10,000}$ is taken to be equal to ½ PMF (= 3042 m^3/s). Thus, the spillway hydraulic design has been done for a routed flow corresponding to $Q_{10,000}$ = ½ PMF = 3042 m^3/s. But, parameters like crest level of control structure and width of spillway have been selected with intent of preventing overtopping even under Q_{PMF} (WWDS, 1992b). In other words, in the original design of Tendaho Dam the spillway is designed to pass the ½ PMF and the full PMF is assumed to pass through the spillway without overtopping of the dam due to the selected crest level and spillway width. That is partly the reason why it is opted to investigate both floods ($Q_{10,000}$ and Q_{PMF}) in the current study.

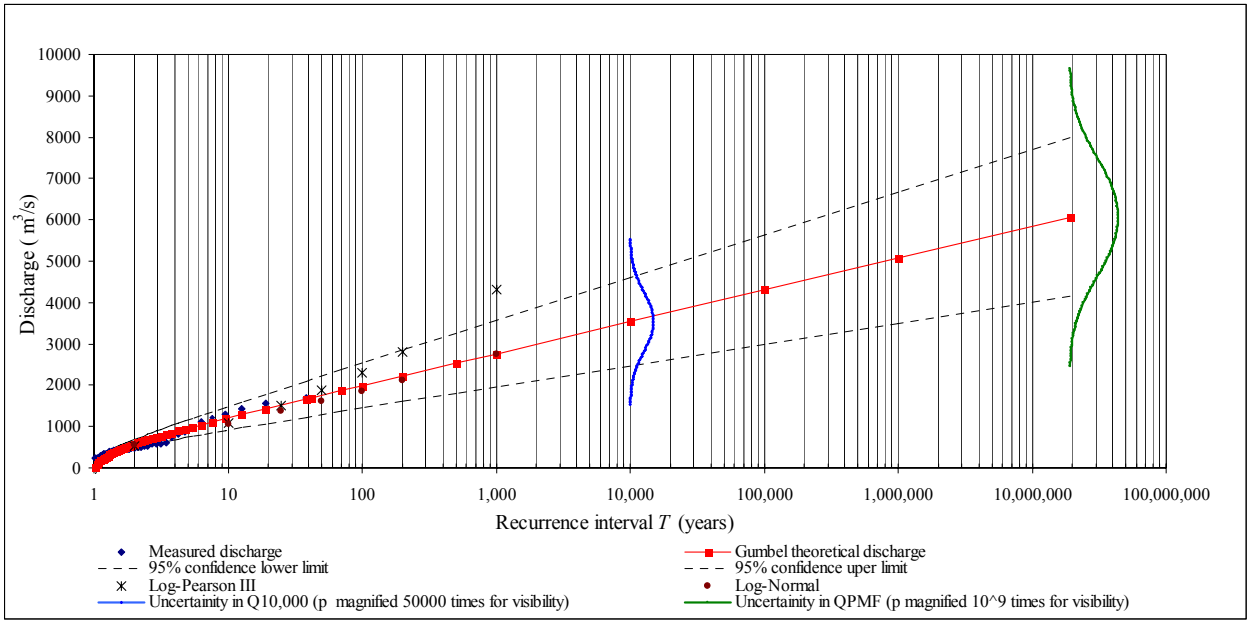


Figure 7-1. Flood frequency analysis, with $Q_{10,000}$ and Q_{PMF} pdfs.

7.3.1.2 Uncertainty in flood hydrograph (FH) shape

Besides Q_P magnitude the uncertainty in the shape of its hydrograph, which characterizes its passage with time (flow conditions), has a decisive influence on routing behaviors of reservoirs. This shape is determined by the response of catchments to rainfall inputs. It characterizes the passage of Q_P with time as a flow in three phases –surface, inter, and base. Recorded hydrographs show many kinks and multiple peaks, reflecting storm and catchment peculiarities and their complex interaction. They have shapes much complex than the simple single-peaked FHs commonly used in design. Figure 7-2 shows selected measured FHs at Tendaho. In designs such complex FH are resolved into simple deterministic analytical shapes, like the ones shown in Figure 7-3. However, such deterministic shape is not an accurate representation of reality. An analytical FH often used to simulate natural FHs is the gamma function, Eq.(7-5) (Ponce 1987).

$$Q = Q_b + (Q_P - Q_b) \cdot (t / t_P)^m \cdot e^{[(t_P - t) \cdot (t_g - t_P)]} \qquad (7\text{-}5)$$

Where Q is flow rate at time t; Q_b and Q_P are base and peak flows; t_p is time-to-peak; t_g is time-to-centroid; $m = t_p/(t_g-t_p)$. Zero time depicts the beginning of the FH. t_p is measured from $t = 0$ to the time at which Q_P is attained. t_g is measured from $t = 0$ to the t separating FH into two equal volumes.

For mid size catchments direct runoff FH can alternatively be represented using Unit Hydrographs (UH), a concept first introduced by (Sherman 1932). UH can be calculated either by: (1) directly, using rainfall-runoff data or (2) indirectly, using synthetic hydrograph formula. UH establishes a relationship between flow characteristics of unit depth of surface runoff resulting from unit depth of effective rainfall, lasting for specified duration, on a catchment. This relationship can be transformed to a relationship representing Surface Flow Hydrograph (SFH) of any design depth and duration other than unity, this topic is adequately covered in many hydrologic books, like (Chow, 1988), (Ponce, 1987), (Linsley, 1962). In overtopping analysis what is sought to be established is the total flood hydrograph (FH). The difference between FH and SFH is the Q_b ($FH = SFH + Q_b$).

Recently, use of pdfs for derivation of synthetic UH has received much attention. For example, Nadarajah (2007) explores the use of eleven most flexible pdf's for UH derivation. For each pdf he derived an expression for the unknown pdf parameters in terms of measured values of Q_P, Q_b and t_b. Bhunya et al. (2006) investigated potentials of four pdf's to the same purpose. Haktanir et al. (1999) studied applicability of two and three parameter beta distributions as synthetic UH. Klein et al. (2006) demonstrated the application of superimposed gamma distributions for stochastic generation of multi-peak hydrographs in flood design of dams. He used the simulated multi-peak UH together with design storms of defined duration and return period.

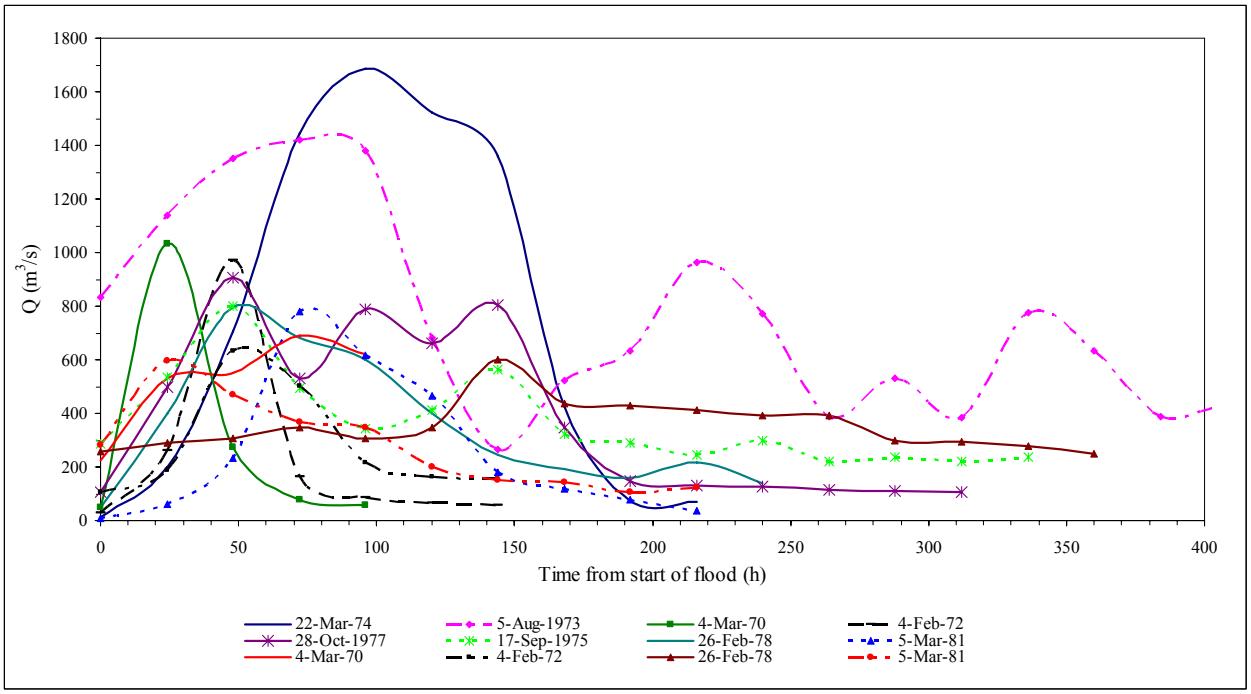


Figure 7-2. Measured FHs shape variability.

Consequently, characterization of Q_P FH shape uncertainty can basically be done either using analytical equations for describing FH, like Eq.(7-5) or using synthetic UH together with quantified Q_b and storm randomness. The equations describing FH and UH can be treated as FRV with their respective component parameters, like Q_p, Q_b, t_p and t_g, considered to be RVs. UH method is used mainly for mid sized catchments. Originally, Sherman (1932) used it for basins varying from 1300 to 8000 km^2. Linsley et al. (1962) mention an upper limit of 5000 km^2 in order to preserve accuracy. Ponce (1987) states that use of UH has been linked to midsize catchments greater than 2.5 km^2 and less than 250 km^2 but its application in larger catchments is not precluded, though overall accuracy is likely to decrease.

In this study, for estimating Q randomness at every Δt, it is opted to use FH represented by Eq.(7-5) with all its variables, not only Q_P, considered as RV. This choice is made because of the significantly large Tendaho drainage area (62088 km^2) that makes application of UH inaccurate. Eq.(7-5) component parameters (Q_b, Q_P, t_p, and t_g) randomness is determined by analyzing recorded extreme FHs at Tendaho. Out of the 35 year record 48 isolated Q_P FHs are used in the analysis. The statistics concerning the parameters of the selected FHs is given in Table 7.1 and their best fit pdfs are given in Table 7.2. An alternative approach was to fit pdfs on flow records for each routing time step Δt of the selected FHs. However, the result from this approach and from using synthetic FH represented as Eq.(7-5) is found to be reasonably similar. To see effects of taking all or part of FH variables on randomness of FH two scenarios are assessed:

a) Q_P taken as the lonely RV. Here, Q_P of $Q_{10,000}$ and Q_{PMF} are represented using pdfs (Table 2),
b) All variables of Eq.(7-5) taken as RV.

Consequences of considering either only Q_p or combinations of the other FH parameters as RV are shown in Figure 7-3.

Positively skewed shape of natural FHs shows that t_g depends on t_p. Therefore, t_g is not kept as independent random parameter. In this regard three cases are tested:

1) t_g taken as t_p + average of (t_g - t_p),
2) Fitting a pdf on the measured t_g - t_p values and taking t_g as RV,
3) Fully deterministic t_g.

Table 7.1: Statistics on measured FH variables.

	Main Rainy Season (June-September)			Short Rainy Season (February-April)			Mean
Variable	Min	Max.	Mean	Min.	Max.	Mean	
Q_b (m^3/s)	42	330	181	7	281	87	138
Q_P (m^3/s)	269	1420	579	413	1687	702	641
t_p (h)	24	144	63	24	96	46	55
t_g (h)	60	249	127	24	180	77	107
t_b	216	576	337	96	624	297	338

Table 7.2: FH random variables best fit pdfs.

Variable	PDF	PDF Parameter			Deterministic	Limitations
		Mean	Std.			
Q_P (m^3/s)	Normal	3535	645.98		3042	PMF ≥ Q_P≥0
PMF (m^3/s)	Normal	6066	1164.7		6084	PMF ≥ Q_P≥0
		Location	Scale	Shape		
Q_b (m^3/s)	Gamma	-14.39	40	3.37	100/60*	Q_P≥ Q_b≥0
t_p-t_g (h)	Gamma	-37.5	54.65	1.49	43	t_g≥ t_b≥ 0
		Min.	Mode	Max		
t_p (h)	Triangular	20.65	24	202	36/60**	t_g≥ t_b≥ 0

*Q_b for main /short rainy season used in the original deterministic design
**t_p and t_b used for the Q_{P10000}/Q_{PMF} in the original deterministic design

The Q randomness at each t is evaluated using pdfs of the variables in Eq.(7-5) and the stochastic models and solution procedure discussed below in section 7.4. Figure 7-3 shows the analysis results plotting mean $Q(t)$ and sample pdfs' of Q at t = 36, 90, 204 and 360 hours indicating randomness of FH (p values are magnified to enable their plotting on same axis with the hydrograph coordinates visibility). Figure 7-3 clearly shows that considering Q_P as the only random variable over estimates the peak flow and under estimates the flood base time. The deterministic approach has the similar effect. Considering randomness of all the variables gives a more realistic solution. The effect of this on $P_{f,\,OT}$ is discussed in results section. Flood routing and $P_{f,\,OT}$ evaluation is done using the stochastic FH represented by such $Q(t)$ pdfs.

The spreadsheet files and Mathematica codes used for making the different calculations for stochastic evaluation of dam overtopping as discussed in this chapter are given in the

appendix CD-ROM under several self-explanatory sub-folders placed under the folder (:\overtopping analysis).

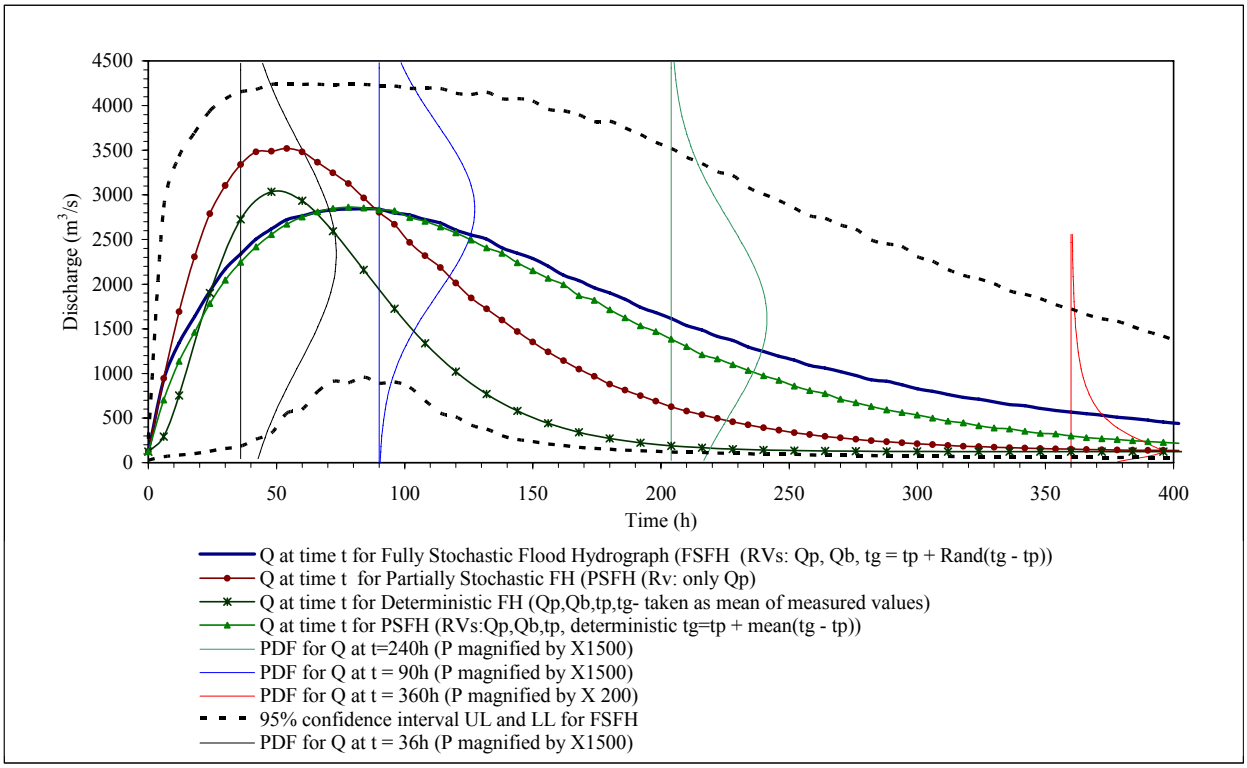


Figure 7-3. Synthetic FHs shapes variability.

7.3.1.3 Flood routing

The rise in h_o due to flood routing, i.e. h_f, is an important load parameters deciding the chance of dam overtopping (see Eq.(7-2)). A stochastic flood routing is done with intent of estimating h_f. For reservoir routing volume-elevation characteristic of reservoir, outflow-elevation characteristics of spillways and outlets and initial storage conditions should be known. For stochastic routing of Tendaho Dam a deterministic storage-elevation curve at year zero is found from studies made during its initial deterministic design (Figure 5-11). And The composite Q_{out}-H relation of the spillway and outlet structure is given as:

$$Q_{out} = Q_{outlet} + Q_{spillway} = Q_{outlet} + \frac{2}{3} \cdot C \cdot \sqrt{2 \cdot g} \cdot L \cdot H^{3/2} \qquad (7\text{-}6)$$

Where C is discharge coefficient (0.722), L spillway effective length (29.5 m), H head in m, g = 9.81 m/s^2. Q_{out} is peak irrigation outlet discharge (78 m^3/s) corresponding to peak irrigation requirement for which the outlet is designed (WWDSE, 20005b). Because of the outlet intake location, which provides at least 5 m head above MDDL, this amount is assumed to be granted during peak flood events. Routing can be done using any of established methods (iteration, Goodrich, puls, etc.) depending on the size of reservoir, the time step Δt chosen, and the accuracy required (Novak et al., 2003). All routing methods are based on continuity equation,

given as $Q_{in} - Q_{out} = dV/dt$. In finite difference form for small time Δt , $\overline{Q}_{in} - \overline{Q}_{out} = \Delta V / \Delta t$, Where $\overline{Q}_{in}$, $\overline{Q}_{out}$ and ΔV are average inflow, outflow and change in storage in time Δt. A common form of this equation used in reservoir routing is:

$$\left((2 \cdot V_2)/\Delta t\right) + Q_{out2} = Q_{in1} + Q_{in2} + \left(2 \cdot V_1/\Delta t\right) - Q_{out1} \qquad (7\text{-}7)$$

In the stochastic flood routing and $P_{f,\ OT}$ determination h_f randomness is determined following modified pulse method by using Eq.(7-7) with inflows ($Q_{in1,\ 2}$) taken as RV and represented by stochastic FH; in addition volumes ($V_{1,\ 2}$) and composite outflows ($Q_{out1,\ 2}$) are also taken as RV. $Q_{out1,\ 2}$ are computed using Eq.(7-6). $V_{1,\ 2}$ and $Q_{out1,\ 2}$ are computed for a relevant stochastic H in Δt . Δt = 6 h is selected following suggestion in (Subramanya, 1988). Mathematica5.2. codes are written for doing the necessary computations. The solution implementation architecture is given in Figure 7-7.

7.3.2 Initial water level (h_o)

The chance of overtopping due to flood inflow and wave events obviously depends on the available storage volume at the beginning of the flood event (h_d - h_o). h_o is a RV whose value depends on previous flood or draught, abstractions, evaporation, environmental releases and sedimentation. Two scenarios of h_o are considered and results compared:

(1) Assuming Q_P arrives when the dam is at FRL (water level in the reservoir is static at 408 m). This is a common assumption in dam design.

(2) Assuming h_o is RV and Q_P could inter at any level with some probability (p).

To determine h_o randomness long-term synthetic series of monthly flow (50 yrs) is generated using Thomas-Fiering synthetic monthly flow generation model (Thomas and Fiering, 1962). Equivalently similar models for generating monthly flow sequences can be used, e.g. (Şen, 1978), Thomas-Fiering model is selected simply because it was the one used during the original design of the dam. Then monthly water balance model is constructed to estimate water levels after meeting irrigation demand (60,000 ha), environmental release (5 m^3/s) and evaporation losses (5.8 to 9.4 mm/day) using:

$$V_j = V_{j-1} - Q_j + P_j - E_j - ER_j - IR_j - SP_j \qquad (7\text{-}8)$$

Where: V_j and $V_{j\text{-}1}$ are reservoir storage at end of time period j and j - 1; Q_j, P_j, E_j, R_j, IR_j, SP_j are inflow, rainfall, evaporation over/from reservoir, environmental releases, irrigation release, spill in time period j respectively, all in volumes. The length of time period adopted is j_{n+1} - j_n = 10-days. The cdf that defines h_o randomness is then determined from data fitting done on the maximum water levels found from this water balance model (Figure 7-4).

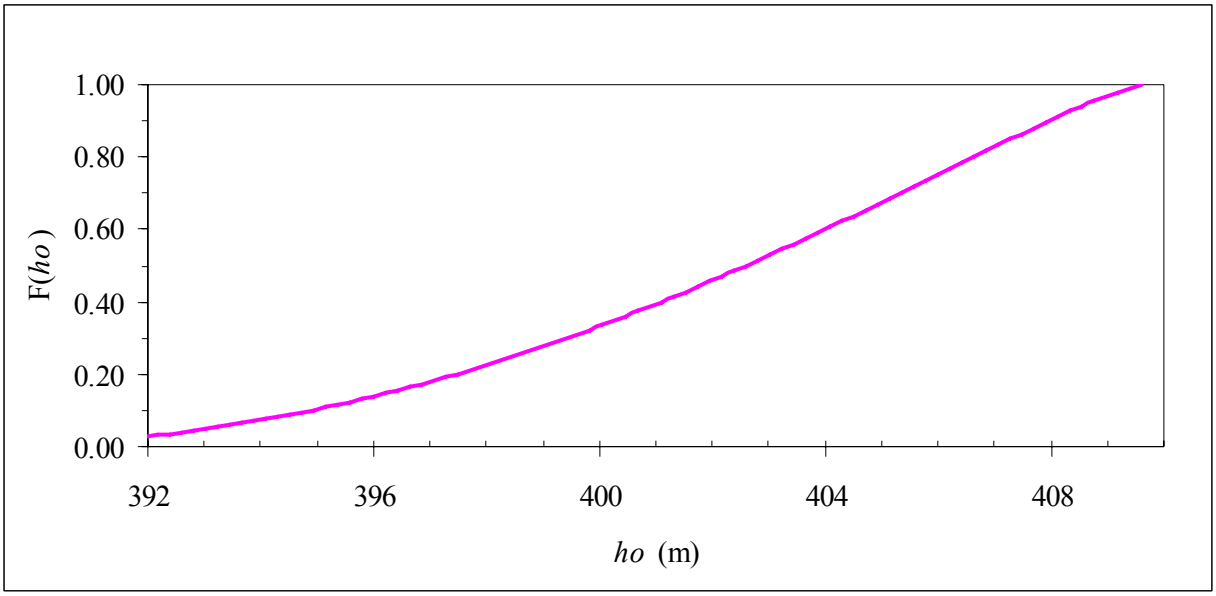


Figure 7-4. cdf (h_0) *cdf ho: Beta distribution (max. = 409.62, min = 388.29, α = 2.22, β = 1.24).*

7.3.3 Wave height-wave run up and wind setup (h_r, h_s)

To guard against overtopping, apart from the flood surcharge (h_f), adequate freeboard has to be provided to accommodate waves. The total wave surcharge is computed as a sum of the height of wave run up (or design wave height whichever is greater) h_r and the wind setup h_s.

7.3.3.1 Wave height and run-up (h_r)

Historical background of h_r computation and details of the commonly used h_r formulas is given in (Senturk, 1994). Different h_r formula can give different results. For a deterministic mind set these variations seem significant and at times they are confusing. But, when the natural wind speed (U) uncertainty is accounted for, results from the different methods are found to be well bound by fairly similar pdf. For the computations h_r the British Institute of Civil Engineers (ICE) standards and procedures are adopted with the only exception of taking wind speed (U) as a random RV. The method uses a standard calculation of the significant wave height (H_s) and modifies this to obtain a wave surcharge allowance that is appropriate to extreme conditions comparable to those during design flood. ICE (1996) can be referred for complete discussion on formulas and procedures. However, the basic steps and equations used in the analysis are provided below.

7.3.3.2 Wind speed and direction

The ICE standard relies on the 50-year maximum hourly wind speed (U_{50}, in m/s) for its wave height-wave run-up and wind setup calculations. This, U_{50} in m/s has to be adjusted to mean annual maximum hourly wind speed at sea level as required in the standards (see Table 5.5). A serious of adjustment factors are used to get the appropriate mean annual maximum hourly wind speed.

In order to satisfy the criteria set out in Table 5.5 first an adjustment factor (f_T) is required to provide the estimate of the mean annual maximum hourly wind speed. Factor for adjusting the 50-year maximum hourly wind speed to other return periods and the mean maximum hourly values are given in Table 7.3.

Table 7.3: Wind speed ratios (f_T) for selected recurrence interval (ICE 1996).

Return period (years)	Wind speed ratio relative to 50-year (f_T)
Mean annual	0.79
5	0.85
10	0.90
20	0.95
50	1.00
100	1.04
200	1.08

The required adjustment for altitude (f_A) to get the mean annual maximum hourly wind speed at sea level is obtained from the equation:

$$f_A = 1.0 + (0.001 \cdot alt) \qquad (7\text{-}9)$$

Where '*alt*' is the reservoir altitude in meters above sea level.

A further adjustment is also required which reflects the increase in wind speed over open water as opposed to over land. This adjustment (f_{ew}) is obtained from Table 7.4, as a function of fetch. Senturk (1994) suggests a similar approach. For Tendaho dam data on wind speed over open water was not available. Hence, the available over land *U* record (1969 - 2000) is increased by multiplying it with a coefficient $f_{ew} = 1.26$ to reflect the *U* over water. The effective fetch F_e for Tendaho reservoir is 5.97 km and its maximum fetch F_{max} = 38.7 km.

To develop a fully arisen reservoir wind must blow for a certain duration; the shorter the F_e and higher the *U* the shorter the time. Consequently, it might be necessary to convert the wind speed to a more appropriate duration for reservoirs. Novak (2003) gives typical values of 1h for F_e of 3 km, 3 h for F_e = 20 km. Martin and Pohl (1998) give the minimum required duration (in minutes) to develop a fully arisen sea as 10 times the F_e, where F_e is in km; i.e. for every kilometer of the effective fetch length at list 10 minutes duration wind blow is necessary. Based on this Martin and Pohl (1998) give factors for converting one-hour-mean wind velocity, which is appropriate to only fetch length of about 6 km, into shorter duration or impact times when the fetch is less than 6 km (Table 7.6). For Tendaho with *Fe* = 5.97 km, thus wind speed *U* of 1h duration is appropriate. Therefore, no additional duration adjustment factor (f_D) is required.

Table 7.4: Ratio between wind speeds over water and over land (f_w) (ICE 1996).

Fetch(m)	1000	2000	4000	8000	12,000
Wind speed ration (f_w) (over water/over land)	1.10	1.16	1.23	1.29	1.31

Table 7.5: Factor for converting one-hour-mean wind speed to shorter duration (ICE 1996).

Fetch(km)	Minimum duration (minutes)	Factor for converting one hour mean to other duration
6	60	1.0
2	20	1.05
1	10	1.1
0.5	5	1.2

A final wind direction adjustment factor (f_N) may be considered to be appropriate. This allows for the orientation of the principal axis of the reservoir with respect to wind direction. The values of this direction factor (f_N) are given in (ICE 1996). For Tendaho in the original design a conservative value of 1.0 is assumed. However, seeing the predominant wind direction in the main flood season that is south to south westerly wind an f_N value of about 0.73 might be appropriate.

$$U = f_T \cdot f_A \cdot f_W \cdot f_D \cdot f_N \cdot U_{50} \qquad (7\text{-}10)$$

Frequency and confidence interval (CI) analysis is conducted on the overland U data adjusted over water speed as discussed above. The T = 50 years U is found to be distributed normal [μ = 90.95 km/h, σ = 16.08 km/h] (Figure 7-5). This cdf is used for computing h_r and h_s distribution. For Tendaho dam the product of the adjustment factors is $f_T \cdot f_A \cdot f_W \cdot f_D \cdot f_N = 0.79 \cdot 1.4 \cdot 1.26 \cdot 1.0 \cdot 0.73 \approx 1.0$. In the original design 50 years maximum hourly over water U of 152 km/h has been adopted.

7.3.3.3 Significant wave height (H_s), design wave height and wave run-up (h_r)

The significant wave height (H_s) is the mean height of the highest third of all waves. About 14% of the waves are higher than H_s. ICE 1996 recommends the use of Donelan/JONSWAP method for the prediction of significant wave height (H_s) given as:

$$H_s = \frac{U \cdot F_e^{0.5}}{1760} \qquad (7\text{-}11)$$

Where U is the adjusted wind speed in meter per second and F_e is the effective fetch in meter. The effective fetch is defined as:

$$F_e = \frac{\sum x_i \cdot \cos\theta}{\sum \cos\theta} \qquad (7\text{-}12)$$

To use Eq.(7-12) the reservoir is divided in to sectors of equal angle both to right and left of the maximum fetch length (usually also the line of wind direction), for example at an increment of 10^0. Then angle θ is measured from the maximum fetch line to end of each sector. And x_i is the fetch length of each sector.

7.3.3.4 Design wave height and wave run-up (h_r)

The Donelan/JONSWAP method will produce a significant wave height for extreme conditions on the reservoir. It is then necessary to modify this height to give a design wave height. Three main factors are involved, i.e. influence of structure and land form near the dam, tolerance of the dam to overtopping and wave carry-over and wave run-up on the upstream of the dam.

The predicted significant wave height (H_s) is exceeded by about 14% of waves, some 6% exceed $1.2H_s$ and the maximum wave height may approach $1.67H_s$. Table 7.6 shows factors to be applied to H_s in order to estimate the design wave height (h_d). In case of Tendaho dam no Stillwater or wave surcharge carryover is permitted. As a result a design wave height $h_d = 1.67 \cdot H_s$ is taken to be appropriate.

Table 7.6: Design wave height (ICE 1996).

Dam type	Top of dam	Design wave height (h_d)	Percentage of waves above h_d
concrete/masonry	-	$0.75H_s$	33
Rockfill	surface road	$1.0H_s$	14
Earthfill with reinforced downstream face	surface road	$1.1H_s$	9
Earthfill with random grass downstream face	surface road	$1.2H_s$	6
Earthfill with random grass downstream face	Grass crest	$1.3H_s$	4
All dam types- no still water or wave surcharge carryover permitted		$1.67H_s$	

Waves will run-up the upstream slope of a dam to different heights depending on its inclination, smoothness and permeability. In (ICE 1996) wave run-up factor (R_f), which is the ratio of wave run-up height to design wave height (h_d) is provided graphically as a function of the upstream dam slope inclination and the nature of the face (permeability and roughness). For Tendaho dam considering its upstream slope inclination of 1:3.5 that is covered with damped rock riprap with high permeability and roughness, an R_f value of 0.6 is considered appropriate.

In the final computation the highest of the design wave height (h_d) or the wave run-up ($R_f \cdot h_d$) will be taken as the appropriate design wave height-wave run-up (h_r) value.

7.3.3.5 Wind set up (h_s)

Wind setup (h_s) can be calculated using the Zuider Zee formula (Novak, 2003):

$$h_s = (U^2 \cdot Fe_{\max} \cdot \cos\alpha)/(62000 \cdot H) \qquad (7\text{-}13)$$

Where, H is average reservoir depth (m), U (km/h) and Fe_{max} maximum fetch. Notionally there are two RVs U and H. However, the influence of taking H randomness is found to have

little effect on the computed total wave surcharge (h_r + h_s) (see Figure 7-6). Therefore, the computation of total wave surcharge (h_r + h_s) ultimately took only U as a RV.

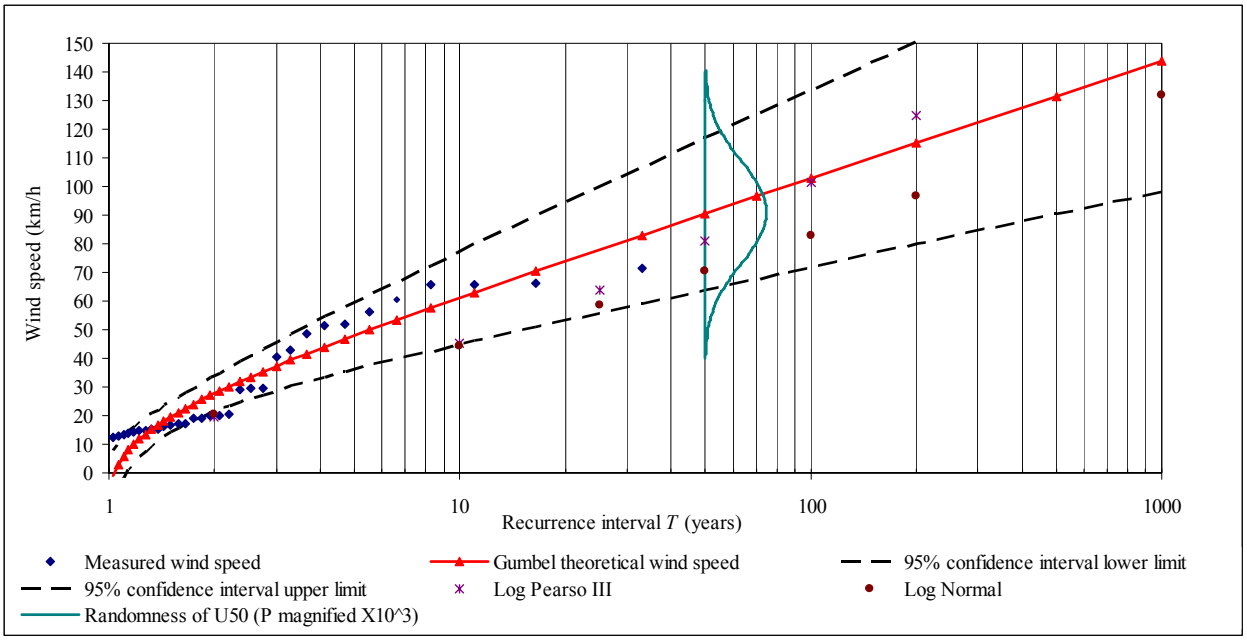


Figure 7-5. U_{50} randomness.

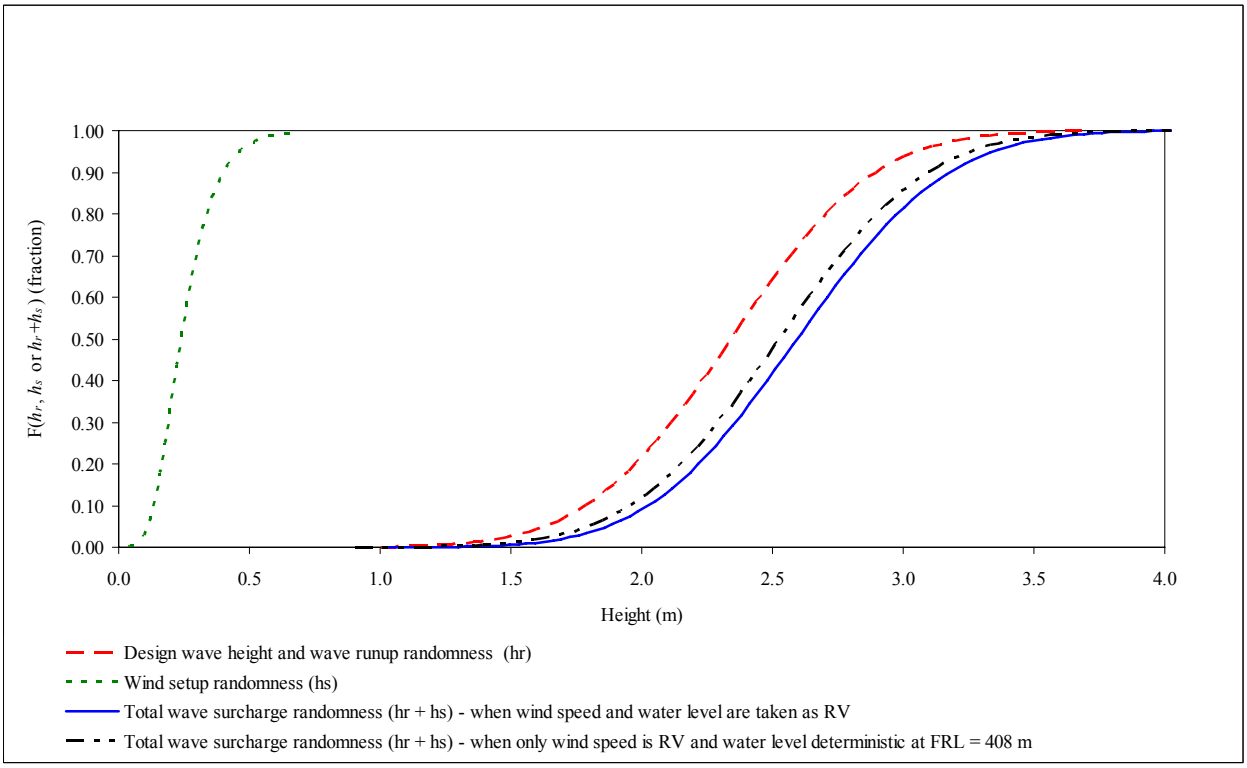


Figure 7-6. cdf for *h_r normal (μ=2.34, σ=0.44), h_s max. extreme (likeliest = 0.21, scale = 0.09), $h_r + h_s$ normal (μ= 2.6, μ=0.45).*

7.4 Stochastic models and solution procedure

A complete evaluation of Tendaho Dam overtopping reliability (Z) (Eq.(7-1)) and consequently its $P_{f,\,OT}$ and $R_{f,\,OT}$ (Eq.(7-2) and (7-3)) is done for two design discharges ($Q_{10,000}$ and Q_{PMF}) by using the four stochastic methods-MCSM, the two approximate moment analysis methods (FOSM and SOSM methods) and the new to the purpose Analytical Solution for Determining Derived Distributions (ASDD) method. The FRV used for determining cdf of Z and $P_{f,\,OT}$ is Eq.(7-1), which has four RVs (h_o, h_f, h_s, h_r) that need to be replaced with corresponding pdfs or moments. The pdfs for these four RVs is determined considering the multitude of approaches for estimating underling random variables discussed earlier in this chapter. In addition, the computation of the pdfs of the four random variables involves applications of the four stochastic methods at several levels as presented in Figure 7-7 and discussed below. In the discussion of the implementation architecture below, whenever a FRV one or more variables is defined, that means the calculation using the specific FRV is made using all the four stochastic methods (MCSM, FOSM, SOSM, and ASDD) even though the discussion uses one or the other method in the explanation.

The pdf of Q_{in} at every Δt is calculated using Eq.(7-5) which has four RVs (Q_b, Q_P, t_p, and t_g).The calculation of h_f pdf at the end of every time step (Δt) uses Eq.(7-7), which itself is FRV of four RVs ($Q_{in1,\,2}$, V_1, $Q_{out\,1}$) and an additional one RV h_o in the first Δt . For fixing pdfs of h_f and Q_{in} at end of every Δt the same four stochastic methods are used. V_1, Q_{out1} are functions of computed h_f in every Δt and are used in computing pdf of $\left(2 \cdot V_1 / \Delta t - Q_{out\,1}\right)$ that will later be used in the next Δt computations. Results from the multitude of approaches employed and the different stochastic methods are compared and explained. Figure 7-7 gives the solution implementation architecture in more or less self explanatory manner. A brief description of the architecture and stochastic models is given below. Additional references on the probabilistic methods can be found in chapters 3 and 4 (also in Negede and Horlacher, 2008a) and in (Horlacher and Negede, 2008). And, detailed demonstration of the methods application for FRVs with different number of variables is given in chapter 6 and appendix 9.4. Therefore, repeating the demonstration here is avoided. The description of the architecture that contains five layers is:

1) The intention is to determine the prevailing $P_{f,\,OT}$ of dams. This is possible through calculating or estimating uncertainty of reliability provided against overtopping, i.e. through computation of pdf of Z using Eq.(7-1). This pdf define randomness in outputs of Eq.(7-1). Z is a function of four load random variables h_o, h_f, h_s and h_r. Hence, determining Z's pdf, and thus $P_{f,\,OT}$, involves computation with pdf's of these four load random variables whose pdf is yet unknown. The uncertainty of these load random variables is a result of propagated inherent randomness in underlying physical variables. The load parameters are not primary random variables exhibiting an inherent randomness. Rather their randomness comes from underlying variables uncertainty. Thus, it is not possible to get a data set through measurement, which could enable statistical analysis, on these four secondary load parameters prior the operation of the dam. However, uncertainty of these load parameters can be estimated by propagating uncertainty in underlining primary load parameters. These primary load parameters are U, Q_p, Q_b, and FH variables

(t_p and t_g). The uncertainty of these primary load parameters can be modeled with relative ease using pdfs derived from measured data sets.

Consequently, a logical approach to evaluate randomness (pdf) of the secondary load random variables of Eq.(7-1) (h_o, h_f, h_s, h_r) shall commence from quantifying uncertainty in individual primary load random variables U, Q_p, Q_b, t_p and t_g. And then the uncertainty in the underlying primary load random variables can be propagated through the functional relations (design equations) for estimating h_o, h_f, h_s, and hr. Here, it has to be noted that h_o is an exception. It is a result of a water balance interaction of a given dam. Therefore, at a design stage data can only be found through water balance modeling. Consequently, underlying uncertainty in variables influencing h_o are propagated in a water balance model constructed on synthetic data unlike measured data as in the case of the other secondary load variables.

Accordingly, the first layer of the implementation architecture presents the primarily measured or modeled raw data series/set, i.e. Q_{out}, Q_{in}, FH (t_p, t_g , and Q_b), Q_p and U. The Q_{in} series comes from synthetic monthly flow generation models. Flood hydrograph shape (FH) parameter t_p, t_g, Q_b data set is found from measured peak flow hydrographs. The FH for peak flows can be constructed from long term daily flow series and the corresponding variables (t_p, t_g, Q_b) can be read from the FH data set. Q_p and U series are found from hydrological and meteorological stations, respectively.

The relevant data, analysis and intermediate results of this layer are given in the appendix CD-ROM in the file (:\flood analysis\flood-wind initialWL volume elevation sediment frequency analysis).

Once the data series/sets are at hand the next step is to quantify their respective randomness using pdfs or using their moments (mean and standard deviation). This step is described in 2 below and is represented by layer 2 in the architecture.

2) Layer 2 of the architecture deals with construction of pdf for the primary load parameters (t_p, t_g, Q_b, h_o, Q_p, and U). The pdf of t_p, t_g, Q_b, h_o is constructed through distribution fitting. The data for t_p, t_g, Q_b, h_o, and Q_p comes from historic records and the data for h_o comes from water balance modeling done using synthetically generated flows for 50 years (see section 7.3.2). Topics of distribution fitting to data and uncertainty modeling using distributions are discussed in appendix 9.1.3 and 9.1.1.5, respectively. The pdfs of Q_p and U are constructed by first doing frequency analysis to estimate the central value of the design Q_p and U. Then confidence interval analysis is done to fix the respective pdfs for the design Q_p and U. This step is discussed in section 7.3.1.1. Therefore, the output of this step is pdf or moments for the underling load variables t_p, t_g, Q_b, Q_p, U and h_o. For applications in MCSM and ASDD formulation pdf construction is mandatory. But, for application in FOSM and SOSM formulation computing only moments of the data series/set using discrete statistics is adequate.

The relevant data, analysis and intermediate results of this layer are given in the appendix CD-ROM in the file (:\flood analysis\flood-wind initialWL volume elevation sediment frequency analysis\Tendaho flood frequency analysis - v6.xls) and under the folder (:\flood analysis\stochastic deterministic hydrographs) in several files with self-explanatory names.

3) Among parameters determining value of h_f, which is one of the random load parameter in Eq.(7-1), are inflow flood hydrograph (FH), reservoir stage discharge relation, and spillway hydraulics-operation. Therefore, uncertainty in the underlying FH shape parameters (t_p, t_g, Q_b, Q_p), which is evaluated in step 2 above, has to be propagated through Eq.(7-5). This equation defines the interaction between these variables in defining flood flow pattern (flood hydrograph). Propagating uncertainty of the FH shape parameters in Eq.(7-5) in turn allows estimating uncertainty in the design flood hydrograph. Thus, using the underlying variables uncertainty (t_p, t_g, Q_b, Q_p) a stochastic flood hydrograph (SFH) is constructed. This is the task in layer 3 of the architecture. For propagating uncertainty in underling random variables and determining SFH, Eq. (7-5) is considered as a function of four random variables, i.e. the flow magnitude at any time t is $Q_t = Q(\boldsymbol{X})$, where $\boldsymbol{X} = (t_p, t_g, Q_b, Q_p)$. Thus, the randomness in Q_t is given by a Q_t pdf or Q_t moment computed by taking Eq.(7-5) as a FRV of four random variables and by applying the stochastic models discussed in sections 3.1 (MCSM), section 3.2.4 (FOSM and SOSM) and section 4.3.2 (ASDD for multivariate problem).

For the MCSM estimation a simulation is run 10,000 times by taking random samples of the four random variables (t_p, t_g, Q_b, Q_p) from their respective pdfs determined in step 2 above. The computational framework for MCSM is developed on spreadsheet and computer program Crystal Ball 7.2 is used to do the actual simulation. The SOSM and FOSM approximations are found by solving the following equations after replacing the appropriate variables and moments (estimated in step 2) in the general equations for SOSM and FOSM approximation of means and variances, i.e. Eqs. (3-14), (3-22), and (3-23), (3-21).

The expression for the SOSM approximation for the mean and variance of Q_t takes the form:

$$\mu Q_{t,SOSM} = Q(\mu_{tP}, \mu_{tg}, \mu_{Qp}, \mu_{Qb'}) + \frac{1}{2}\cdot\left(\nu_{tp}\cdot\left[\frac{\partial^2 Q}{\partial t_p^{\,2}}\right]_{\mu} + \nu_{tg}\cdot\left[\frac{\partial^2 Q}{\partial t_g^{\,2}}\right]_{\mu_\mu}\right)$$

$$+\nu_{Qp}\cdot\left[\frac{\partial^2 Q}{\partial Q_p^{\,2}}\right]_{\mu} + \nu_{Q_b}\cdot\left[\frac{\partial^2 Q}{\partial Q_b^{\,2}}\right] \qquad (7\text{-}14)$$

$$\nu Q_{t,SOSM} = \left(\nu_{tp} \cdot \left[\frac{\partial Q}{\partial t_P} \right]_\mu^2 + \nu_{tg} \cdot \left[\frac{\partial Q}{\partial t_g} \right]_\mu^2 + \nu_{Q_P} \cdot \left[\frac{\partial Q}{\partial Q_P} \right]_\mu^2 + \nu Q_b \cdot \left[\frac{\partial Q}{\partial Q_b} \right]_\mu^2 \right)$$

$$-\frac{1}{4} \cdot \left(\nu_{tP} \cdot \left[\frac{\partial^2 Q}{\partial {t_P}^2} \right]_\mu + \nu_{tg} \cdot \left[\frac{\partial^2 Q}{\partial {t_g}^2} \right]_\mu + \nu_{QP} \cdot \left[\frac{\partial^2 Q}{\partial {Q_P}^2} \right]_\mu + \nu_{Qb} \cdot \left[\frac{\partial^2 Q}{\partial {Q_b}^2} \right]_\mu \right)^2 \quad (7\text{-}15)$$

Similarly, the expression for the FOSM approximation for the mean and variance of Q_t takes the form:

$$\mu Q_{t,FOSM} = Q(\mu_{tP}, \mu_{tg}, \mu_{Qp}, \mu_{Qb}) \quad (7\text{-}16)$$

$$\nu Q_{t,FOSM} = \left(\nu_{tp} \cdot \left[\frac{\partial Q}{\partial t_P} \right]_\mu^2 + \nu_{tg} \cdot \left[\frac{\partial Q}{\partial t_g} \right]_\mu^2 + \nu_{Q_P} \cdot \left[\frac{\partial Q}{\partial Q_P} \right]_\mu^2 + \nu Q_b \cdot \left[\frac{\partial Q}{\partial Q_b} \right]_\mu^2 \right) \quad (7\text{-}17)$$

$[\ldots]_\mu$ means evaluate partial derivatives at mean of X_i. $[\ldots]_\mu^2$ means evaluate the partial derivative of τ at mean of X_i.and square it.

Similarly, MCSM and ASDD method are applied considering Eq.(7-5) as a function of four random variables $Q_t = Q(\boldsymbol{X})$, where $\boldsymbol{X} = (t_p, t_g, Q_b, Q_p)$. Application of the stochastic methods for multivariate FRV is adequately described and demonstrated in chapter 6 and appendix 9.4. Therefore, repeating the demonstration here is avoided. However, the computer program written to do the computations in Mathematica 5.2 environment is given in the CD-ROM attached. The relevant data, analysis and intermediate results of this layer are given in the appendix CD-ROM in the folders (:\flood analysis\stochastic deterministic hydrographs), (:\flood analysis\ASDD method), (:\flood analysis FOSM-SOSM methods), and (:\flood analysis\ MCS method) under several files with self-explanatory file names.

The relevant Q_t values can be determined based on the routing interval selected. In the case of Tendaho dam $\Delta t = 6h$ is used for routing. As a result the randomness in Q for every 6 hour interval along the maximum flood base time is computed in the case of Tendaho dam.

The out put of this layer is a stochastic flood hydrograph (SFH), which is actually pdfs or moments giving the possible values of flood flow rate Q_t at regular interval of $\Delta t = 6h$. For examples, as shown in Figure 7-3 for t = 240, 90 and 360 h. These computed pdfs or moments of Q_t will be used in the stochastic reservoir routing scheme run for the determination of h_f randomness in layer 4 of the architecture.

4) In layer 4a stochastic flood routing is done with intent of estimating h_f randomness. For the stochastic reservoir routing using Eq.(7-7) pulse method is employed. Understanding of the computation scheme will be facilitated if the problem is put in the classical tabular form as shown in Table 7.7. This table gives part of instant values of Monte Carlo Simulation run just for describing the relation between variables and the process. Otherwise, the ASDD, FOSM, SOSM based stochastic routing is done in Mathematica 5.2 environment. The Mathematica codes for doing the stochastic routing based on the different methods are given in the appendix CD-ROM in the folders (:\flood analysis\ASDD method), (:\flood analysis FOSM-SOSM methods), and (:\flood analysis\ MCS method) under several files with self-explanatory file names.

Equations, relevant random variables and solution procedures for the stochastic routing done in layer 4 are explained as follows. The inflow Q_{in} pdf is known for every time step. It is computed in step 3 above. The initial water level h_o pdf is also known from step 2 above. From dam storage-elevation-area relationship a regression equation can be fitted to give an explicit relation between water level and gross storage volume (V).

At t = 0 the pdf of h_o is known. It is determined in step 2 above. Thus, at t = 0 the known pdf of h_o can be transformed to pdf of storage volume using the storage-elevation-area equation. This is possible by considering the storage-elevation-area curve as uni-variate function of random variable (n = *1*) and applying Eq.(4-9) (the only RV being h_o). The application of is demonstrated in section 6.3.1. Similarly, FOSM and SOSM methods can be applied using moments of h_o and Eq.(4-9) as outlined in Eqs. (3-14), (3-22), and (3-23), (3-21). The demonstrations of the application of these equations are given in appendix 9.4. From this computations, therefore, at t = 0 pdf of storage volume (V) is also known.

Similarly, at t = 0, h_o pdf can be transformed to Q_{out} pdf by considering Eq.(7-6) as uni-variate FRV and applying the same set of stochastic methods. The only random variable being water level at t = 0 (h_o). Also, the pdf of flood inflow Q_{in} at t = 0 time interval is known. Thus, the pdf of the expression $\left(2 \cdot V_1 / \Delta t - Q_{out1}\right)$, which we call it here as *S_ind.* (for storage indication), can be calculated by taking the expression as FRV of two random variables $\boldsymbol{X}(V_1, Q_{out1})$. The pdf of *S_ind.* will later be used in the next Δt computation.

Therefore, at t = 0 the pdf for all variables in the right hand side of the routing equation Eq.(7-7) are known. This means for this Δt the expression on the left hand side $\left((2 \cdot V_2)/\Delta t + Q_{out}\right)_2$, we call it here *R_ind.*, can be given as a FRV of three random variables (n = *3*) , mathematically $\left((2 \cdot V_2)/\Delta t + Q_{out}\right)_2$ = $\boldsymbol{X}(V, Q_{in1+}\ Q_{out2,}\ S\_ind.)$. Consequently, the pdfs of these three random variables $\boldsymbol{X}(V, Q_{in1+}\ Q_{out2,}$ S_ind.) can be analytically manipulated using ASDD method as per the guideline given in Eq.(4-14) and the pdf of *R_ind.* can also be computed. Or when FOSM and SOSM methods are used moments of the three random

variables can be used to get the moments of *R_ind.* Thus, at this level the pdf or moments of all expressions in the routing Eq. (7-7) are known at routing interval $t = 0$.

Once the *R_ind*. pdf is known it can be converted to water level (flood surcharge, h_f) pdf because there is a one two one relation between water level and the expression $\left((2 \cdot V_2)/\Delta t + Q_{out}\right)_2$. The *S_ind.* pdf computed at $t = 0$ time interval will be used in the next time step. This same process will be repeated until we reach the last routing time step. The same approach is used in classical routing methods except with the difference of doing computations using pdfs and moments applied here.

In layer 4b the wave height-run-up (h_r) pdf is computed by considering the significant wave height equation Eq.(7-11) as uni-variate FRV. The only random variable being wind speed (U). In this same layer 4b wave set-up (h_s) pdf is computed by considering Eq.(7-13) as FRV of one random variables. The random variables being wind speed (U). Notionally there are two RVs U and H. However, the influence of taking H randomness is found to have little effect on the computed total wave surcharge ($h_r + h_s$) (see also section 7.3.3.5).

Therefore, at the end of layer 4, pdfs of all variables of the reliability function Z, Eq.(7-1), are known. And, this pdf of Z can be computed by taking Eq.(7-1) as an FRV of four random variables (h_o, h_f, h_s, h_r), whose pdfs are known from the preceding steps.

5) In layer 5 takes Eq.(7-1) as a FRV of four random variables and determines the pdf of Z. Once the pdf or cdf of Z is known, then the $P_{f,\,OT}$ can be read from it for selected design dam height (h_d).

Table 7.7: Modified pulse reservoir routing.

T (h)	Q_{in} (m^3/s)	$Q_{in1} + Q_{in2}$ (m^3/s)	Q_{out} (m^3/s)	$(2 \cdot V_1 / \Delta t - Q_{out1})_1$ (m^3/s)	$((2 \cdot V_2) / \Delta t + + Q_{out})_2$ (m^3/s)	Y_n, (m) absolute	$Y_n + h_r + h_s$	Y_n (m) relative to spillway crest
0	40	40	1423	170784	170825	407.81	410.33	7.81
6	322	362	1378	168390	171146	407.83	410.35	7.83
12	967	1289	1351	166977	169679	407.73	410.25	7.73
18	1683	2651	1350	166927	169627	407.72	410.24	7.72
24	2314	3997	1374	168176	170924	407.81	410.33	7.81
30	2790	5104	1417	170447	173280	407.98	410.49	7.98
36	3095	5885	1473	173387	176332	408.18	410.70	8.18
42	3240	6335	1535	176652	179722	408.41	410.93	8.41
48	3252	6492	1598	179947	183144	408.64	411.16	8.64
54	3160	6412	1658	183043	186360	408.86	411.38	8.86
60	2994	6154	1711	185775	189197	409.05	411.56	9.05
66	2778	5772	1751	188045	191546	409.19	411.70	9.19
72	2535	5313	1780	189798	193357	409.29	411.81	9.29
78	2280	4815	1799	191014	194613	409.36	411.88	9.36
84	2027	4307	1811	191699	195321	409.40	411.91	9.40
90	1783	3810	1814	191881	195509	409.41	411.93	9.41
96	1555	3338	1809	191602	195220	409.39	411.91	9.39
102	1346	2901	1798	190907	194503	409.35	411.87	9.35
108	1157	2503	1780	189850	193411	409.29	411.81	9.29
114	990	2147	1758	188480	191996	409.21	411.73	9.21
120	842	1831	1730	186851	190312	409.11	411.63	9.11
126	713	1555	1696	185013	188406	408.99	411.51	8.99
132	603	1316	1658	183014	186329	408.86	411.38	8.86
138	507	1110	1617	180891	184124	408.71	411.23	8.71
144	427	934	1574	178677	181825	408.56	411.07	8.56
.								
.								
.								

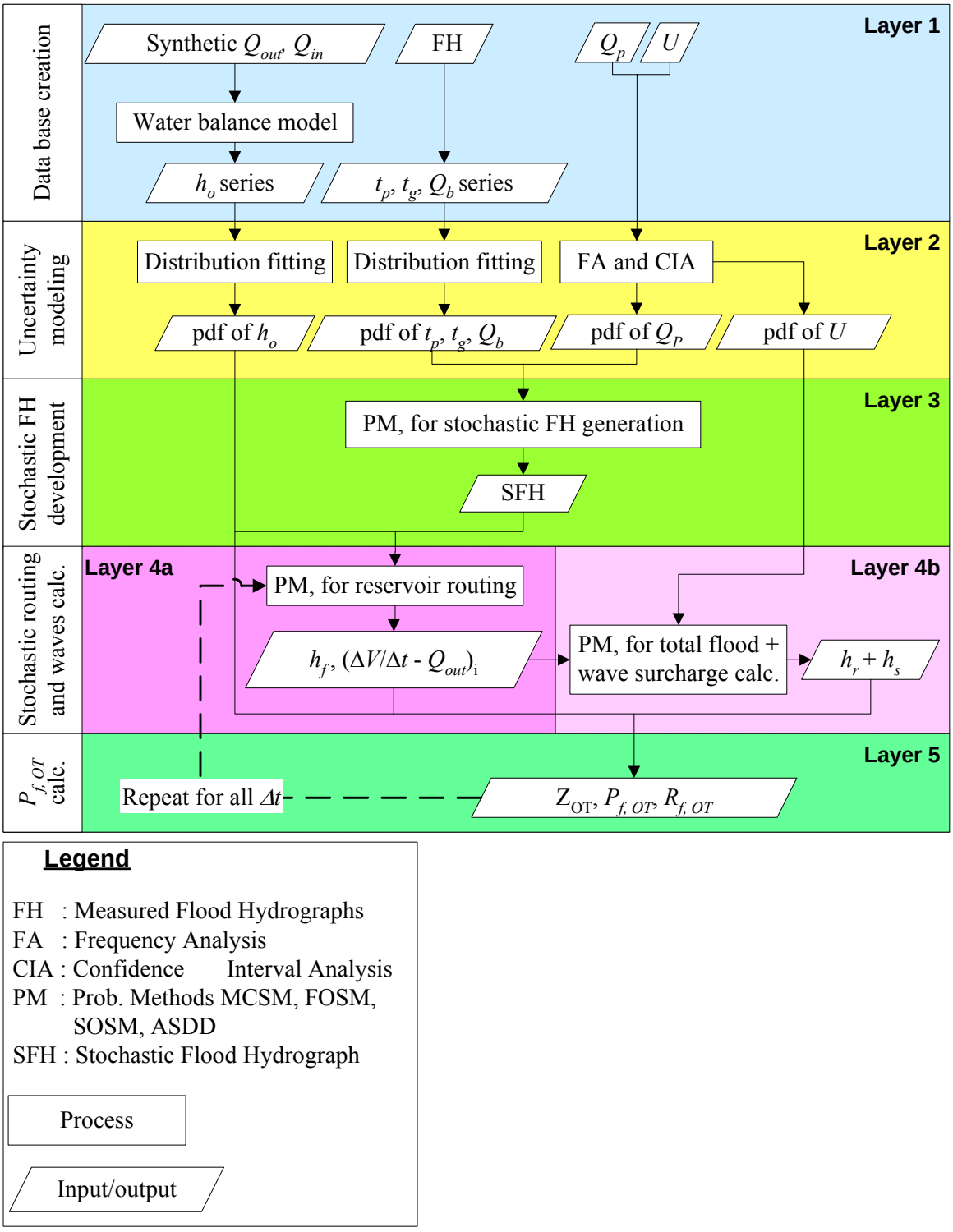


Figure 7-7. Implementation architecture.

7.5 Results and discussions

Summary of the results from all stochastic models, for $Q_P = Q_{10,000}$ case, and the result from MCSM for $Q_P = Q_{PMF}$ case are given in Table 7.8.

In the original deterministic design of Tendaho Dam, providing a spillway hydraulic capacity sufficient to withstand the event of $Q_{10,000}$ is assumed adequate, when the $Q_{10,000}$ is taken to be equal to ½ PMF (= 3042 m^3/s). Thus, the spillway hydraulic design has been done for a routed flow corresponding to $Q_{10,000}$. But, parameters like crest level of control structure and width of spillway have been selected with intent of preventing overtopping even under Q_{PMF} (WWDS 1992b). However, from the current study it is apparent that while the dam is safe to slightly over designed for a $Q_{10,000}$ events, it nevertheless has serious capacity short coming in an event of Q_{PMF}. Figure 7-8 provides plots of cdfs for reservoir water level surcharge due to combined effects of $Q_{10,000}$ flood, wave and initial water level. $P_{f, OT}$ can be read from Figure 7-8 for any selected dam crest level (adopted crest height at Tendaho is 412.5 m a.m.s.l).

In Figure 7-9 and Figure 7-10 consequence of different design considerations are presented in a self explanatory way. The design considerations evaluated are: (1) effect of using design flood of either $Q_{10,000}$ or Q_{PMF}, (2) significance of considering concurrent wave surcharges beside floods, and (3) effects of using different combinations of variables of FH as RV. In Figure 7-9 and Figure 7-10, due to space limitations, plots are provided only for MCSM results.

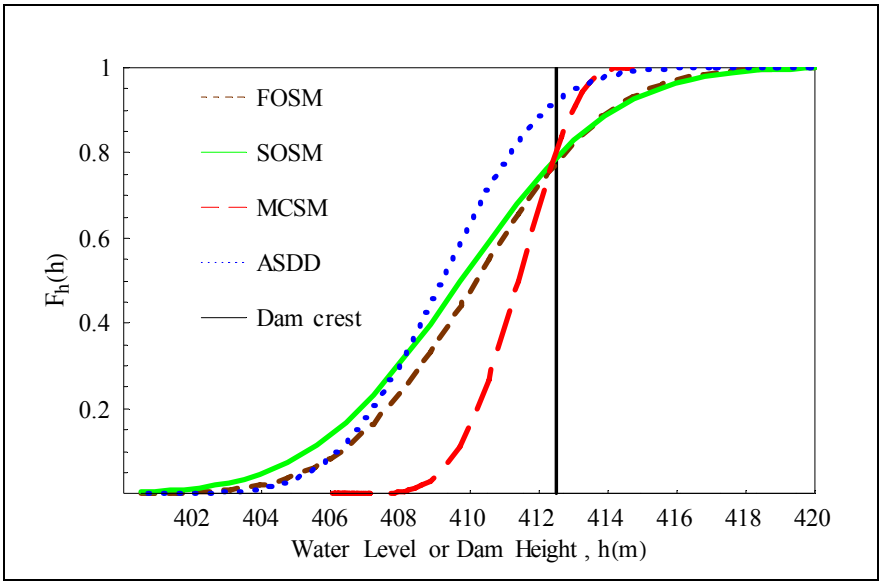


Figure 7-8. $P_{f, OT}$ due to $Q_{10,000}$ flood plus total wave surcharge.

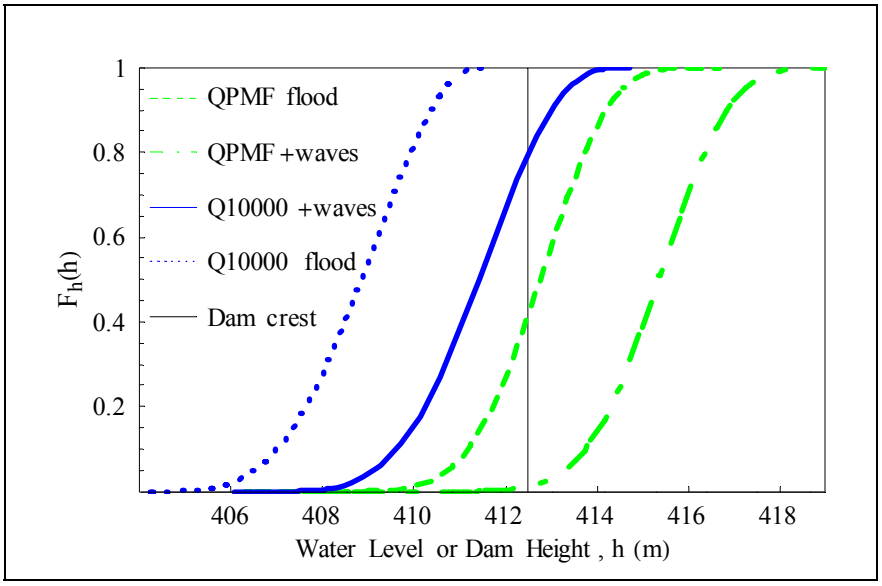


Figure 7-9. $P_{f,\,OT}$ due to $Q_{10,000}$ and Q_{PMF} with and without waves.

$P_{f,\,OT}$ estimated by the different methods are slightly different. But each method offer valid set of procedures that follow internally consistent logic, provided their assumptions are well scrutinized. The ASDD method makes minimum assumptions and it is near to exact. It out perform the other three methods by its theoretical foundation, accuracy and reproducibility.

The classical methods, when compared with ASDD method, overestimated $P_{f,\,OT}$ on average by about 13% and $R_{f,\,OT}$ by 0.0649%. The difference in $P_{f,\,OT}$ between $Q_{10,000}$ and Q_{PMF} flood is big. For example, when only h_f is considered (excluding occurrence of concurrent waves), the difference is $P_{f,\,OT}$ of 0 to 6% for $Q_{10,000}$ versus 58% for Q_{PMF} (see Figure 7-9). This shows that while the dam is safe to slightly over designed for a $Q_{10,000}$ events, it however has serious capacity short coming in an event of Q_{PMF}. Acceptability of such chances of failures can only be judged when there is safety standard explicitly defining tolerable volume of failure based on economic, social, environmental, political, etc criteria. In any case, such dosing of imbedded volume of failure in designs is possible only when using probabilistic techniques like the ones demonstrated. Using such analysis a dam height corresponding to quantified level of $R_{f,\,OT}$ can be selected. From the results it is clear that deterministic approach obscures reality and is poor as an option.

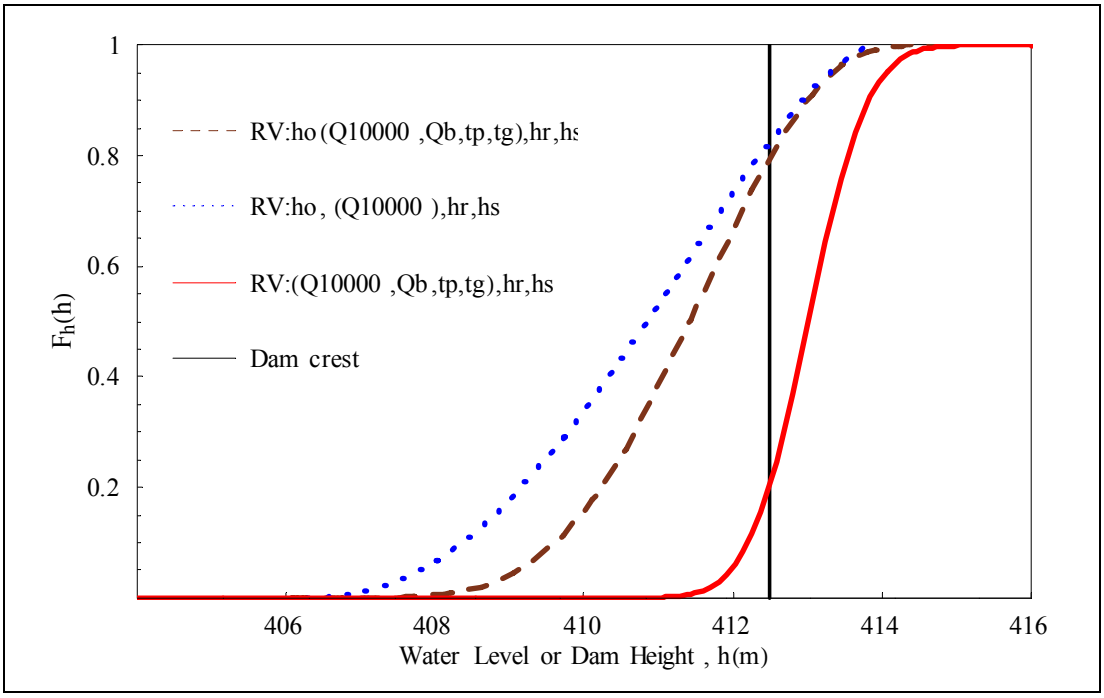


Figure 7-10. Effects of taking elements of FH and h_o as RV.

Table 7.8: Evaluated overtopping performance randomness and $P_{f,OT}$ values.

Method	Cause of water level (WL) surcharge	μ_{WL}	σ_{WL}	$P_{f,OT}$	Actual overtopping Prob. in 50 yrs, $R_{f,OT}$
MCSM	$Q_{10,000}$ Flood	408.76	1.26	0	0
	Total Waves ($h_r + h_s$)	2.60	0.45	-	-
	$Q_{10,000}$ Flood + ($h_r + h_s$)	411.36	1.28	0.2011	2.01×10^{-5}
	Q_{PMF} Flood	412.72	1.15	0.5865	-
FOSM	$Q_{10,000}$ Flood	407.62	3.02	0.0531	5.31×10^{-6}
	Total Waves	2.60	0.45	-	-
	$Q_{10,000}$ Flood + ($h_r + h_s$)	410.19	3.05	0.2244	2.24×10^{-5}
SOSM	$Q_{10,000}$ Flood	407.17	3.44	0.0606	6.06×10^{-6}
	Total Waves	2.60	0.45	-	-
	$Q_{10,000}$ Flood + ($h_r + h_s$)	409.76	3.47	0.2144	2.14×10^{-5}
ASDD	$Q_{10,000}$ Flood	406.61	2.30	0.0052	5.2×10^{-7}
	Total Waves	2.60	0.45	-	-
	$Q_{10,000}$ Flood + ($h_r + h_s$)	409.21	2.34	0.0802	8.02×10^{-6}

Considering Q_p as the only RV in FH under estimated $P_{f,OT}$ (see Figure 7-10). However, the difference between computed $P_{f,OT}$ values for the cases where Q_P is the only RV and where all FH variables (Q_P, Q_b, t_p, t_g) are considered as RV decays with an increase in dam height (see Figure 7-10). The decrease in influence of FH elements randomness with an increase in dam height can partly be due v-shapes of river valleys. At higher heights there is a relatively greater storage area per unit depth that will have an attenuating effect on FH fluctuations. Thus, effect of taking all or part of FH variables as RV is especially significant in middle range dams and this is a range where usually design decisions are made based on economic evaluations. Apparently this effect is dependant on shapes of storage elevation curves.

In the same Figure 7-10 it can be seen that influence of h_o randomness is more significant than influences of FH elements. Effect of FH shape parameters randomness is dwarfed by h_o randomness. However, this fact may be relevant to only Tendaho's peculiar characteristics; i.e. a dam for irrigation purpose in tropical climate with bimodal rainfall pattern causing considerable seasonal fluctuations of stored volumes. In places where such operational water level fluctuations are not significant the randomness induced by the hydrograph elements might be equally significant.

7.6 Chapter conclusions

The presented probabilistic methods and implementation architecture are proved to be applicable for dam overtopping probability evaluation. The approaches followed to drive a fully stochastic hydrograph and randomness of initial water level and waves together with the stochastic routing method applied are found to be practical. The analyses made exposed facts that could potentially be obscured when using classical deterministic dam design approaches. The proposed approaches produce more realistic solutions. Using these methods engineers could design systems to meet a quantified level of performance (volume of failure), which allows economic optimization and uniform distribution of safety across the systems safety chain. Moreover, realizing and quantifying performance randomness helps for explaining problems encountered during dam operation and it assists in making intelligent dam upgrading and rehabilitation decisions. The ASDD method out perform the classical methods by its theoretical foundation, accuracy and reproducibility. However, it is a good practice to compare results from the different probabilistic methods. The error induced by using different methods may not be bigger than that due to influences of data and model completeness. These later two problems are in fact equally shared by deterministic methods and therefore it does not make the deterministic method any better. All the probabilistic methods discussed are much better than deterministic approach.

8 CONCLUSIONS AND RECOMMENDATIONS

8.1 On the deterministic and stochastic/analytical methods

Traditionally uncertainty in dam design practice is assumed to be accounted for through use of safety factors. By selecting generous safety factors dam performance is assumed to remain in an acceptable range and it is presumed that dams never fail. This classical safety factor based deterministic design procedure is poor as an option. It is not accurate and transparent. It does not adequately describe reality, it potentially obscure variability in design parameters and thus provide vague performance estimates. This approach does not allow for transparent accounting of uncertainties and for numerical quantification of safety. It does not permit for optimizing safety and economy through associating a quantified level of risk to alternative design options. If design is complex, sometimes safety factors can compound to cause over design. Other times, they disguise uncertainty in critical design parameters and lull engineers with false sense of safety, whereas there still is one or more under designed components in the system. Using this method it is hard to give uniform reliability distribution in the system safety chain, i.e. one or more weak links could exist in the safety chain that put the entire system only as safe as the weakest element in the chain. In general, it is unsuitable to the 21st century societal requirements and standards. In this research these shortcomings are sufficiently explained using a case study by comparing results from the classical deterministic method with results from proposed modern stochastic/analytical methods.

Nowadays, in most modern engineering codes and society, the appropriateness of the safety factor approach is being questioned. However, dam engineering as a profession is not adequately presented with demonstrated analytical and probabilistic techniques, implementation strategies and application tools to meet the growing demand for transparency and precision in dam safety analysis. It is a thorny issue to analyze uncertainty in design parameters and integrate it in risk based design formulation. Until the tools and knowledge gained through research become adequate enough to deal with uncertainties in dams design and to permit good quantitative descriptions of safety issues, design practices will remain entirely in the realm of subjective judgments. The introduction of risk based design and probabilistic approaches will not make design practices error proof and absolutely accurate; at least at its current level of detail and precision. It will have to go a long way to claim that and become a widely practiced approach but even a tiny improvement may help.

This research dealt with the determination of failure probabilities of dams, which are fundamental ingredients in dam safety analysis, through the use of a multitude of rigorous probabilistic and analytical approaches and updated design procedures. It addresses techniques of doing probabilistic design and safety analysis. It conceives a variety of approaches for customizing classical design equations, which are originally set for deterministic design applications, so that they suit the intended probabilistic analysis. A multitude of tailor-made reliability equations and solution procedures that will enable the implementations of stochastic and analytical methods of computing failure probability have been crafted and tested. Various ways of analyzing and representing uncertainty in parameters of design equations, pertinent to three dominant dam failure causes (sliding, overtopping and seepage), are identified and a suite

of stochastic models capable of capturing the parameters uncertainty, to better facilitate evaluation of failure probabilities (functional and structural failures) are tested.

The research basically studied three classical stochastic models (MCSM, FOSM and SOSM methods) and applied them for modeling dam performance and evaluating failure probabilities in line with the above mentioned dominant dam failure causes. Moreover, the research presented an exact new to the purpose analytical method of transforming design parameters distributions to a distribution bounding dam performance randomness. From this analytically transformed distribution failure probabilities can easily be computed. This method is called ASDD method. The method's basic principles are proved, integrated and well presented in such a way that they are understandable by practicing engineers and in a way convenient for implementation. Generic implementation architectures that are easy to follow are prepared.

Applicability of the methods in dam engineering is demonstrated using a real life case study from Tendaho Dam, Ethiopia and considering the three dominant dam failure mechanisms. Codes and tools for doing the necessary probabilistic computations on the major dam failure mechanisms are provided. Applications of the techniques using engineering, mathematical software and supporting technologies is demonstrated. The performance of the various stochastic and analytical methods is compared with each other and with results from classical deterministic design approach.

The results indicate that the proposed stochastic and analytical approaches provide a valid set of procedures, internally consistent logic and they produce more realistic solutions than the classical deterministic safety factor based approach. Using these approaches engineers could design dams to meet a quantified level of performance (volume of failure) and set a balance between safety and economy.

However, it has to be noted that risk analysis is an enhancement but not substitute to good engineering intuition. What good will a risk based design optimize the hell out of a bad idea? Clearly the critical ingredient that determines the appropriateness of a solution is the problem relevance and the solution approach, which both relay on a good engineering intuition. The best strategy is one that emphasizes both good engineering intuition and risk based design as complement of each other. The recommendation given here in is thus, good engineering intuition shall be supplemented with transparent risk based probabilistic design instead of the safety factor based deterministic design approaches. In this regard, the research is assumed to bridge the gap between the available probability theories in one hand and the suffering risk based design in dam safety evaluation on the other hand. Out of the suite of stochastic/analytical approaches studied the ASDD method out perform the classical methods (MCSM, FOSM and SOSM methods) by its theoretical foundation, accuracy and reproducibility. However, when compared with deterministic approach, each of the stochastic approaches gave more realistic solution.

8.2 On seepage and sliding probability analyses

The stochastic and analytical methods presented in this research are found to be applicable for assessing sliding probability of embankment dams. There is a reasonable match among the sliding failure probability ($P_{f,s}$) determined by the different methods. It is proved that

the absolute safety suggested by the deterministic method, as granted by a computed factor of safety against sliding ($F_{s,\ s}$) of 1.72, is just naive. The stochastic/analytical methods estimate a prevailing sliding failure probability ($P_{f,s}$) that ranges between 2.4 and 4.35%. Using the computed $P_{f,\ s}$ it is possible to do risk analysis and economic optimization that will enable intelligent dam rehabilitation and upgrading decisions.

In case of seepage flux (q) computation, results from FOSM and SOSM showed some divergence from those found from ASDD method. FOSM and SOSM do not provide a distribution of known standard type for characterizing the randomness of outputs of FRV, say the equation for q. They do not either give arrays of possible values for performance, like MCSM, on which distributions that characterize performance can be fitted. Rather results from FOSM and SOSM are mostly presented using Normal distribution. The Normal distribution is constructed from computed first two moments and by assuming Central Limit Theorem (CLT), which states interaction of independent but not identically distributed parameters tend to be normally distributed (when no one distribution is dominant). The possible explanation for the deviation of the FOSM and SOSM results from that of ASDD results is the error induced by this assumption. This assumption is not plausible in seepage analysis, because the dominant design parameter permeability (K) has highly skewed distribution. The accuracy of the result from FOSM and SOSM methods is dependent up on the degree of linearity of the FRV of interest near the mean values of its random variables. The more linear the function is the more accurate the result will be. The appropriateness of CLT assumptions can be evaluated by computing skew and higher order moments for checking symmetry and squatness. In case of seepage flux computation, MCSM also showed problem of convergence. In such circumstances ASDD gives the most reliable result. Moreover, the ASDD method helped to explain and illustrate the shortcomings of the stochastic simulation and moment approximation methods.

8.3 On dam overtopping probability analysis

The presented stochastic/analytical methods and the implementation architecture proposed for their application in dam overtopping probability evaluation are proved to be valid. The approaches followed (1) to drive a fully stochastic hydrograph, (2) to estimate randomness of initial water level and wave heights and (3) the deployed stochastic routing method are found to be practical. The analysis made exposed facts that could potentially be obscured when using classical deterministic dam design approaches. The proposed approaches produce more realistic solutions.

The overtopping probability ($P_{f,\ OT}$) estimated by the different methods are found to be slightly different. But, each method offer valid set of procedures in that they follow internally consistent logic, provided their assumptions are well scrutinized. The ASDD method makes minimum assumptions and it is near to exact. It out perform the other three methods by its theoretical foundation, accuracy and reproducibility. The classical methods, when compared with ASDD method, overestimated $P_{f,\ OT}$ on average by about 13% and $R_{f,\ OT}$ by 0.0649%. It is good practice to compare results from the different probabilistic methods. The error induced by using different methods may not be bigger than that due to influences of data and model completeness. These later two problems are actually equally shared by deterministic methods and therefore, it does not make the later any better.

The difference in the computed $P_{f, OT}$ for $Q_{10,000}$ and Q_{PMF} flood is big; for example, $P_{f, OT}$ of 0 to 6% versus 58% respectively (when only h_f is considered-excluding occurrence of concurrent waves). Therefore, while Tendaho dam is safe to slightly over designed for Q_{10000} events, it however has serious capacity short coming in an event of Q_{PMF}.

Considering Q_p as the only RV influencing flood hydrograph shape (FH) resulted under estimation in the computed $P_{f, OT}$. It is found that the gap in values of computed $P_{f, OT}$ for the cases where only Q_P is taken to RV and when all FH variables (Q_P, Q_b, t_p, t_g) are considered as RV, decays with increase in dam heights. The possible explanation for the decrease in influence of FH elements randomness with an increase in dam height can partly be the v-shapes of the river valleys. At higher heights there is a relatively greater storage area per unit depth that will have an attenuating effect on FH fluctuations. Apparently this effect is dependant on shapes of storage elevation curves. Thus, effect of taking all or part of FH variables as RV is especially significant in middle range dams and this is a range where usually design decisions are made based on economic evaluations. In addition, it is observed that influence of h_o randomness is more significant than influences of FH elements. Effect of FH shape parameters randomness is dwarfed by h_o randomness. However, this fact may be relevant to only Tendaho's peculiar characteristics; i.e. a dam for irrigation purpose in tropical climate with bimodal rainfall pattern causing considerable seasonal fluctuations of stored volumes. In places where such operational water level fluctuations are not significant the randomness induced by the hydrograph elements might be equally significant.

8.4 Recommendations and outlooks

The risk-based dam design approach, which for decisions bases on finding economic optimum points between cost of dam structures and cost of damage in case of failure, will have to go a long way to become an attractive approach when dealing with large dams where rupture would endanger downstream population (especially if there is a large town or city for which "absolute safety" will be demanded- though no method guarantee this). However, it can be attractive approach when there is no risk of loss of life, as with dams in remote areas, near seas or lakes; or those small dams impounding small reservoirs; or those lower dams in wide valleys where rise in water level due to dam failure is not significantly larger when compared with the same flood occurring with out the dam; or when there is reliable warning and evacuation system for the downstream population. The later cases, those which make the approach attractive at its current level of detail and precision, are prevalent in most developing countries with scarce financial resources, countries like Ethiopia, countries where extensive dam construction projects are kicking off in remote areas and where a number of dam upgrading and rehabilitation projects are foreseen. However, to establish generality of the conclusions about the risk-based design approach methods probabilistic analysis and representation of damages/consequences should be researched and equally developed. This area is a researchable field with equal breadth and depth as the problem of probabilistic evaluation of dam failures itself.

The scope of this research is limited to analysis of failure probabilities due to interaction of uncertain strength and load design parameters using stochastic models that take probability theory as their foundation. Nonetheless, uncertainties in strength and load parameters could be of different type. Some has to do with processes that are inherently unpredictable (inherently

random), like occurrence of flood, occurrence of earthquake, patterns of the weather etc. Such uncertainties are characterized by purely objective information content, they satisfy statistical laws and subjective influences are not taken into account. They are some times described as aleatory uncertainties. Such uncertainty can effectively be treated using probability theory. In contrast in other design parameters uncertainties could result from limited knowledge, which has to do with uncertainty about models, uncertainty due to information deficit, uncertainty due to lack of understanding of physics of a system etc. For example, when only small number of observation is available, or when boundary conditions are subject to arbitrary fluctuation, or when system overview is incomplete. These are those like, uncertainty on whether or not a climate changes, likelihood of geological problems- faults and soft lenses etc. Such uncertainties are characterized by subjective influences and they have non-statistical properties. Thus, probability theory is not a perfect uncertainty model to describe them. In such cases, other uncertainty models like fuzzy set theory and theory of fuzzy random variable could be used most effectively. Therefore, it is recommended to extend the implementation and theory to include other uncertainty models than probability theory, such as fuzzy set theory and theory of fuzzy random. However, it has to be noted that the classification of uncertainties into two components, those resulting from natural variation and limited knowledge, is just a modeling decision. It is just to give the implication that probabilities assigned to natural variations are statements about frequencies of occurrence in time or space, which can better be modeled using stochastic models, and those probabilities assigned to limited knowledge are statements about degree of belief. Otherwise, both categories are measured as mathematical probabilities.

The ability of dams to survive hazards and loads does not only depend on capacities provided for its components during design but also on certain other features that come later in its life cycle; such as construction, operation, surveillance and maintenance. For a properly designed dam reliance has to be made on construction, operation, surveillance and maintenance for it to remain in safe condition. Moreover, in the case studies not all uncertain parameters are taken as random variables. This is limited either because of significance of a variables influence on the final result or because of lack of data. For instance, in stability analysis dynamic earthquake loading is not considered because of lack of the necessary records at the case study site. These factors are also uncertain parameters but they are not dealt with in this research. Method of characterizing and quantifying influences of the above mentioned missing features and parameters worth investigation.

Acceptability of computed probability of failure can only be judged when there is safety standard that explicitly define tolerable volume of failure based on economic, social, environmental, political etc criteria. For successful implementation of risk based design such standards need to be researched, developed and enforced.

It is advisable to extend the implementation structures to include all failure mechanizes of dams, other than the ones considered in this research case study, and demonstrate them.

9 APPENDICES

9.1 Appendix to chapter two and three

9.1.1 Statistical and mathematical background

To assist a better understanding of the procedures for solving Eq. (2-3) using either the numerical approximate methods or the analytical exact method it is important to review some topics in statistics and probability theory. This appendix gives such statistical and mathematical background. Generally, reading this appendix before chapters 3 and 4 might facilitate the better understanding of concepts. However, for a reader with sound knowledge on statistics and probability theory reading this section first is not a necessity.

9.1.1.1 Descriptive statistics and data reduction

Descriptive statistics is the art of presenting data in such a way that the useful information they contain become readily available. After taking observations in a laboratory or taking records from hydrometrological stations we have a serious of numbers without much order, for example like the raw data on ϕ' from triaxial tests (see Table 9.1). Such bulk of data has to be condensed or reduced by using techniques like grouping, tally-charting, calculating absolute and relative frequencies etc. This procedure of condensing data is called data reduction. Reduced data are presented in such a way that they clearly present the useful information they contain for the purpose of both conventional deterministic safety factor calculations or for evaluation of safety through probabilistic–based methods. For example, Table 9.2 shows the results of the triaxial test in a reduced form. This table gives the data in a more clear way in such a way that a better understanding about the variability of the values with is possible.

Table 9.1: example raw data on shell material angle of friction ϕ' (degrees).

35	45	48	43	40	42	42	38	41	45	45	40
44	40	46	38	44	40	40	42	44	41	46	48

For presenting data in reduced form mostly the following statistics and graphical representations of statistics are used:

- absolute frequencies (n_j) and relative frequencies (f_j)[6],
- cumulative frequencies (N_j) and relative cumulative frequencies (F_j),
- class and class marks (for data represented in a grouped form)[7],

[6] When only one set of observation is under consideration absolute frequency or the absolute cumulative frequency is sufficient. When comparing two, or more, set of observations, not of the same size, it is easier to work with relative frequencies. Relative frequencies (absolute frequencies divided by the total number of observations) eliminate the influence of sample size.

[7] Number of classes must be chosen with care. A small number of classes end to a great deal of information that is lost. Usually 5 to 10 classes work well as a first idea. When we have n observations a suitable number of classes is given by $\sqrt{n}$, rounded of to the nearest integer. Preferably we take classes of equal width. However, this is not a necessity and often at tails wider classes are used.

– histograms and ogives,
– descriptors or moments (see section 9.1.1.6).

Table 9.2: Reduced data on shell material angle of friction ϕ' (degrees).

Class limit	(Class boundaries]		Class mark X_j	Absolute frequency n_j	Relative frequency f_j	Cumulative frequency N_j	Relative cum. frequency F_j
(34.5-37]	34.5	37.0	35.75	1.00	0.04	1.00	0.04
(37-39.5]	37.0	39.5	38.25	2.00	0.09	3.00	0.13
(39.5-42]	39.5	42.0	40.75	10.00	0.43	13.00	0.57
(42-44.5]	42.0	44.5	43.25	4.00	0.17	17.00	0.74
(44.5-47]	44.5	47.0	45.75	5.00	0.22	22.00	0.96
(47-49.5]	47.0	49.5	48.25	1.00	0.04	23.00	1.00
sum				23.00	1.00		
Mean	42.16			Skew	2.16		
variance	8.99			Coeff. Var.	14.06		
Stand. Div.	3						

In section 5.4 relevant reduced data, in relation to loading and resistance terms, of the case study dam are presented. Additional information on data reduction techniques in engineering can be found in (Devore and Peck 1986; Ang and Tang 1975; Ehrenberg 1986; Benjamin and Cornell 1970).

9.1.1.2 Probability

Ansell (1994) describes that probability is introduced through the axiomatic approach due to Kolmogorov. This approach considers an experiment and every possible outcome (result) of the experiment. The set which is the totality of all the outcomes will be defined as the *sample space*. *Elementary event* is an outcome of the experiment that appears in the list of possible outcomes (sample space). An *event* (realization) is a collection of one or more elementary events (outcomes) which can be considered as a subset of the sample space.

As an example of an experiment assume a laboratory test to determine the shear strength parameter ϕ' (angle of friction, degrees) of a shell material for a dam. The results for example could be any thing between 35^o and 48^0 rounded to the nearest whole number, i.e. 35, 38, 39,….,48. If we perform the experiment once, it will always give one of the results in the above list. However, if we perform this same experiment a number of times we will notice that some of the results occur more often than others (see Table 9.1).

The list of possible out comes can be simplified to:

G: the value of ϕ' is between 35^o and 41^o (which implies ϕ' can be represented by the average 38^o)
F: the value of ϕ' is between 41^o and 48^o (which implies ϕ' can be represented by the average 44.5^o)

If one perform the results a number of times and record the results (Table 9.1). This will result something like:

G	G	F	F	G	F	F	G	G	F	F	G
G	G	F	G	F	G	G	F	F	G	F	F

After each experiment we can calculate the relative frequency of the result G has occurred.

1/1; 2/2; 2/3; 2/4; 3/5; 3/6; 3/7; 4/8; 5/9; 5/10; 5/11; 6/12;7/13; 8/13;…

After n experiments, in which n_G times the result G has occurred, the relative frequency of G is equal to: $f_G(n) = n_G/n$. It is clear that the value of $f_G(n)$ will always be between 0 and 1. When n becomes very large, the value of $f_G(n)$ will remain in a neighborhood of a certain value, with only very small fluctuations. This is the value we call the probability of the result G. The notation used is $P(G)$. Probability is idealized relative frequency.

i. Simple set theory definition

To understand the relations between two or more events understanding of simple results of set theory is advantage.

1. *The complement of the event A*: all possible outcomes not belonging to an event *A*. Notation $\bar{A}$. Called as not *A*. It is the event that *A* does not occur.
2. *Mutually exclusive events* (also called disjoint events): events that can not happen together. So elementary events are always mutually exclusive.
3. *Sure event*: the event that happens always. Notation Ω.
4. *Impossible event*: the event that never happens. Notation ϕ.
5. *The union of two events A and B*: the event that *A* happens of *B* happens or both *A* and *B* events happen. Notation $A \cup B$ we say *A* or *B*.
6. *The intersection of two events A and B*: the event that both *A* and *B* happen. Notation $A \cap B$ or *AB*. We say *A* and *B*.

ii. Classical definition of probability

Suppose that an experiment with m equally likely possible results is performed. Suppose further that for the event A there are h favorable results (in other words the event A is the union of h equally likely elementary events). Then the probability of A is equal to:

$$P(A) = \frac{h}{m}$$

Clearly such a quotient, which defines a probability, satisfies the following rules:

$$0 \le P(A) \le 1$$

$$P(A)=1$$

$$P(\phi)=0$$

$$P(A)+P(\overline{A})=1$$

$$P(AUB)=P(A)+P(B), when\ A\ and\ B\ are\ disjont\ events.$$

Many experiments have an infinite number of possible outcomes. For instance, if the experiment is to measure the shear strength parameter ϕ'. For example, that any value between 30^o and 50^o is equally likely. Specify the event A as $\phi' < 45^o$.

The total number, m, of possible outcomes of the experiment is equal to ∞. The total number of outcomes favorable for the event A, h_A, is also equal to ∞. So the quotient h_A/m, which would define the probability of A is, not defined. But, as all outcomes are equally likely, it is reasonable to assume that the four events: ϕ' between 30 and 35; ϕ' between 35 and 40; ϕ' between 40 and 45 ; ϕ' between 45 and 50. All have the same probability. Together these four must add up to one, so each of the four probabilities will equal to ¼. With this trick it is possible to use the classical definition of probability also in cases where the number of outcomes is infinite. Thus the definition for probability can be rephrased as: favorable length divided by possible length.

Another example of a continuous sample space is the following: assume a 50 cm radius circular dartboard. The probability of hitting the board within the circle of radius 30 cm is then found as favorable area divided by possible area. $=\pi\cdot(30)^2/\pi\cdot(50)^2=0.36$. Note that the probability of hitting exactly on the circle with radius 30 cm is equal to zero because the favorable area is equal to zero. In the case of hitting dartboard example the sample space is continuous. A continuous sample space always has infinite number of elements. A discrete sample space can have a finite number of elements (possible outcomes) or an infinite number.

iii. Axiomatic definition of probability

Probability is a mathematical form of model building. As long as we stick to the rules, we can build any model we like. The rules, called axioms, are simple and the same as those that followed from working with empirical definition, which are give above using relative frequencies or classical definitions.

Axioms of probability:

1. $P(\Omega) = 1$; the total amount of mass is 1;
2. $P(A) \geq 0$; for any event A the amount of mass in region is 0 or more;
3. $P(AUB)=P(A)+P(B)$, if the event A and B are disjoint.

These axioms tells for every outcome (elementary event) E_i = *1, 2,…* in the countable sample space a real non-negative number $P(E_i)$, the probability of E_i, is defined. The probabilities are assigned so that their sum $\sum_i P(E_i)$ is unity. Hence, every outcome has a probability which is not greater than one. This definition can be extended to events by defining probability P(A) of an event A, say, to be equal to the sum of the probabilities of the collection of outcomes which define A.

iv. Some basic rules

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

De Morgan Rules:

$$\bar{A} \cap \bar{B} = \overline{A \cup B} \text{ and } P(\bar{A} \cap \bar{B}) = P(\overline{AUB})$$

$$\bar{A} \cup \bar{B} = \overline{A \cap B} \text{ and } P(\overline{A \cup B}) = P(\overline{A \cap B})$$

Associative/Distributive Properties:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$A = (A \cap B) \cup (A \cap \bar{B})$ and as $A \cap B$ and $A \cap \bar{B}$ are disjoint events, it follows that

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

Some generalization of the above formula:

$$\begin{aligned} &P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) \\ &P(\bar{A} \cap \bar{B} \cap \bar{C}) = P(\overline{A \cup B \cup C}) \\ &P(\bar{A} \cup \bar{B} \cup \bar{C}) = P(\overline{ABC}) = 1 - P(ABC) \end{aligned} \qquad (9\text{-}1)$$

v. Conditional probability

Sometimes having defined probability with respect to one sample space, it is desired to restrict attention to outcomes from subset of the sample space. For example, probability that a dam collapses given overtopping. It would be possible to redefine the probabilities with respect to this new sample space. However, it is simpler to use the conditional probability.

Definition of conditional probability: provided $P(B)$ is positive (not zero) conditional probability of the event A given B is defined by:

$$P(A|B) = \frac{P(AB)}{P(B)} \qquad (9\text{-}2)$$

Where $P(AB)$ is the probability of the event given by the intersection of A and B, which is defined by the outcomes that are contained in both the events A and B. As long as we stick to the same condition, all the formulas for probabilities remain valid.

$$0 \le P(A|B) \le 1$$

$$P(A \cup C|B) = P(A|B) + P(C|B) - P(AC|B) \text{ etc}$$

Conditional probabilities can be very useful in the determination of absolute probabilities. Absolute is sometimes added to stress that it is not a conditional probability that is meant. So P(A) is (absolute) probability of the event A; $P(A \mid B)$ is the conditional probability of the event A given that B has happened and $P(A|\overline{B})$ is the conditional probability of A given B has not happened. For calculating of $P(AB)$ the following can be used:

$$P(A \mid B) = \frac{P(AB)}{P(B)} \text{ so}$$

$$P(AB) = P(A \mid B) \cdot P(B)$$

But also for the calculation of the (absolute) probability of the event A this detour can be useful. From $P(A)=P(A \cap B)+P(A \cap \overline{B})$ follows, using the last result:

$$P(A) = P(A \mid B) \cdot P(B) + P(A\overline{B}) \qquad (9\text{-}3)$$

Note, $P(A \mid \overline{B}) = \frac{P(A\overline{B})}{P(\overline{B})}$ from this,

$$P(A\overline{B}) = P(A \mid \overline{B}) \cdot P(\overline{B})$$

Thus,

$$P(A) = P(A \mid B) \cdot P(B) + P(A \mid \overline{B}) \cdot P(\overline{B}) \qquad (9\text{-}4)$$

The idea can be generalized, by splitting up the sample space Ω into partition with more than two events. A partition $|B$ of Ω is collection of events B_i $i = 1, 2, 3, \ldots, n$ such that:

$$\bigcup_{i=1}^{n} B_i = \Omega$$

$$B_i B_j = \phi \text{ for all } i \neq j,$$

$$P(B_i) \neq 0 \text{ for all } i,$$

The generalization is known as the theorem of total probability: if $B_1, B_2, \ldots, B_n$ form a partition of Ω then:

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i) \qquad (9\text{-}5)$$

If an experiment can be spited in two phases it is usually simplest to define the partition $|B$ in such a way that the events B_i correspond to the outcomes of the first phase.

It is always easy to explain concepts in probability using simple classical examples, like die, coins etc. In the same tone the concept of total probability can be explained with the following simple example.

Example 9-1-1: There are three boxes: A, B, and C. Box A contains two silver coins; box B contains one silver coin and one gold coin; box C contains two gold coins. One of the boxes is chosen arbitrarily. Determine the probability that the chosen coin is gold.

The event that the coin chosen is of gold will be denoted as G. Choosing box A will be the event A; the events B and C can be defined in the same way. Clearly A, B and C form a partition of Ω. So,

$$P(G) = P(G \mid A) \cdot P(A) + P(G \mid B) \cdot P(B) + P(G \mid C) \cdot P(C) = 0.\frac{1}{3} + \frac{1}{3}.\frac{1}{3} + 1.\frac{1}{3} = \frac{1}{2}$$

Conditional probabilities always work smoothly if the events appearing in the condition are related to an earlier stage of the experiment. Sometimes it is necessary to calculate conditional probabilities whereby the oldest event appears before the vertical line. In the last example for instance $P(B \mid G)$, the probability that box B was chosen, given we end up with a gold coin. In this the *Baye's theorem* can be used.

If $B_1, B_2, \ldots, B_n$, form a partition of Ω then:

$$P(B_j \mid A) = \frac{P(A \mid B_j) \cdot P(B_j)}{\sum_{j=1}^{n} P(A \mid B_j) \cdot P(B_j)} \qquad (9\text{-}6)$$

Applying this to the example,

$$P(B \mid G) = \frac{P(G \mid B) \cdot P(B)}{P(G \mid A) \cdot P(A) + P(G \mid B) \cdot P(B) + P(G \mid C) \cdot P(C)} = \frac{\frac{1}{2}.\frac{1}{3}}{0.\frac{1}{3} + \frac{1}{2}.\frac{1}{3} + 1.\frac{1}{3}} = \frac{1}{3}$$

9.1.1.3 Random variables, probability (density) functions and distribution functions

In many cases it is convenient to characterize a result of an experiment or realizations of a random natural process by numerical values. In this regard the concept of random variables is widely applied. In engineering physical dimensions characterizing load and resistance parameters are normally presented using numbers found either from records of measurement of physical dimensions and natural phenomenon (such as flood) or from laboratory experiments (for example angle of friction ϕ'). The measurements and experiments can naturally be continued indefinitely to give infinite results. The records found either from limited measurements or experiments are actually samples from a population.

A random variable assign a numerical value to every possible out come of an experiment or measurement. When performing the experiment or measurement it will, in general, be impossible to predict the result value with certainty. As soon as the experiment is finished the value, called the realization of the random variable, is known. The set of all possible values of these realizations is called the *range* of the random variable. Random variables are denoted by capital letters *X*, *Y*, *Z* etc. For designating their arbitrary realizations corresponding lower case letters are used *x*, *y*, *z*. A simply example can effectively explain the concept of random variables:

Example 9-1-2: assume there is a small probability (say 0.1) that there will be flood greater than a certain magnitude q^o on a day in December. Suppose it is required to look at flows during three days in December. If, on a day, a flood of $Q \geq q^o$ is measured "*s*" is recorded, otherwise, an "*f*". Introduce a random variable *X* representing the number of flood days among the three. Table 9.3 gives all possible outcomes of the observation with their probabilities and the corresponding realizations *x* of the random variable *X*. Independence between the days is assumed.

Table 9.3: Example for modeling using random variables.

Possible outcome	sss	ssf	sfs	fss	sff	fsf	ffs	fff
probability	0.001	0.009	0.009	0.009	0.081	0.081	0.081	0.729
Realization of *X*	3	2	2	2	1	1	1	0

From Table 9.3 it can be seen that:

$$P(X=0) = 0.9 \cdot 0.9 \cdot 0.9 = 0.729$$

$$P(X=1) = 0.1 \cdot 0.9 \cdot 0.9 = 0.081$$

$$P(X=2) = 0.1 \cdot 0.1 \cdot 0.9 = 0.009$$

$$P(X=3) = 0.1 \cdot 0.1 \cdot 0.1 = 0.001$$

The result can be summarized as:

$$P(X) = P(X=x) = \frac{3!}{x! \cdot (3-x)!} \cdot (0.1)^x \cdot (0.9)^{3-x} \qquad x = 0, 1, 2 \text{ or } 3$$
$$= 0 \qquad \text{else.}$$

This is actually a Binomial distribution with parameters $n = 3$ and $P = 0.1$. The range of this experiment, which represent the set of all possible values of random variable, is [0, 3], i.e. $0 \le x \le 3$.

The function $P(x) = P(X = x)$ is called a *probability function*. The plot of the probability function of example 2-2 is shown in Figure 9-1 (left side). In this example the random variable is discrete because it assumes discrete values (0, 1, 2, 3). Next to the probability function *(cumulative) distribution function* of random variables is used. Distribution function, notation $F(x)$, is defined as:

$$F(x) = P(X \le x) \qquad (9\text{-}7)$$

Therefore, from this definition, the distribution function for example 9-1-2 follows as:

$$P(X \le 0) = 0.729$$
$$P(X \le 1) = 0.729 + 0.243 = 0.972$$
$$P(X \le 2) = 0.729 + 0.243 + 0.009 = 0.999$$
$$P(X \le 3) = 0.729 + 0.243 + 0.009 + 0.001 = 1$$

The graph of the distribution function for example 9-1-2 is shown in Figure 9-1 (right side).

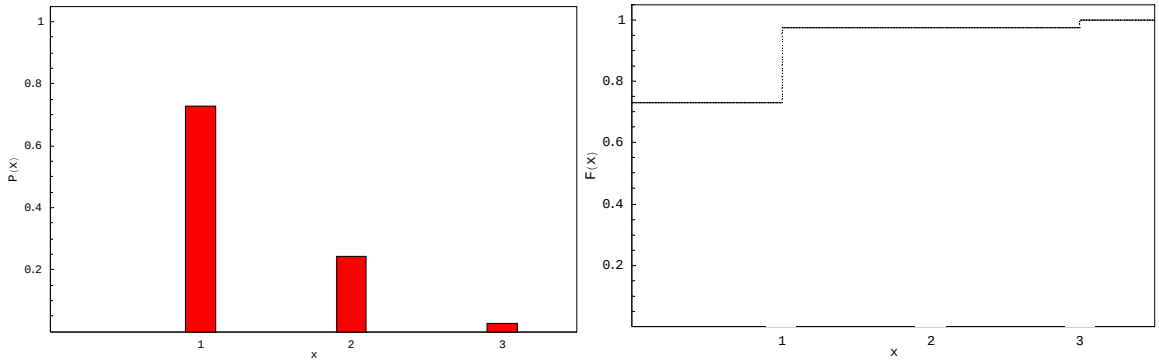


Figure 9-1. Modeling discrete random variables (example).

A probability function (pdf) is only defined for a discrete random variable. For continuous random variable, say Z, $P(z) = P(Z = z) = 0$ for all values of z (see the discussion in the last paragraph of section 9.1.1.2 (ii)). Distribution function (cdf) of a continuous random variable however does not create such problem. So, in case of continuous random variables, it will be easier to start from distribution functions (cdf) and try to understand the underlining principles and formulate a way of expressing probability functions (pdf) for continuous random variables. Again it will be easier to take an example and explain things. In this case we can again look at the case of hitting dartboard, which has been mentioned in the last paragraph of section 9.1.1.2 (ii).

Example 9-1-3: Assume the case of throwing an arrow on 50 cm radius dartboard. Consider the probability of hitting the board within a central circle of radius of 30 cm. Assume the player has not practiced this game before and each point of the board is equally likely to be

hit. In this case the sample contains all points (*x, y*) for which $x^2 + y^2 \le 50^2$. Define the random variable *Z* as the distance, in cm, between the origin and the point that is hit. The range of *Z* is the interval from 0 to 50. It can be seen that:

$$F(0) = P(z \le 0) = 0, \text{ and}$$

$$F(30) = P(z \le 30) = \frac{\pi \cdot 30^2}{\pi \cdot 50^2} = 0.36$$

For any value of *z* which is smaller than zero $F(z) = 0$, because "$Z < 0$" is an impossible event. Also for any value *z* greater than 50 $F(z) = 1$, because "$Z < z$" is a sure event (assume the player is not stupid to miss the entire board at all). For *z* between 0 and 50:

$$F(z) = \frac{\pi \cdot z^2}{\pi \cdot 50^2} = \frac{z^2}{2500}$$

Putting the pieces together the distribution function $F(z)$ is given as:

$$\begin{aligned} F(z) &= 0 \quad , z < 0 \\ &= \frac{z^2}{2500}, 0 < z < 50 \\ &= 1 \quad , z \ge 50 \end{aligned}$$

Its plot is given in Figure 9-2 (left side). In this case there are no jumps in the graph of the distribution function, the function *F*(*z*) is continuous. In the discrete case (Figure 9-1) the probability function gave the increase of the distribution function. In the case of continuous random variables, only the rate of increase of *F*(*z*) is visible. This rate of increase is of course given by the first derivative $\frac{dF(z)}{dz}$. This new function of z is called the *probability (density) function* of the random variable *Z*, notation $f(z)$. The expression for this probability density function $f(z)$ is given below and its graph is provided in Figure 9-2 (right side).

$$\begin{aligned} f(z) &= 0 \quad , z < 0 \\ &= \frac{z}{1250}, 0 < z < 50 \\ &= 0 \quad , z > 50 \end{aligned}$$

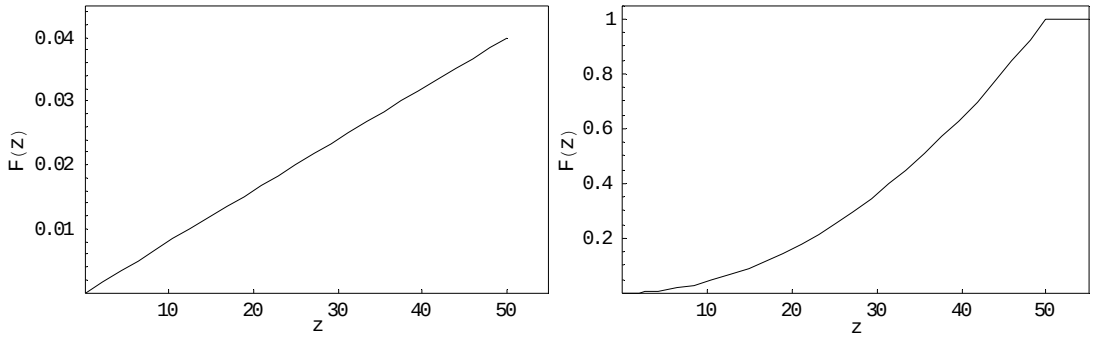


Figure 9-2. Modeling continuous random variables (example).

Results of the above discussion can be summarized as In Table 9.4, in such a way that they give the generic relations for probability (density) and distribution functions of continuous and discrete random variables.

Table 9.4: Basic expressions for probability (density) and distribution functions.

Discrete random variables	**Continuous random variables**
Distribution function $P(x)$	Distribution function $F(x)$
Probability function $p(x)$	Probability density function $f(x)$
$p(x) = p(X = x)$	$f(x) \cong P(x < X \le x + dx)$
$P(x) = p(X \le x)$	$F(x) = P(X \le x)$
$P(x) = \sum_{t \le x} p(t)$	$F(x) = \int_{-\infty}^{\infty} f(t)dt$
$p(x) = P(x^+) - P(x^-) = P(X \le x^+) - P(X \le x^-)$	$f(x) = \frac{d}{dx} F(x)$
$0 \le P(x) \le 1$ for all x	$f(x) \ge 0$ for all x
$\sum_{x} p(x) = 1$	$\int_{-\infty}^{\infty} f(x)dx = 1$
$P(a < X \le b) = \sum_{a < x \le b} p(x)$	$P(a < X \le b) = \int_{a}^{b} f(x)dx$
$F(x)$ is non decreasing; $F(-\infty) = 0$; $F(+\infty) = 1$; $P(a < X \le b) = P(b) - P(a)$	

From the above discussion what can be seen is that a (probability) density function or distribution function can be defined for events (physical phenomenon) using the outcomes of an experiment or using data from records. This can be done by starting from reality and defining probability associated with events provided the range of the random variable is known. In practice, however it is mostly difficult to clearly demarcate the range of random variables and to estimate the probability associated with every event, unlike the simple example cases of the dartboard or the hypothetical three day flood monitoring. Therefore, in engineering practice, mathematical models of probability distributions are commonly used. Computational convenience combined with empirical verifications supports the use of such mathematical functions to approximate probability distributions. Hartford and Baecher (2004) states that a comparatively small set of mathematical functions can be used to fit a broad range of frequency distributions encountered in practice.

Baecher and Christian (2003) tabulated the mathematical equations, ranges and moments for a variety of commonly used probability density functions. By far among the most useful distributions are Normal distribution, log Normal distribution, Exponential distribution, Beta distribution, Gamma distribution, Weibull distribution, although a great many other models exist. Johnson and Koltz (1969, 1970, 1972) and (Johnson et al. 1992) give details on many of these models. Brief and basic information on characteristics, functions and applications suitability of the most common probability density functions in engineering practice is tabulated in appendix 9.1.2.

9.1.1.4 Uni- and multi-variate probability theory

When more than one variable is of concern, the concepts of probability (density) and distribution function described in section 9.1.1.3 and summarized in Table 9.4 directly extends to the joint probability density function and distribution functions of multiple variables. For example for a bi-variate case:

$$\begin{aligned} P_{X,Y}(x,y) &= P((X = x)\ and\ (Y = y)) \\ F_{X,Y}(x,y) &= P((X \le x)\ and\ (Y \le y)) \end{aligned} \qquad (9\text{-}8)$$

The joint probability density function is the continuous analogy of the joint probability density function:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) \qquad (9\text{-}9)$$

The marginal distribution of one variable irrespective of the other(s) is found by integrating (summing in the discrete case) over the distribution of probability in other variable(s), for example:

$$f_X(x) = \int_Y f_{X,Y}(x,y)dy \qquad (9\text{-}10)$$

The conditional distribution of one variable given a particular value of the other variable(s) is found from the relative probability density with the other variable(s) fixed by renormalizing the probability distribution so that its integral is unity:

$$f_{X|Y}(x \mid y_o) = \frac{f_{X,Y}(x,y)}{f_Y(y_o)} \qquad (9\text{-}11)$$

For the special case in which x and y are independent, the following relations hold:

$$f_{X|Y}(x \mid y) = f_X(x) \qquad (9\text{-}12)$$

$$f_{Y|X}(y \mid x) = f_Y(y) \qquad (9\text{-}13)$$

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \quad (9\text{-}14)$$

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y) \quad (9\text{-}15)$$

$$F_{X|Y}(x \mid y) = F_X(x) \quad (9\text{-}16)$$

$$F_{Y|X}(y \mid x) = F_Y(y) \quad (9\text{-}17)$$

In case where the variables are dependent, these simple multiplicative relations do not hold. For more variables than two all the above relations can be extended straight by putting more variables in the list of expressions.

9.1.1.5 Uncertainty modeling and analysis using probability distributions

It has been said time and again that the values of most engineering parameters used in design are random variables and therefore actual values in time and space are likely to be different from the nominal value used during deterministic design. In risk analysis it is required to model the uncertainty of these parameters from which performance and engineering response is calculated using probability models. Measuring values of these parameters repeatedly (in laboratory or by recording natural events) gives data on the frequency of occurrence of the values of the parameters. If there are sufficient data values the frequency can be rescaled to give a probability and it can be represented using any of the best fit mathematical probability density functions discussed in appendix 9.1.2. Considerations while choosing representative distribution to various dam engineering parameters are discussed below.

The following steps are suggested for selecting a probability distribution for modeling specific engineering design parameter. These steps are practically applied to model uncertainty of relevant design parameters for the case study dam (section 5.5):

1. Gathering raw data for all uncertain parameters of interest.

2. Organize the data into a meaningful format and plot it as a frequency distribution on a chart. To create a frequency distribution the total number of available raw data is divided into groups (intervals). There is a recommendation for selection of number of intervals based on the total number of data (n); usually a rule of thumb $\sqrt{n}$ is used to fix the number of intervals. Therefore, when you have a physical parameter a rough estimate of the meaningful range of variation of a parameter can be estimated (for example, between 35^o and 45^o in case of angle of shear resistance for shell material) and then suitable interval width can be selected (for example every 2^o for angel of shear resistance) from this too the approximate number of data required can be determined. The total number of raw data (n) to fit a distribution shall not be less than 15 (Crystal Ball, 2005 a and b). The

groups (intervals) are plotted on the horizontal axis of the chart and the number of frequencies of data values in an interval is listed on the vertical axis.

3. The frequency distribution can be plotted as a probability distribution. The probability distribution shows the number of sample data in each interval as a function of the total number of sample data. To create the probability distribution the number of sample data in each interval is divided by the total number of sample data and is listed on the chart's vertical axis. From the probability distribution the likelihood (probability) that a variable, drawn at random from a population, assumes a given value can be seen.

4. Select a mathematical probability distribution function, among the likes listed in appendix 9.1.2, that can resemble the probability distribution created in step 3 above. Though plotting is one guide to selecting among the standard probability distributions. The following steps provide the complete process used for selecting the distribution that best describe the uncertain variable.
    a) Considering the variable in question list everything known regarding the conditions surrounding the variable. For this, knowledge from theory, laboratory results and own judgment shall be used.
    b) Descriptions and behaviors of available standard distributions (like normal, lognormal, beta, Weibull etc.) is reviewed in quest for a distribution that features the conditions listed in the former step for the variables in question.
    c) Based on the above two set of information select a distribution that characterizes the variable. A distribution characterizes a variable when the conditions of the distribution match those of the variable. In this regard formal goodness-of-fit test techniques, like Chi square (χ^2), Aderson Darling and Kolmogorov-Smirnov tests, could be of help (see also appendix 9.1.3). However, in many cases there are physical conditions that suggest appropriate forms for probability distributions function of an uncertain quantity. In such cases there may be convincing and logical reasons for favoring one distributional form over the other, no matter the behavior of limited number of observed data or best-fit tests suggest. The goodness-of-fit test topic is adequately treated in many classical statistics and reliability engineering books, like in (D'Agostino et al. 1986, Benjamin et al. 1970, Crowder et al. 1991, and Lawless 1982).
    d) Then the parameter values for the distribution are determined. The conditions of the variable describe the values for the parameter of the selected standard distribution. Each distribution has standard set of parameter that defines its probabilistic density function. For example, the parameters that define a normal distribution are mean and standard deviation. These parameters are determined from the data.
5. For sufficient data available (at least 15 for each variable) distribution function fitting, short listing and ranking can be done using computer program Crystal Ball 7.1. But, final selection shall be based on own judgment on whether the conditions of a distribution match those of the variable.

6. Seeing the range of a variable truncation of a distribution might be necessary. By default Crystal Ball 7.1 fits the distribution from negative infinity to positive infinity. Truncating changes the bounds or limits and characteristics of the distribution. Thus the mean and

other parameters of a truncated distribution are automatically adjusted by Crystal Ball 7.1 according to the new range.

Ultimately, the selected probability distributions for respective design parameters are described using their respective parameters, probability density and distribution functions and their set of moments.

a) the parameters of a given type of distribution are the mathematical parameters in the formula for the distribution (not to be confused with design parameters).
b) the probability density function describes the basic shape and location of the distribution.
c) the cumulative distribution function allows reading off the area under the probability density function in a given range. This area represents the probability that the random variable will lie in this range.

The interplay between probability distribution modeled design parameters in random process (performance function, function of random variable) is the topic of chapter 3 and 4.

9.1.1.6 Moments for data set and distributions

Moments of a distribution (data set) are a way of summarizing the important characteristics of the distribution as a single number, without having to cope with too much detail. The first few lower moments are generally of most interesting like the mean and variance in probabilistic design using moment analysis methods. Definitions of the first few moments and coefficients based on them are provided below.

i. Moments of observed data

Non-Central Moment (the mean): the mean is a measure of central tendency (location). The mean of set of observations or realizations $x_1, x_2, x_3, \ldots..,x_n$ is $\bar{x}$ and it is calculated as:

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{k}.x_j \quad \text{(ungrouped)} \qquad (9\text{-}18)$$

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{k} n_j.x_j = \sum_{j=1}^{k} f_j.x_j \quad \text{(grouped)} \qquad (9\text{-}19)$$

Where X_j = class mark (in case of grouped data) or observation (in case of ungrouped data), n_j is number of observations in a class, $n = \sum_{j=1}^{k} n_j$ and $f_j = \frac{n_j}{n}$.

Central Moment (the variance): the variance is a measure of the variability (dispersion). The variance of a data set, σ^2, is calculated as:

$$\sigma^2 = \frac{1}{n-1}\sum_{j=1}^{k} n_j.\left(x_j - \bar{x}\right)^2 = \frac{n}{n-1}\sum\left(x_j - \bar{x}\right)^2 \qquad (9\text{-}20)$$

$\bar{x}$ and σ^2 are the most useful characteristics of a set of data. The mean gives an indication about the order of magnitude of the observed values. The variance gives an indication about whether the observed values are close together or widely varying. Moments, both central and non-central, can also be computed for distributions representing random variables.

Central Moment (skew and Kurtosis): the skew (*s*) is the third central moment. Its unit is the cube of the unit of the random variable and hence may be positive or negative. Skewness for a data set is calculated as:

$$s = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3 \text{ (ungrouped)} \qquad (9\text{-}21)$$

$$s = \sum_{j=1}^{k} f_j \cdot \left(x_j - \bar{x}\right)^3 = \frac{1}{n}\sum_{j=1}^{k} n_j * \left(x_j - \bar{x}\right)^3 \text{ (grouped)} \qquad (9\text{-}22)$$

Kurtosis is the fourth central moment. It measures the "squatness" (peakdness) of a distribution. Kurtosis of a data set is calculated as:

$$s = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4 \text{ (ungrouped)} \qquad (9\text{-}23)$$

$$s = \sum_{j=1}^{k} f_j \cdot \left(x_j - \bar{x}\right)^4 = \frac{1}{n}\sum_{j=1}^{k} n_j * \left(x_j - \bar{x}\right)^4 \text{ (grouped)} \qquad (9\text{-}24)$$

ii. Moments of distributions

Non-Central Moments: The k^{th} (non-central) moment $\mu'_{(k)x}$ of a probability distribution $f(x)$ about the origin is:

$$\mu'_{(k)X} = \int_{-\infty}^{\infty} x^k \cdot f(x)dx \qquad (9\text{-}25)$$

The first non-central moment ($k = 1$) is called the mean. The mean of a random variable X will be denoted by μ_X, or simply μ when the context is clear. The mean is also the expectation of X, denoted $E[X]$.

Central Moments: the nth central moment $\mu_{(k)X}$ of a probability distribution $f(x)$ about its mean μ is:

$$\mu_{(k)X} = \int_{-\infty}^{\infty} (x-\mu)^k \cdot f(x)dx \qquad (9\text{-}26)$$

The first central moment of any distribution is zero. The second (k = 2) central moment is the *variance* (v_X). The third (k = 3) is the *skew* (s_X) and the fourth (k = 4) central moment is *kurtosis* (k_X). Variance measures the spread of the distribution. A zero variance thus implies a deterministic variable. Its unit is the square of the random variable and hence is always positive. Skew measures symmetry of a distribution. Its unit is the cube of the unit of the random variable and hence may be positive or negative. A positively skewed distribution has its longer tail to the right. A negatively skewed distribution has its longer tail to the left. Skew is sometimes used to test how valid it is to assume a given distribution is symmetric and hence perhaps approximatable by the normal distribution. Kurtosis measures the squatness of the distribution. It is useful in differentiating different types of symmetric distributions, for example normal and uniform distributions.

iii. Coefficients based on moments

Standard deviation: the standard deviation of a distribution is the positive square root of the variance. It has the same dimensions as the mean but it is the variance that is the more fundamental quantity. The standard deviation of a random variable X is denoted by σ_X. For well-behaved distribution roughly 2/3 of its values fall in the region $(\bar{x}-\sigma, \bar{x}+\sigma)$.

Coefficient of variation and variance ratio: the coefficient of variation is the ratio of the standard deviation and the mean, and is thus a measure of the relative spread of a distribution. This ratio is dimensionless and so may often be used to cast formulae in a dimensionless form. The coefficient of variation of a random variable X will be denoted by $\hat{X}$. The variance ratio is the dimensionless ratio of the variance to the square of the mean. And is denoted by $\hat{x}^2$.

Coefficient of skewness and coefficient of kurtosis: the coefficient of skewness is the dimensionless ratio of the skew to the cube of the standard deviation. The normal distribution has a coefficient of skewness of zero. The exponential distribution has a coefficient of skewness 2. Coefficient of Skewness (β) is calculated as:

$$\beta = \frac{\mu_3^x}{(\sqrt{\mu_2^x})^3} = \frac{\mu_3^x}{\sigma^3} \qquad (9\text{-}27)$$

The coefficient of kurtosis (γ) is dimensionless ratio of the kurtosis to the fourth power of the standard deviation. It is calculated as:

$$\beta = \frac{\mu_4^x}{(\sqrt{\mu_2^x})^4} = \frac{\mu_4^x}{\sigma^4} \qquad (9\text{-}28)$$

The coefficient of kurtosis measures the peakdness of the type of distribution. Uniform distribution has a kurtosis coefficient of 1.8, triangular 2.4, normal 3, and exponential 9.

9.1.1.7 Expectations of random variables and arbitrary functions

In probability theory and statistics the expected value (also called as expectation, mathematical expectation, mean, or first moment) of a continuous random variable is the integral of the random variable with respect to its possibility measure. For discrete random variables this is equivalent to the probability –weighted sum of the possible values.

The expected value may be intuitively understood by the law of large numbers: the expected value, when it exists, is the limit of the sample mean as sample size grows to infinity. It can be mathematically defined as: if X is a random variable then the expected value of X denoted $E(X)$ is defined as:

$$E(X) = \int_{-\infty}^{\infty} X \cdot f(x)dx \qquad (9\text{-}29)$$

Where $f(x)$ is the probability density function of X. The limits $-\infty$ to ∞ can be replaced by appropriate probability space over which random variable X is defined. If X is a discrete random variable with probability mass function $p(x)$, then the expected value becomes:

$$E(X) = \sum_i x_i \cdot p(x_i) \qquad (9\text{-}30)$$

The expected value of an arbitrary uni-variate function of X, $g(X)$, with respect to probability density function $f(x)$ is given by:

$$E(X) = \int_{-\infty}^{\infty} g(x) \cdot f(x)dx \qquad (9\text{-}31)$$

The expectation of a multivariate function $g(x, y,...)$ of multiple random variables X, Y, …with probability density function $f(x, y,...)$ is denoted $E[g(x, y,...)]$ and is defined as integral:

$$E[g(x,y,...)] = \int_{-\infty}^{\infty} g(x,y,...) \cdot f(x,y,...)dx\,dy... \qquad (9\text{-}32)$$

9.1.1.8 Relationships of expected values and moments

Moments of random variables can be defined as expectations. From the definition of expectation operator given in equation (9-31) for uni-variate function $g(x)$, where X is any random variable with cumulative distribution function (cdf) $F_X(x)$, and $g(x)$ is integrable with respect to $F_X(x)$ on $(-\infty, \infty)$ particular choices of $g(X)$ will be of interest in defining the moments discussed above as expectations. Let k be a nonnegative integer. Then for any real number a, let $g(X) = (X - a)^k$. The k^{th} moment of $F_X(x)$ about a is $E\{(X - a)^k\}$. If $a = 0$, then $E\{X^k\} = \mu_k'$ is called the k^{th} non central moment of $F_X(x)$. In particular μ_1' is usually denoted by

μ and is designated the mean of $F_X(x)$. If $a = \mu$, then $E\{(X-\mu)^k\} = \mu_k$ is called the k^{th} central moment of $F_X(x)$. If $k = 2$, then $E\{(X-\mu)^2\} = \sigma^2$ is the variance of $F_X(x)$. Therefore, the following relations can be written as a summary:

Non-central moments as Expectations

$$\begin{aligned} E[1] &= 1 \\ E[X] &= \mu_X \\ E[X^k] &= \mu'_{(k)X} \end{aligned} \qquad (9\text{-}33)$$

Central Moments as Expectations

$$\begin{aligned} E[(X-\mu_X)^k] &= \mu_{(k)X} \\ E[X-\mu_x] &= 0 \\ E[(X-\mu_X)^2] &= \nu_X \\ E[(X-\mu_X)^3] &= s_X \\ E[(X-\mu_X)^4] &= k_X \end{aligned} \qquad (9\text{-}34)$$

9.1.1.9 Properties of expectation for constants, linear sum and products

The following properties are either a straight follow up of the general definition of expectation of a random variable and or can be proved from Eq. (9-32):

Property 1: Constants

The expectation of a constant *c* is the constant itself.

$$E[C] = c \qquad (9\text{-}35)$$

Property 2: Expectation of a linear sum

If *a, b,* … are constants, and *X*, *Y*,… are independent random variables then the expectation of a linear sum of random variables can be given as:

$$E[a \cdot g_1(X,Y,...) + b \cdot g_2(X,Y,...) + ...] = a \cdot E[g_1(X,Y,...)] + b \cdot E[g_2(X,Y,...)] + ... \qquad (9\text{-}36)$$

Proof: from definition of expectation:

$$E[a \cdot g_1(X,Y,...) + b \cdot g_2(X,Y,...)] = \int_{-\infty}^{\infty} a \cdot g_1(x,y,...) + b \cdot g_2(x,y,..) \cdot f(x,y...) dx$$

From general principle of integral of a sum and applying Eq.(9-32),

$$= a \cdot \int_{-\infty}^{\infty} g_1(x,y,...) \cdot f(x,y...)dx + \int_{-\infty}^{\infty} b \cdot g_2(x,y,...) \cdot f(x,y...)dx$$

Therefore, the following general relation hold true:

$$E\left[a \cdot g_1(X,Y,...) + b \cdot g_2(X,Y,...)\right] = a \cdot E\left[g_1(X,Y,...)\right] + b \cdot E\left[g_2(X,Y,...)\right]$$

Property 3: Expectation of product of independent random variables

If X, Y, … are independent random variables, then:

$$E\left[g_1(X) \cdot g_2(Y)...\right] = E\left[g_1(X)\right] \cdot E\left[g_2(X)\right]... \qquad (9\text{-}37)$$

Eq. (9-31) to (9-37) are important relations for deriving moments for special case design equations that are function of several random variables made of a linear sum and/or positive integer powered random variables and constants (discussed in section 3.2).

9.1.1.10 Relations of central and non-central moments

Using properties of expectations for product and linear sum of independent random variables Eq. (9-36) and(9-37) and relations of expectations to moments Eq. (9-33)and(9-34), central moments can be expressed in terms of non-central moments and vice versa. This will provide important relation that will be used to drive general relations for moments of simple multi-variate function of random variables (see section 3.2).

Central moments in terms of non-central moments

The central moment $E[(x-\mu_X)^k] = \mu_{(k)X}$ can be written in terms of Binomial Expansion[8] as:

$$E\left[\sum_{i=0}^{k}\left(Binomial[k,i] \cdot x^i \cdot (-\mu_x)^{k-i}\right)\right] = \sum_{i=0}^{k}(Binomial[k,i] \cdot E\left[x^i\right] \cdot (-\mu_x)^{k-i}] \qquad (9\text{-}38)$$

Proof:

$$(X-\mu_X)^k = \sum_{i=0}^{k}\left(Binomial[k,i] \cdot x^i \cdot (-\mu_x)^{k-i}\right)$$

It follows,

[8] Binomial Expansion of power of sums is given as $(x+y)^n = \sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i}$. Where the Binomial Coefficient $\binom{n}{i}$ is given as $\binom{n}{i} = \frac{n!}{i!(n-i)!} = \frac{n.(n-1)....(n-i+1)}{i.(i-1)....1}$. And $\binom{n}{i} = 0$ if $i < 0$ or $i > n$.

$$E\left[(X-\mu_X)^k\right]=E\left[\sum_{i=0}^{k}\left(Binomial[k,i]\cdot x^i\cdot(-\mu_x)^{k-i}\right)\right]$$

Noting that $Binomial[k,i]$ and $(-\mu_x)^{k-i}$ are constants and applying Eq. (9-37) gives:

$$E\left[\sum_{i=0}^{k}\left(Binomial[k,i]\cdot x^i\cdot(-\mu_x)^{k-i}\right)\right]=\sum_{i=0}^{k}(Binomial[k,i]\cdot E\left[x^i\right]\cdot(-\mu_x)^{k-i}]$$

Using Eq. (9-38) the variance can be expresses in terms of the non central moments as:

$$\nu_X=E[(X-\mu_X)^2]=E[X^2+2\cdot X\cdot\mu_X+\mu_X^2]=E[X^2]-2\cdot E[X]\cdot\mu_X+\mu_X^2=E[X^2]+\mu_X^2 \qquad (9\text{-}39)$$

Similary the skew can be given as,

$$s_x=E\left[(x-\mu_x)^3\right]=\left[x^3-3\cdot x^2\cdot\mu_x+3\cdot x\cdot\mu_x^2-\mu_x^3\right]=E\left[X^3\right]-3\cdot E\left[X^2\right]\cdot\mu_X+2\cdot\mu_X^3 \qquad (9\text{-}40)$$

Using Eq. (9-39) and (9-40) expressions for expectations of second and third power random variables can be given as:

$$E\left[X^2\right]=\mu_X^2+\nu_X$$
$$E\left[X^3\right]=s_X+\mu_X^3+3\mu_X\nu_X \qquad (9\text{-}41)$$

From similar steps the expression for expectation of fourth power of a random variable is:

$$E[X^4]=k_X+4s_X\mu_X+\mu_X^4+6\mu_X^2\nu_X$$

Non-central moments in terms of central moments

Applying the same principles the non-central moments $E[x^k]=\mu'_{(k)X}$ can be written in terms of the central moments $E\left[(x-\mu_X)^k\right]=\mu_{(k)X}$ as:

$$E\left[X^k\right]=E[((X-\mu_X)+\mu_X)]v^k=\sum_{i=0}^{k}\left(Binomial[k,i]\cdot E\left[(X-\mu_X)^k\right]\cdot(\mu_X)^{k-1}\right) \qquad (9\text{-}42)$$

Proof: adding and subtracting a constant of same magnitude on a variable once keeps the variable unchanged. Thus,

$$X=X-\mu_X+\mu_X$$

Therefore, we can write $X^k=(X+\mu_X-\mu_X)^k$. Thus, $E\left[X^k\right]=E\left[((X-\mu_X)+\mu_X)^k\right]$. Using binomial expansion:

$$= E[\sum_{i=0}^{k}\left(Binomial[k,i]\cdot(X-\mu_X)^i\cdot\mu_X^{k-i}\right)] = \sum_{i=0}^{k}\left(Binomial[k,i]\cdot E\left[(X-\mu_X)^k\right]\cdot(\mu_X)^{k-1}\right)$$

Therefore,

$$E[X^k] = E[((X-\mu_X)+\mu_X)^k] = \sum_{i=0}^{k}\left(Binomial[k,i]\cdot E[(X-\mu_X)^k]\cdot(\mu_X)^{k-1}\right)$$

9.1.2 Review on common probability density functions in engineering practice

A brief summary of probability models commonly used for modeling uncertainty of engineering parameters is given summarized below. The contents of this section are mainly modified from (Hartford and Gregory, 2004), (Crystal Ball, 2005), (Lange, 1997), (Crowder et. al, 1991) and (Leitch, 1995).

Normal distribution (adding random variables, standard law of errors)

A random variable X has a Normal distribution with expected value μ, standard deviation σ (variance $\nu = \sigma^2$) if its probability density function is given by:

$$f_X(x) = \frac{1}{\sigma\cdot\sqrt{2\pi}}\cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < +\infty \qquad (9\text{-}43)$$

and its distribution function by:

$$F(t) = P(x \le t) = \int_{-\infty}^{t} f_X(x)dx$$

An analytical expression for this last integral exists only for very special values of t. Its value has to be determined numerically, which is usually expressed in terms of error function (erf) as:

$$F(t) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \qquad (9\text{-}44)$$

Normal distribution is also known as the Gaussian distribution. It gets principal applications in calculations involving sum of random variables. For example, in calculating the limiting shear strength of a potential slip circle through an earthen embankment dam, the mass over the slip surface is usually divided in to slices and the contributing strength of each of the slices are added together to estimate the total resistance. The contributing strength of each of these slices is known only with some uncertainty, say up to a mean and a variance. What distributional form is appropriate for the total resistance? In fact, as n becomes large, the distribution of the sum of independent random variables asymptotically approaches a Normal distribution, almost regardless of the distribution of the underlying variables. This is reflected in the central limit theorem (CLT). If you take random variables that follow any distribution with bounded variance, then the CLT shows that the mean of a large number of these variables always

approaches a normal distribution (Feller, 1967), (Hartford and Gregory, 2004). Thus, for uncertainties, such as the average strength across a slip circle which averages random variations, the Normal distribution is an appropriate model.

For example, if uniformly distributed *n* random variables are considered and normalized sum of 2, 3, …, 10 of these variables is considered, i.e. if two, three, four, …ten variables are summed, and then divided by the number of variables in the sum (for example, when two variables sum is considered the normalized sum is $X = (X_1 + X_2)/2$, when three variables are considered $X = (X_1 + X_2 + X_3)/3$ etc). The normalized sum *X* distribution tends towards normality ("bell-shape") as the number of variables in sum increases. *X* with *n = 10* is more normally distributed than the one with *X* with *n = 2*, where n refers to the number of random variables in the summation.

Lognormal distribution (multiplying variables):

Some calculations in design and risk analysis involve the product of random variables, resulting equations of the form:

$$Z = X_1 \cdot X_2 \cdot ...,\cdot X_n$$

Where, the X_i's are random variables. Taking the logarithm of each side leads to an equation involving the sum of logarithms:

$$\log Z = \log X_1 + \log X_2 + ... + \log X_n$$

Thus, since *logZ* is the sum of logarithms of the X_i terms, the distribution of *logZ* should approach Normality (and hence *Z* approach log normality) as *n* becomes large.

If the mean and standard deviation of the data themselves are μ and σ, respectively, and the means of the logarithms of the data are ζ and λ, respectively, the following relations apply and are useful in working with lognormal distributions (for transforming arithmetic to log parameters):

$$\zeta^2 = \ln\left(1 + \frac{\sigma^2}{\mu^2}\right);\ \lambda = \ln(\mu) - \frac{1}{2}\zeta^2;\ \mu = \exp\left(\lambda + \frac{1}{2}\zeta^2\right)$$

The probability density function of a lognormal distribution is:

$$f(z) = \frac{1}{z \cdot \lambda \cdot \sqrt{2\pi}} e^{-\frac{(\ln(z)-\zeta)^2}{2\lambda^2}} \text{, for x > 0} \qquad (9\text{-}45)$$

Where ζ and λ are mean and standard deviations of the variables natural logarithm. It is not possible to integrate this function in a closed form and the cumulative distribution function of lognormal distribution has to be found numerically. The cumulative distribution is given as:

$$F(z) = \frac{1}{2} + \frac{1}{2} erf\left[\frac{\ln(z) - \zeta}{\lambda\sqrt{2}}\right] \qquad (9\text{-}46)$$

Lognormal distribution is a single tailed distribution of any random variable whose logarithm is normally distributed. If X is a random variable with normal distribution, then $Z = e^X$ has a lognormal distribution; likewise, if Z is lognormally distributed, then $\log Z$ is normally distributed. The base of the logarithm function does not matter: if $\log_a^Z$ is normally distributed, then so is $\log_b^Z$, for any two positive numbers $a, b \neq 1$. It is a distribution suited to a variable which can increase without bound, but is limited to a finite positive value at the lower limit. Also, a variable might be modeled as lognormal if it can be thought as the multiplicative product of many independent random variables each of which are positive. The distribution is positively skewed (most values closer being closer to the lower limit). The logarithms of the variable yield a normal distribution. It gets principal applications in dam engineering for example to model uncertainty in river flows.

Binomial Distribution (success and failure)

Binomial distribution is a discrete distribution. It describes the number of successes in a fixed number of trials. For each trial only two outcomes are possible- success (*s*) or failure (*f*). The trails are independent and the probability of success (*P*) remains the same for each trial. The random variable *X* representing the number of success among *n* results of the trial then has a Binomial distribution, with parameters *n* and *P*. The range of *X* is {*1, 2, 3, ... , n*}. To calculate P(*X* = *k*) we need the number of combinations from *n* objects of which *k* are of type *s* and the remaining *n* - *k* are type *f*. An outcome of the experiment like:

$$\underbrace{s\ s\ s \ldots s\ s}_{k\ times}\ \underbrace{f\ f\ f \ldots f\ f}_{(n-k)\ times}$$

Because the trials are independent the probability of such an outcome can be given as product $P^k \cdot (1-P)^{n-k}$. The above combination of *s* and *f* successions is just one out of the $\binom{k}{n}$ possible combinations of getting *k* successes out of *n* total trials, each having a probability $P^k \cdot (1-P)^{n-k}$ (see example 2-2). Moreover, they correspond to mutually exclusive events, so the probability function of the random variable *X* representing the chance of getting *k* number of failures out of *n* total trials is written as:

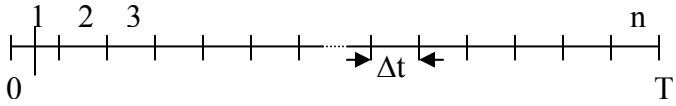
$$P(X = k) = \binom{k}{n} \cdot P^k \cdot (1-P)^{n-k} \qquad (9\text{-}47)$$

In which $\binom{k}{n} = \frac{n!}{k!(n-k)!}$ is the number of combinations of *n* things taken *k* at a time. The mean of *X* is $E[X] = n \cdot P$ and the variance is $\nu = n \cdot P \cdot (1-P)$. There are tables of Binomial distribution. For various values of n and p they give the values of $F(k) = P(X \leq k)$.

For example, the number of dam failures in any year is the sum of all the failures individual dams across the world. For purpose of illustration, assume that these failures are independent from one another; that is the failure of the dam is assumed for the purpose of data modeling in no way to affect potential failures of other dams; and, furthermore, that no common cause initiating event simultaneously affects failures of multiple dams. Let the probability of any one dam failing in a particular year be *P*. The probability that K dams fail in one year is then P^k. The probability that the remaining $n - k$ not failing is $(1-P)^{n-k}$. Thus, the probability of k dams out of a total of *n* dams failing is the product of *k* failing the complement $n - k$ not failing, or $P^k \cdot (1-P)^{n-k}$. Since there are many different combinations of how *k* dams can be chosen from a set of n total dams, the probability distribution function of the number of failures of *k* out of *n* dams in any given year is given as: $P(X=k) = \binom{k}{n} \cdot P^k \cdot (1-P)^{n-k}$.

Poisson distribution (random occurrence- number of times an event occur in a given interval)

Let the expected number of occurrences of a certain event *D* be proportional to the length of the (time) interval considered. For example, earthquake or flood in a period of years, breakdowns of systems, number of defects per square meter, number of fractures in abutment rock per square meter etc. Calling the intensity of occurrence λ, means that during time *T* (length *L*) the expected occurrence is $\lambda \cdot T$ (when considering length $\lambda \cdot L$). The time interval, say (*0, T*), (or the length (0, *L*)) can be divided into small non-overlapping intervals of equal length Δt (or Δl).



If *n* the number of intervals is big enough, the length of each interval will be small that during such a short time the event *D* can not happen more than once or the probability that *D* happens more than ones is so small that it can be neglected. As a further assumption consider that whatever happens in one interval is independent of what happens in other intervals. Now the whole thing can be modeled as Bernoulli-experiment[9]:

1. The experiment consists of $n = T/\Delta t$ independent trials;
2. In each trial (interval of length Δt) the event *D* occurs or does not occur;
3. For each trial the probability that *D* occurs is $P = \lambda$

If *X* is the number of times *D* occurs during the whole period (*0, T*) then *X* has a Binomial distribution with parameters $T/\Delta t\ (=n)$ and $\lambda \cdot \Delta t\ (=P)$. The expected value of $X, E[z] = n \cdot P = \frac{T}{\Delta t} \cdot (\lambda \cdot \Delta t) = \lambda \cdot T$, corresponding with the expected number of occurrences of *D* during (*0, T*). This result is independent of *n*. Let *n* go towards infinity, which means that Δt goes

[9] A Bernoulli experiment is a random experiment, the outcome of which can be classified in but one of two mutually exclusive and exhaustive ways, mainly, success or failure (e.g., female or male, life or death, non-defective or defective).

to zero, for $n \cdot \Delta t = T$ is constant. After some mathematical manipulations, the probability function of *X* changes to:

$$P(x) = P(X = x) = \frac{(\lambda \cdot T)x}{z!} \cdot e^{-\lambda \cdot T}, \qquad x = 0, 1, 2, ... \qquad (9\text{-}48)$$

From this follows that $E[x] = \lambda \cdot T$ and $\nu = \lambda \cdot T$.

Hartford and Gregory (2004) mention examples of application of Poisson distribution in surveying rock joints in dam abutment by taking core samples across lines set on rock outcrops, and the number of fractures intersected by the sampling line or boarding is recorded. Thus the probability of number of fractures per unit length can be estimated by Poisson distribution.

Exponential distribution (random occurrence- interval until occurrence)

The random variable *X* denotes the number of times an event *D* occurs during the time interval (*0, T*) (or length (*0, L*)). Suppose *X* has a Poisson distribution with parameter $\lambda \cdot t$, so $E[X] = \lambda \cdot t$.

Now consider the continuous random variable *Z*, representing the waiting-time until the first occurrence of the event *D*. If during (*0, t*) there are no occurrences of *D*, we wait longer than t for the first occurrence of *D*. So we get the equality:

$$P(X = 0) = P(Z > t)$$

from which follows:

$$P(X \geq t) = e^{-\lambda \cdot t}$$

This in turn gives the distribution function of *Z*:

$$F(t) = P(Z < t) = 1 - P(Z > t) = 1 - e^{-\lambda \cdot t}, \; t \geq 0$$

It is clear that $F(t) = 0$, for all values of $t < 0$, for a waiting-time of course can never be negative.

The density of *Z* follows as $\frac{d}{dt} F(t)$. Therefore, a random variable, *Z* has an exponential distribution with parameter $\lambda \; (\lambda > 0)$ if:

$$F(t) = 1 - e^{-\lambda \cdot t}, \;\; t \geq 0;$$
$$= 0 \qquad , \;\; t < 0 \qquad (9\text{-}49)$$

$$f(t) = \lambda \cdot e^{-\lambda \cdot t}, \; t > 0$$
$$= 0 \qquad , \; t < 0 \qquad (9\text{-}50)$$

$$E[Z] = \frac{1}{\lambda}$$

$$\nu = \left(\frac{1}{\lambda}\right)^2$$

Hartford and Gregory (2004) mention examples of application of exponential distribution in estimating spacing among adjacent rock fractures in dam abutments.

Extreme value and Weibull distribution (tails of probability distribution)

In probability theory and statistics, the generalized extreme value distribution (GEV) is a family of continuous probability distributions developed with in extreme value theory[10] to combine the Gumbel, Fréchet and Weibull families also known as type I, II and III extreme value distributions. By the extreme value theorem[11] the GEV distribution is the limit distribution of properly normalized maxima of a sequence of independent and identically distributed random variables. Because of this, the GEV distribution is used as an approximation to model the maxima of long (finite) sequence of random variables.

For example, the peak annual flow in a stream is the largest discharge to occur in any of the weeks in the year. Presume, for the purpose of discussion, the peak weekly flows are independent of one another (this is clearly not the case for daily flows, since flood hydrographs typically have a base time of more than one day). If the peak weekly discharge, say, can be modeled by a Normal pdf, what is the appropriate distributional form of the peak annual discharge? In essence, the question reduces to the distribution of the maximum value within samples of size 52 drawn from Normal population. Such problems are said to involve the distribution of extreme values. The classical reference on the statistics of extreme values is (Gumbel, 1958).

The generalized extreme value distribution has cumulative distribution function that is given as:

[10] Extreme value theory is a branch of statistics dealing with the extreme deviations from the median of probability distributions.

[11] In statistics, the Fisher–Tippet–Gnedenko theorem (also the Fisher–Tippet theorem or the extreme value theorem) is a general result in extreme value theory regarding asymptotic distribution of extreme order statistics. The maximum of a sample of independent and identically-distributed random variables after proper renormalization converges in distribution to one of 3 possible distributions, the Gumbel distribution, the Fréchet distribution, or the Weibull distribution. The role of extreme types theorem for maxima is similar to that of central limit theorem for averages.

$$F(x;\mu,\sigma,\xi)=\exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\} \tag{9-51}$$

For $1+\xi\left(\frac{x-\mu}{\sigma}\right)>0$, where $\mu\in\Re$ is the location parameter, $\sigma>0$ the scale parameter and $\xi\in\Re$ the shape parameter.

The density function is consequently:

$$f(x;\mu,\sigma,\xi)=\frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}-1}\cdot\exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\} \tag{9-52}$$

Again for $1+\xi\left(\frac{x-\mu}{\sigma}\right)>0$.

The shape of the extreme value distribution is sensitive to the shape of the tail of the parent distribution. The pdf of the largest values within a sample is sensitive to the shape of the upper tail of the parent pdf, and likewise the pdf of the smallest value is sensitive to the lower tail. For a normal distribution these are symmetric, but this is not the case for all parent distributions. Thus, the shape parameter ξ governs the tail behavior of the distribution. The sub-families defined by $\xi\to 0$, $\xi>0$ and $\xi<0$ correspond, respectively, to the Gumbel, Fréchet and Weibull families, whose descriptions, pdfs and cdfs are given below.

Gumbel or Type I extreme value distribution

The type-I limiting distribution arises for the largest variable from a parent distribution with an exponentially decaying upper tail, that is, an upper tail that falls off as:

$$F_X(x)=1-\exp(-g(x)) \tag{9-53}$$

In which *g(x)* is an increasing function of *x*. For example, the Normal, gamma and exponential distributions are all of this type. Gumbel showed that for large *n* this distribution approaches:

$$F_X(x;\alpha,\beta)=\exp\left\{-\exp\left[\frac{x-\alpha}{\beta}\right]\right\}=e^{-e^{\frac{-(x-\alpha)}{\beta}}} \tag{9-54}$$

$$f_X(x;\alpha,\beta)=\frac{1}{\sigma}\exp\left(\frac{-(x-\alpha)}{\beta}-\exp\left(-\frac{1}{\beta}(x-\alpha)\right)\right) \tag{9-55}$$

For $-\infty < x < \infty$, where α is location parameter and $\beta > 0$ is a scale parameter. The parameters of the distribution are typically estimated from observed data, the presumption being that the extreme variable arises from a large but perhaps unknown number *n* of, say, large stream discharges during the course of a year.

Fréchet or Type II extreme value distribution

Type II limiting distribution arises for the largest variable from a parent distribution with the upper tail that falls off as:

$$F_X(x;\mu,k) = 1 - \beta\left(\frac{1}{x}\right)^k \tag{9-56}$$

Gumbel showed that for large *n* this type II distribution approaches:

$$F_X(x;\mu,k) = e^{\left(\frac{-\mu}{x}\right)^k} \tag{9-57}$$

Weibull or Type III extreme value distribution

This limiting distribution arises for either the largest or smallest variable from apparent distribution with a limited tail, which is a tail that falls off as:

$$F_X(x) = 1 - c(w-x)^k \tag{9-58}$$

In which $x \le w$ and $k > 0$. In practice this distribution is mostly used to model smallest values, having a lower tail of the form $F_X(x) = c(w-x)^k$ for $x > \varepsilon$,where ε is the lower limit for *x*. For example the gamma distribution is of this form. Gumbel showed that for large *n* this type III approaches:

$$F_X(x;\alpha,\beta) = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] \tag{9-59}$$

$$f(x;\alpha,\beta) = \frac{\alpha}{\beta}\cdot\left(\frac{x}{\beta}\right)^{\alpha-1}\cdot e^{-\left(\frac{x}{\beta}\right)^\alpha} \tag{9-60}$$

For $x > 0$, β and α are positive parameters, β being a scale and α being a shape parameter. Weibull distribution is a flexible distribution, which can also be generalized to three parameters Weibull distribution with an inclusion of a location parameter *L* as in (9-62), enabling it to be adjusted to cover all stages of reliability (factor of safety) distributions. A shape parameter of, i.e. $\alpha = 1$ gives an exponential distribution and as shape parameter of 3.25 gives approximation to normal distribution. This distribution finds principal applications in situations involving wear, fatigue and fracture.

$$F_X(x;\alpha,\beta,L) = 1 - \exp\left[-\left(\frac{x-L}{\beta}\right)^{\alpha}\right];\ L \le x, \beta > 0, \alpha > 0 \tag{9-61}$$

$$f(x;\alpha,\beta,L) = \frac{\alpha}{\beta}\cdot\left(\frac{x-L}{\beta}\right)^{\alpha-1}\cdot e^{-\left(\frac{x-L}{\beta}\right)^{\alpha}};\ L \le x, \beta > 0, \alpha > 0 \tag{9-62}$$

Gamma distribution (waiting time to failure):

In probability theory and statistics, the gamma distribution is a two-parameter family of continuous probability distributions. It has a scale parameter λ and a shape parameter α. If α is an integer then the distribution represents the sum of α independent exponentially distributed random variables, each of which has a mean of λ.

The gamma distribution is frequently a probability model for waiting times; for instance, in life testing, the waiting time until failure is a random variable that is frequently modeled with a gamma distribution. The equation defining the probability density function of a gamma-distributed random variable *x* is:

$$f(x;\lambda,\alpha) = \frac{e^{\frac{-x}{\lambda}}\cdot x^{-1+\alpha}\cdot\lambda^{-\alpha}}{\Gamma[\alpha]} \tag{9-63}$$

Where Γ is the gamma function.

Beta distribution:

The random variable *X* is said to have a beta distribution if its pdf is given by:

$$f(x;\alpha,\beta) = \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\cdot\Gamma[\beta]}\cdot x^{\alpha-1}\cdot(1-x)^{\beta-1} = \frac{1}{\mathrm{B}[\alpha+\beta]}\cdot x^{\alpha-1}\cdot(1-x)^{\beta-1},\ 0 \le x \le 1 \tag{9-64}$$

Where Γ is the gamma function. The beta function, B, appears as a normalization constant to ensure that the total probability integrates to unity. The integrals needed to calculate the cumulative function cannot be done analytically except in special cases. When X_1 and X_2 have independent gamma distributions with equal scale parameters, the random variable $\frac{X_1}{X_1+X_2}$ a follows the beta distribution B with parameters [α, β], where α and β are the shape parameters of the gamma variables. Beta distribution is frequently used in Bayesian analysis.

9.1.3 Fitting distributions to data

In this research Crystal Ball software is used for doing the majority of distribution fitting work. For available historical or measured data Crystal Ball's distribution fitting feature can

substantially simplify the process of selecting a probability distribution. Not only is the process simplified, but the resulting distribution more accurately reflects the nature of your data than if the shape and parameters of the distribution were estimated.

In distribution fitting, Crystal Ball automatically matches the available data against each probability distribution. A mathematical fit is performed to determine the set of parameters for each distribution that best describe the characteristics of your data. The quality or goodness of each fit is judged using one of several standard goodness-of-fit tests. The distribution with the highest ranking fit is chosen to represent your data. It is possible to review the distributions sorted in order of their fit tests using the comparison chart. This chart shows the fitted distributions superimposed over your data so you can visually check the quality of the fits. Several chart preferences make it easier to pinpoint discrepancies in the fits. If desired, you can override the highest ranking probability distribution with another one of your choice.

In ranking the distributions, using Crystal Ball one can use any one of three standard goodness-of-fit tests:

- *Anderson-Darling*: this method closely resembles the Kolmogorov-Smirnov method, except that it weights the differences between the two distributions at their tails greater than at their mid-ranges. This weighting of the tails helps to correct the Kolmogorov-Smirnov method's tendency to overemphasize discrepancies in the central region.

- *Chi-Square*: this test is the oldest and most common of the goodness-of-fit tests. It gauges the general accuracy of the fit. The test breaks down the distribution into areas of equal probability and compares the data points within each area to the number of expected data points. The chi-square test in Crystal Ball does not use the associated *p*-value the way other statistical tests (e.g., t or F) do.

- *Kolmogorov-Smirnov*: the result of this test is essentially the largest vertical distance between the two cumulative distributions.

The goodness-of-fit test topic is adequately treated in many classical statistics and reliability engineering books, like in (D'Agostino et al. 1986, Benjamin et al. 1970, Crowder et al. 1991, and Lawless 1982).

9.2 Appendix to chapter four

Eq.(4-13) proof: For any set B, $P\{Y \in B\} = \int_B \ldots \int f_Y(y_1, y_2, \ldots, y_n) dy_1\, dy_2 \ldots, dy_n$, compare this with Eq.(4-2) and (9-32). From Eq.(9-38):

$$P\{Y \in B\} = \int_{h^{-1}(B)} \ldots \int f_Y(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots, dx_n = \sum_j \int_{h^{-1}(B)} \ldots \int f_X(x_1, x_2, \ldots, x_n) dx_1\, dx_2 \ldots, dx_n \tag{9-65}$$

The transformation from each $h_{1j}^{-1}(B)$ to the set B is a one-to-one transformation hence, applying the well known theorem in transformation of multiple integrals discussed in advanced calculus books, for example (Adams 2006), results in:

$$P\{Y \in B\} = \sum_j \int_B \ldots \int f_x\left(h_{1j}^{-1}(y_1, y_2, \ldots y_n), \ldots, h_{nj}^{-1}(y_1, y_2, \ldots y_n)\right) \cdot \left|J_j\right| dy_1, \ldots, dy_n \tag{9-66}$$

Since the set B is an arbitrary set and since the integrand gives a correct probability for any such set B in the range of $h(\boldsymbol{X})$, it follows that Eq.(9-66) is equal to $f_Y(Y_1, Y_2, \ldots, Y_n)$.

9.3 Appendix to chapter five

9.3.1 Tendaho Dam main features

1. Important elements:
   - Crest length 412m
   - Crest width 10m
   - Dam height above deepest foundation 53 m
   - Dam height from toe 47 m
   - Volume of fill material $1.37 \cdot 10^6$
   - Reservoir capacity $1.86 \cdot 10^9$
   - Water surface area at FRL 1,700 km^2
   - Top of parapet wall elevation 412.5 m a.m.s.l
   - Dam crest elevation 412 m a.m.s.l
   - Core crest elevation 411 m a.m.s.l
   - Maximum water level (MWL) 410.4 m a.m.s.l
   - Maximum retention level (FRL) 408 m a.m.s.l
   - Minimum drawdown level (MDDL) 396 m a.m.s.l
   - River bed level (OGL) 369 m a.m.s.l
   - Dam toe elevation 365 m a.m.s.l (after stripping 4m from OGL)
   - Core bottom elevation 359 m a.m.s.l
   - Spillway crest level 400 m a.m.s.l
   - Irrigation intake centerline elevation 391 m a.m.s.l
   - Fill volume $3.958 \cdot 10^6$ m^3
   - Upstream slope 3.5:1
   - Downstream slope 2.5:1 and 2:57:1

Table 9.5: Dam component elevations in relative terms from dam toe elevation.

	absolute elevation m a.m.s.l	relative elevation m a.m.s.l
Wave wall (parapet wall)	412.5	47.5
Dam crest	412	47
Core crest	411	46
MWL	410.4	45.4
FRL	408	43
MDDL	396	31
OGL	369	4
Dame toe	365	0
Core bottom elevation	359	-6

2. Spillway information:
   - Crest length 37.5 m
   - Type Ogee type
   - Crest level 400 m a.m.s.l
   - Design discharge 1900 m^3/s

3. Diversion Tunnel:
   - Design discharge — 31 m^3/s
   - Tunnel length — 247.36 m
   - Tunnel diameter- — 6 m
   - Approach channel length — 320.44 m
   - Out let channel length — 357.04 m

4. Intake tower
   - Discharging capacity — 78m^3/s
   - Irrigation command area — 60,000 ha (600 km^2)
   - Irrigation outlet diameter — 6.0 m
   - centerline elevation — 391 m a.m.s.l

5. Main canal
   - Length — 65.0 km
   - Discharging capacity — 78.0 m3/s
   - Bed width — 22. 65 m
   - Water depth — 2.55 m

6. Sugar factory
   - Factory capacity — 26,000 tones cane per day
   - Sugar cane production — $7.02 \cdot 10^6$ tones per year
   - Max. period of operation per year — 270 days

9.3.2 Type, form and sources of data collected on Tendaho Dam during the fieldwork

Data type	Form of data	unit	Amount of data	Data source	Remark
Material laboratory test results and former material study reports					
Material test results from test conducted in Ethiopian Construction Design share company (CDSco.) material laboratory. Samples were collected from stockpiles of the dam shell and core materials and some undisturbed samples were takes as required.	digital and hard copy of test results				tests conducted includes triaxial CU, permeability, compaction, consolidation, Atterburg limits, grain size analysis
Secondary data on material of construction for Tendaho Dam	hard copy	pages	66	WWDSE	this document contains values of basic engineering characteristics of shell and core materials used in original design
Field laboratory test results from construction phase					
Field density test reports from construction site	hard copy	pages	202		this contains field density and moisture content measurements taken at different layers of shell and core zones construction
Meteorological data					
15 years daily rainfall amount, maximum and minimum temperature data for 15 stations in Awash basin	digital data			NAMSA	mean monthly rainfall and temperature records were available for longer series.
15 years daily rainfall intensity, sunshine hours, evaporation, wind speed, RH data for 18 stations in Awash basin	hard copy				mean monthly records were available for longer series
Hydrological and sedementlogical data					
52 stations Awash River instantaneous daily flow data	digital data			MoWR	
Various years Awash River sediment flow data	digital data			MoWR	
Awash basin GIS data					
Awash basin DEM, basin boundary, tributaries and Awash River, soils and lakes	digital				ArcGIS data

Two years 1990 and 2000 satellite image of the Awash Basin	digital				
Design documents					
Tendaho dam and appurtenant works design final report	hard copy	pages	157	WWDSE	
Tendaho dam and appurtenant works drawings	hard copy	pages	47	WWDSE	
Geological, seismological and geotechnical data					
Geo-technical exploratory core drilling work of Tendaho Dam final report	hard copy	pages	78	GSE	By Geological Survey of Ethiopia (GSE)
Tendaho Dam project geological investigation final report	hard copy	pages	78	WWDSE	Contains information on geological setup of the project, geotechnical investigations made, geological evaluation of dam and appurtenances sites
Master plan and other study documents					
Awash River basin flood control and watershed management study project Base line report 5: flood control drainage and irrigation	hard copy	pages	160	MoWR	By Halcrow
Awash River basin flood control and watershed management study project phase 2 summary reports, volume 2, annex B: hydrological studies, annex C: environmental assessment.	hard copy	pages	120	MoWR	By Halcrow
Awash River basin flood control and watershed management study project base line report 6: sediment source and control measures	hard copy	pages	120	MoWR	By Halcrow
Awash River basin flood control and watershed management study project base line report 4: river morphology of Awash and tributaries	hard copy	pages	172	MoWR	By Halcrow
Feasibility study of the lower Awash valley final report-Part II	hard copy	pages	142	MoWR	By Sir Alexander Gibb and Partners London
Feasibility study of the lower Awash valley final report-Part I Annex 2: climate and	hard copy	pages	142	MoWR	By Sir Alexander Gibb and Partners London

hydrology					
Feasibility study, proposals and estimates of cotton development on the area of 60,000 ha in the lower Awash valley: explanatory notes	hard copy	pages	269	MoWR	By Uzbek State Design and Research Institute Department of Land Reclamation and Water Economy
Feasibility Study, proposals and estimates of Cotton development on the area of 60,000 ha in the lower Awash valley: Book 1 Summary	hard copy	pages	269	MoWR	By Uzbek State Design and Research Institute Department of Land Reclamation and Water Economy
Master plan for the development of surface water resources in the Awash basin, final report, Volume I: executive summary	hard copy	pages	44	MoWR	By Halcrow
Master plan for the development of surface water resources in the Awash basin, final report, Volume II: main report	hard copy	pages	199	MoWR	By Halcrow
Master plan for the development of surface water resources in the Awash basin, final report, Volume IV: climate and hydrology	hard copy	pages	199	MoWR	By Halcrow
Master plan for the development of surface water resources in the Awash basin, final report, Volume V, Annex B: geological assessment of dam; Annex C: assessment of groundwater potential	hard copy	pages	99	MoWR	By Halcrow
Master plan for the development of surface water resources in the Awash basin, final report, Volume VIII, Annex i: dams and hydropower; Annex j: irrigation and drainage	hard copy	pages	199	MoWR	By Halcrow
Reference documents					
Manual of soil laboratory testing, Volume 2: permeability, shear strength and compressibility tests	hard copy	450			CDSco
Manual of soil laboratory testing, Volume 3: effective stress tests (ELE international Limited)	hard copy	494			CDSco
Soil Properties: testing, measurement and evaluation	hard copy	315			CDSco
Experimental soil mechanics	hard copy	593			AMU

9.3.3 Selected Tendaho dam core and shell material properties

Quarry location	dam zone	grain size distribution			shear strength		MDD (gm/cm^3)	OMC (%)	permeability (cm/s)
		gravel	sand	fines	c' (kN/m^2)	ϕ' (degrees)			
Area1	shell		1.2	98.8					
Area1	shell		55.06						
		40.7-42.2	-58.4	0.4-2.6					
Area1	shell	75.1	23.6	1.3			1.33		5.21E-03
Area1	shell	9.7	90.1	0.2			1.77		
Area1	shell		97.7	2.3					
Area1	shell	64	31.4	4.6	4	44			3.16E-03
Area1	shell	47.7	52.3		12	41			2.30E-03
Area1	shell	56.1	43.9		4	44			4.60E-03
Area1	shell	62	37	1					
Area1	shell	57	42	1	3	35	1.832	21.25	1.28E-02
Area3	shell		99.8	0.2	19	38			
Area3	shell	27	73		5	46			2.30E-02
Area3	shell	31.3	68.7		14	45			9.50E-03
Area3	shell	37.6	62.4		8	44			1.50E-03
Area3	shell	51.1	48.9		2	48			2.80E-03
Area3	shell	1.7	98.1	0.2					
Area3	shell	18.2	81.5	0.3	12	40			
Area3	shell	35.6	64.4						
Area3	shell				2.5	42			1.01E-03
Area3	shell				4	42			6.31E-02
Area7	shell	41.4	58.6				2.31		
Area7	shell	42.8	55.6	1.8			2.26		
Area7	shell	45.8	53.1	1.1			1.9		
Area7	shell	41.9	56.9	1.2			2		
Area7	shell	51.1	45.1	3.8					
Area7	shell	46.3	53.1	0.6					
Area7	shell				17	40			
Area7	shell				12.4	42			
Area7	shell				16	40	2.087	3	
Area7	shell	70	29	1	4	41	2.168	10.5	4.29E-03
Area7	shell	25	68	7			2.27	8	1.06E-02
Area11	shell	49	51		18	40			
Area11	shell	45.7	54.3		31	40			
Area11	shell	64.5	35.5		4	43	1.33	30.79	5.92E-04
Area11	shell	42.7	57.3		2	45	1.77	15.29	2.84E-03
Area11	shell	69.4	30.6		4	45	1.78	14.25	5.54E-03
Area11	shell	50.1	49.9		5	46	2.255	7.25	4.09E-03
Area11+11A	shell	65	30	5	6	38	1.934	7	3.18E-03
Area 11	shell	0	0.85	99.15					
Area 11A	shell	36.19	59.77	4.04					
Area 2	core		2.3	97.7					
Area 2	core		1	99			1.4	30.43	3.70E-09

Area 2	core		17	83	37	35	1.53	22.25	1.07E-05
Area 2	core		1.1	98.9	2	21	1.25	31.5	1.45E-05
Area 2	core	14	85.5	0.5	66	21			1.50E-05
Area 5	core		33.6	66.4	24	20			
Area 5	core		27.2	72.8	11	18	1.7	19.3	3.7E-09
Area 5	core		14.4	85.6	33	15	1.82	12.4	
Area 5	core		4	96	64	31	1.43	25.25	
Area 5	core		9	91	64	18	1.38	33	3.9E-09
Area 5	core		3.6	96.4	64	31	1.43	28.5	1.44E-06
Area 5 (blended)	core		3	97	4	33	1.623	15.18	4.18E-06
Area 5	core	5			6	31	1.537	23.63	1.37E-06
Area 5	core				95	17	1.757	18	1.23E-06
Area 5	core	18	20	62	20	19	1.72	18.3	2.20E-08
Area 6	core	1.4	94.5	4.1					
Area 6	core		93.4	6.6			1.83	16.4	2.48E-08
Area 6	core		59	41	32	30			1.07E-07
Area 9	core		4.3	95.7			1.4	34	2.33E-09
Area 9	core		2.9	97.1	26	4			
Area 9	core		12.1	87.9	21	28.6	1.4	29.6	1.45E-07
Area 9	core		10.3	89.7			1.49	27.5	
Area 9	core		6.6	93.4			1.35	33.5	1.23E-08

9.3.4 Historical mean monthly, mean annual and annual maximum flow of Awash River at Tendaho

	Mean monthly flow (m^3/s)												Mean annual flow (m^3/s)	Annual max. flood (m^3/s)
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec		
1965	34.6	29.3	40.9	77.9	39.5	21.1	30.1	123.2	94.5	69.3	59.4	43.2	55	435.0
1966	38.7	63.3	54.9	64.2	41.4	32.9	37.5	102.5	113.4	88.9	48.6	33.3	60	321.0
1967	25.5	40.5	25.5	58.6	80.6	22.9	76.4	240.1	80.6	124.2	131.2	62.6	81	570.0
1968	40.4	96.2	65.2	117.0	55.4	45.8	208.3	215.9	147.8	75.6	49.8	41.3	97	745.0
1969	99.4	90.4	84.3	81.5	74.9	77.4	50.5	213.9	128.3	90.6	47.0	28.8	89	519.0
1970	52.1	35.1	139.6	45.2	37.7	25.9	133.8	388.3	197.1	111.8	47.2	27.1	103	872.0
1971	23.8	17.8	15.5	23.0	30.1	19.0	32.5	170.2	186.3	105.4	60.6	31.8	60	486.0
1972	32.5	108.3	45.6	77.2	66.5	48.0	61.4	80.7	75.1	55.3	29.3	19.5	58	827.0
1973	17.5	13.8	6.5	5.3	12.3	3.6	69.0	534.1	93.7	69.1	37.9	21.2	74	1420.0
1974	14.6	11.2	284.4	36.5	22.7	22.9	184.0	379.0	219.9	92.7	49.8	28.7	112	1687.2
1975	33.4	40.2	25.3	97.9	40.1	18.9	107.9	264.9	484.1	224.2	75.8	43.0	121	1100.0
1976	53.4	51.5	53.2	67.8	67.0	53.1	42.7	117.7	85.5	68.9	62.1	42.4	64	237.0
1977	29.5	29.9	23.0	91.2	69.9	21.8	46.7	198.3	112.5	254.7	139.7	64.7	90	1550.0
1978	47.5	175.9	106.3	72.8	68.5	26.2	175.9	147.1	77.9	63.2	55.4	47.1	89	1306.0
1979	105.2	49.0	211.3	76.7	55.6	12.8	61.1	268.2	133.2	135.9	60.9	52.7	102	964.0
1980	30.9	35.8	22.5	30.8	15.7	7.1	42.8	246.2	110.1	84.6	34.1	17.5	57	1205.0
1981	6.3	4.1	193.2	104.4	42.7	20.9	50.6	132.2	128.2	83.1	61.2	41.9	72	462.0
1982	33.0	26.9	64.9	79.4	62.3	15.7	19.1	49.6	53.0	203.6	212.8	62.7	74	417.0
1983	41.8	58.2	42.3	102.6	70.9	50.9	37.0	72.6	81.2	87.3	60.2	46.6	63	248.0
1984	39.8	35.1	25.3	16.8	49.0	15.8	21.5	21.6	44.9	27.4	17.2	40.1	30	292.0
1985	21.3	19.4	20.7	100.4	70.4	30.5	46.5	95.4	203.1	97.2	51.4	40.3	66	442.0
1986	30.9	82.6	81.7	109.7	51.1	46.4	80.9	128.7	109.3	71.1	52.0	32.4	73	352.0
1987	19.8	13.6	200.6	166.3	50.7	57.7	28.9	105.1	57.9	42.9	26.5	25.9	66	594.7
1988	0.8	26.1	7.7	151.4	48.5	30.9	72.1	179.2	172.3	53.3	53.3	28.5	69	580.8
1989	24.8	16.2	22.0	116.9	12.4	17.0	24.8	51.6	53.5	65.5	43.2	31.7	40	231.4
1990	19.4	152.7	72.6	185.2	81.3	33.8	32.3	37.0	43.6	69.9	46.6	29.3	67	513.3
1991	17.3	19.5	61.3	38.6	30.6	21.7	29.0	135.5	68.5	99.6	57.1	25.6	50	515.0
1992	28.8	50.6	38.2	41.7	32.9	16.1	40.8	197.0	133.1	232.9	94.5	37.6	79	466.9

1993	50.6	96.5	56.7	104.2	36.8	91.4	91.0	101.2	162.2	122.6	109.6	54.1	90	531.5
1994	29.5	17.7	43.5	74.5	48.5	30.9	72.1	180.3	133.1	97.1	60.2	85.0	73	-
1995	34.0	43.4	64.2	74.5	48.5	30.9	54.9	180.3	133.1	97.1	60.2	45.5	72	-
1996	34.0	7.9	23.4	12.2	48.5	17.2	7.3	200.6	87.7	42.5	60.2	45.5	49	339.7
1997	12.4	13.5	44.7	53.8	21.0	22.4	34.8	71.6	26.1	119.3	108.7	22.9	46	337.0
1998	32.8	26.4	66.2	12.5	1.3	10.9	164.1	431.3	327.5	105.2	20.3	79.6	106	600.0
1999	44.7	32.7	50.0	24.4	12.9	9.8	131.7	186.4	186.3	84.2	62.1	67.1	74	390.0
2000	38.0	24.6	18.0	9.8	25.2	7.2	71.7	180.3	133.1	97.1	60.2	45.5	59	447.0
2001	17.9	13.3	63.7	62.5	87.2	66.2	112.0	215.3	141.2	68.0	8.3	45.5	75	445.0
2002	15.1	16.2	71.7	196.9	86.3	63.5	51.5	132.0	25.0	12.3	0.9	121.6	66	430.0
2003	23.8	17.8	15.5	27.1	1.7	30.9	142.0	149.5	67.0	9.9	53.3	17.7	46	406.0
Mean	33.2	43.7	65.4	74.1	46.1	30.7	71.2	177.5	125.9	95.0	60.7	43.0	72.2	
Max.	105.2	175.9	284.4	196.9	87.2	91.4	208.3	534.1	484.1	254.7	212.8	121.6	121.3	
Min.	0.8	4.1	6.5	5.3	1.3	3.6	7.3	21.6	25.0	9.9	0.9	17.5	29.5	
stdev	20.3	39.0	61.5	47.4	23.5	20.2	50.3	109.4	84.4	54.2	38.1	20.7	20.3	
CV	0.6	0.9	0.9	0.6	0.5	0.7	0.7	0.6	0.7	0.6	0.6	0.5	0.3	

9.3.5 Monthly flow and sediment data of Awash River at Dubti (based on measurements made by Sogreah between 1962-1964, source (Sogreah, 1965))

Awash Monthly flow (10^6 m^3)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
1962	105.6	102.6	123.8	103.1	81.1	44.3	77.8	520.4	340.8	245.0	127.7	92.3	1964.5
1963	117.8	49.5	57.1	238.8	320.0	130.9	161.7	397.6	571.2	213.7	125.4	109.8	2493.5
1964	129.5	107.0	93.3	183.5	122.0	86.5	593.0	1158.0	685.2	313.1			

Awash at Dubti suspended sediment load (10^3 metric tons)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
1962	84.5	82.1	417.6	240.6	126.3	35.7	877.7	8619.8	3783.0	538.9	203.2	103.2	15112.4
1963	160.1	26.4	100.5	5123.2	5207.5	177.8	1064.3	4791.0	6813.5	684.1	2875.5	247.4	27271.4
1964	357.6	266.0	178.8	2598.6	609.1	160.0	24644.0	32004.0	10561.0	1734.0			

9.3.6 Monthly flow and sediment data of Awash River at Dubti (based on measurements made by MoWR department of hydrology over the period 1985-1987, source MoWR)

Year	Date	flow (m^3/s)	sediment concentration (mg/l)	flow (10^6 m^3/day)	$Q_{sediment}$ (tons/day)
1985	Aug.28	244	34952	21.1	736844
	Sep.1	439	46058	37.9	1746962
1986	May. 14	28	979	2.42	2368
	May.26	27	2407	2.33	5615
	Aug.27	83	8174	7.17	58617
1987	Apr.9	250	26486	21.6	572098
	Apr.11	757	15213	65.4	995003
	Jun.25	44	1407	38.0	5349
Mean		234.000	16960	24.49	515357

9.3.7 Monthly sediment load of Awash River generated using flow-sediment load relationship derived based Sogreah data

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual sediment load (10^6 tones)
1962	258	242	364	245	145	39	132	8315	3308	1612	390	193	15
1963	327	50	68	1524	2884	411	653	4627	10186	1198	375	281	23
1964	402	266	197	859	353	167	11053	47478	15141	2750	514	4364	84
1965	194	108	279	1057	259	62	143	3082	1612	881	586	315	9
1966	247	580	530	695	287	162	231	2067	2396	1516	379	178	9
1967	100	219	100	568	1224	74	1088	13182	1139	3140	3292	705	25
1968	272	1441	770	2565	541	333	9683	10464	4268	1064	399	285	32
1969	1934	1259	1349	1167	1043	1043	442	10254	3136	1580	351	131	24
1970	473	160	4048	323	234	96	3693	37578	7986	2496	356	114	58
1971	86	36	34	75	143	49	170	6233	7069	2194	613	161	17
1972	169	1866	354	1039	805	369	676	1228	977	539	126	56	8
1973	44	21	5	3	20	1	872	75227	1583	875	220	66	79
1974	30	13	19074	203	78	73	7383	35637	10141	1660	399	129	75
1975	180	216	98	1741	267	48	2312	16330	56553	11359	996	311	90
1976	500	369	495	782	820	460	306	2794	1296	870	645	302	10
1977	137	113	80	1492	898	66	372	8690	2356	15001	3777	758	34
1978	387	5364	2234	913	860	98	6701	4535	1059	720	503	380	24
1979	2185	332	9986	1023	546	21	670	16782	3402	3817	620	485	40
1980	152	167	76	141	35	6	309	13924	2248	1361	175	44	19
1981	5	1	8215	2001	307	60	444	3595	3129	1309	625	295	20
1982	175	90	764	1103	699	32	53	425	457	9207	9436	708	23
1983	293	483	301	1927	924	419	224	974	1157	1456	603	372	9
1984	264	160	98	38	414	33	69	69	319	117	39	267	2
1985	67	44	63	1837	912	138	370	1766	8529	1839	428	271	16
1986	152	1034	1262	2231	453	343	1235	3391	2211	932	439	169	14
1987	58	20	8918	5517	446	551	131	2184	554	310	101	103	19
1988	0	84	7	4498	406	141	959	6975	5963	498	463	127	20

1989	94	30	73	2562	21	38	94	464	466	778	294	160	5
1990	55	3946	976	6978	1247	172	167	224	299	898	346	136	15
1991	43	44	673	230	148	65	132	3796	800	1940	538	101	9
1992	130	356	241	271	174	34	278	8570	3398	12340	1610	233	28
1993	444	1450	569	1994	222	1498	1594	2008	5227	3049	2226	514	21
1994	137	36	319	959	406	141	959	7070	3398	1835	603	1373	17
1995	187	255	746	959	406	141	530	7070	3398	1835	603	353	16
1996	187	6	83	19	406	40	6	8919	1370	304	603	353	12
1997	21	20	339	473	65	70	196	945	98	2877	2186	79	7
1998	172	86	797	20	0	14	5755	47219	24134	2185	57	1190	82
1999	339	138	434	84	23	12	3565	7595	7064	1345	646	822	22
2000	238	74	47	11	97	6	948	7070	3398	1835	603	353	15
2001	46	19	732	654	1453	743	2504	10403	3864	845	8	353	22
2002	32	30	949	7970	1421	677	462	3582	89	21	0	2997	18
2003	86	36	34	106	0	141	4204	4697	762	13	463	45	11
Mean	269	506	1590	1401	526	216	1709	10891	5141	2438	896	491	26
Max.	2185	5364	19074	7970	2884	1498	11053	75227	56553	15001	9436	4364	90
Min.	0	1	5	3	0	1	6	69	89	13	0	44	2
stdev	427	1051	3616	1800	552	306	2674	15482	9323	3316	1568	792	23
CV	2	2	2	1	1	1	2	1	2	1	2	2	1

9.3.8 Mean annual sediment load of the Awash River at the Tendaho dam site generated based on two different rating equations.

Year	Mean annual flow (m^3/s)	Sediment load* (10^6 tones/year)	Sediment load** (10^6 tones/year)	Sediment volume*** (10^6 m^3/ year)
1962	63	15	8	11
1963	80	23	13	16
1964	129	84	33	58
1965	56	9	6	6
1966	61	9	7	6
1967	82	25	13	17
1968	98	32	19	22
1969	90	24	16	16
1970	106	58	22	40
1971	61	17	7	12
1972	59	8	7	6
1973	76	79	11	54
1974	115	75	27	52
1975	123	90	31	62
1976	65	10	8	7
1977	92	34	17	23
1978	89	24	16	16
1979	104	40	22	27
1980	58	19	7	13
1981	74	20	11	14
1982	75	23	11	16
1983	63	9	8	6
1984	30	2	2	1
1985	67	16	9	11
1986	74	14	11	10
1987	68	19	9	13
1988	70	20	10	14
1989	41	5	3	3
1990	67	15	9	11
1991	51	9	5	6
1992	80	28	13	19
1993	91	21	16	14
1994	74	17	11	12
1995	73	16	11	11
1996	50	12	5	8
1997	47	7	4	5
1998	109	82	24	56

1999	76	22	11	15
2000	60	15	7	10
2001	77	22	12	15
2002	67	18	9	13
2003	47	11	4	7
Annual average suspended sedimentation load		26	12	18
The bed load is taken to be 5%**** of the suspended load		1.30	0.60	0.90
Total sediment (suspended + bed load) transported to the reservoir/ year		27.38	12.61	18.88
Total deposited in the reservoir/year at a trap efficiency of 90%		24.64	11.35	16.99

* based on (Sogreah 1965) sediment rating equation

** based on sediment rating equation established using the eight recent data points

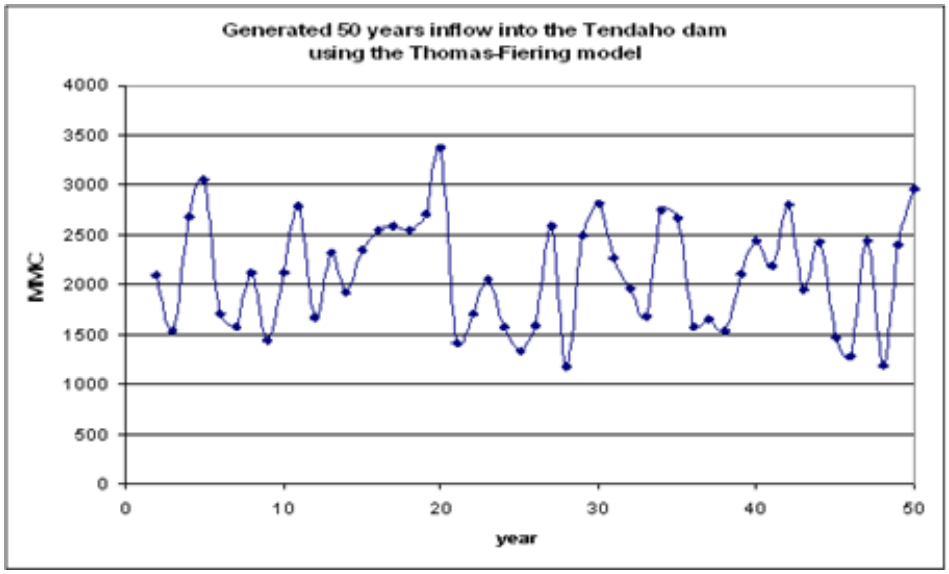
*** Based on Sogreah sediment rating equation and assuming a sediment density of 1.45 ton/m^3

**** Note: Halcrow (1989) mentions that at Tendaho site the bed load is 5% of the suspended load. They indicate in the Lower Awash Valley the sediment load is predominantly silt with fine sand and is dominantly in suspension.

9.3.9 Tendaho Dam elevation storage-area relationship at year 0, 25 and 50

Water surface elevation (m a.m.s.l)	Year zero (10^6 m^3)	After 25 years (10^6 m^3)	After 50 years (10^6 m^3)
384	21	0.0	0
390	125	28.8	0
391	149	36.5	0
392	186	55.1	0
393	237	87.5	1.9
394	297	126.8	9.1
395	366	174.3	22.4
396	444	228.7	41.1
397	528	288.7	63.4
398	617	353.4	88.7
399	713	422.4	116.8
400	815	497.5	149.4
401	924	578.9	187.1
402	1040	665.8	229.1
403	1161	758.3	276
404	1289	856.8	328.3
405	1422	959.8	385.4
406	1561	1070.1	450.7
407	1708	1190.2	528.7
408	1860	1329.6	604.9
409	2017	1464	642
410	2181	1600	717
411	2351	1742	795
412	2528	1891	877

9.3.10 Generated long term (50 years) flows for Tendaho Dam (after WWDSE, 2005)



9.3.11 Historic wind speed records used in Tendaho Dam wave height analysis

Year	Maximum daily above ground wind speed (km/hr)
1969	32
1970	51
1971	52
1972	14
1973	47
1974	16
1975	16
1976	15
1977	44
1978	38
1979	33
1980	16
1981	56
1982	52
1983	41
1984	15

Year	Maximum daily above ground wind speed (km/hr)
1985	23
1986	13
1987	11
1988	12
1989	12
1990	13
1991	13
1992	11
1993	12
1994	10
1995	11
1996	10
1997	23
1998	23
1999	10
2000	40

9.4 Appendix to chapter six

In this section to demonstrate the applications of FOSM and SOSM methods some sample calculations are provided.

9.4.1 SOSM approximation for moments and uncertainty of shear strength (τ)

Effective shear strength at failure for a defined slice of base length l_n ($l_n = b_n/cos\alpha$) that has a unit width along the dam axis direction can be written as in Eq.(6-1). Replacing relevant geometric values or the sample slice in Eq.(6-1) (l_n = 3.28 m, h_n = 17.2 m, h_w = 22.8 m, α_n (degrees) = 6.16, b_n (m) = 3.26, γ_w (kN/m^3) = 9.81) and c' (kN/m^2) = 0, yields the expression for shear strength τ (kN/m) at the bottom of the specific sample slice of base length l_n and having unit width along dam axis as given in Eq.(6-18).Geometric values for the sample slice are given in Table 6.1.

If parameters γ, ϕ' and u_n are taken to be random variables then Eq.(6-18) is a function of random variable (FRV) with $n = 4$ random components $\boldsymbol{X} = (u_n, \phi', \gamma, c')$. This implies that τ is a random variable and will have a cdf $F_\tau(\tau)$ describing its randomness. The problem of determining $F_\tau(\tau)$ is a distribution problem, which can be estimated using FOSM and SOSM methods.

General relations for SOSM approximations of the mean value (μ_Y) and variance (ν_Y) of a generic multivariate FRV $Y = h(\boldsymbol{X})$, where X = (X_1, X_2, … , X_n) is continuous random vector, is given by Eq.(3-14) and Eq.(3-21), respectively. Replacing τ for h and u_n, ϕ' , γ and c' in place of X_1, X_2, X_3 and X_4 in Eq.(3-14) and Eq.(3-21) yields the expression for the SOSM approximation for the mean and variance of τ (see section 3.2.4 for derivation of Eq.(3-14) and Eq.(3-21)):

$$\mu\tau_{SOSM} = \tau\left(\mu u_n, \mu_{\phi'}, \mu_\gamma, \mu_{c'}\right) + \frac{1}{2}\cdot\left(\nu_{Un}\cdot\left[\frac{\partial^2\tau}{\partial u_n^{\,2}}\right]_\mu + \nu_{\phi'}\cdot\left[\frac{\partial^2\tau}{\partial\phi'^2}\right]_\mu + \nu_\gamma\cdot\left[\frac{\partial^2\tau}{\partial\gamma^2}\right]_\mu + \nu_{c'}\cdot\left[\frac{\partial^2\tau}{\partial c'^2}\right]_\mu\right) \qquad (9\text{-}67)$$

$$\nu\tau_{SOSM} = \left(\nu_U\cdot\left[\frac{\partial\tau}{\partial u_n}\right]_\mu^2 + \nu_{\phi'}\cdot\left[\frac{\partial\tau}{\partial\phi'}\right]_\mu^2 + \nu_\gamma\cdot\left[\frac{\partial\tau}{\partial\gamma}\right]_\mu^2 + \nu_{c'}\cdot\left[\frac{\partial\tau}{\partial c'}\right]_\mu^2\right)$$

$$-\frac{1}{4}\cdot\left(\nu u_n\cdot\left[\frac{\partial^2\tau}{\partial u_n^{\,2}}\right]_\mu + \nu_{\phi'}\cdot\left[\frac{\partial^2\tau}{\partial\phi'^2}\right]_\mu + \nu_\gamma\cdot\left[\frac{\partial^2\tau}{\partial\gamma^2}\right]_\mu + \nu_{c'}\cdot\left[\frac{\partial^2\tau}{\partial c'^2}\right]_\mu\right)^2 \qquad (9\text{-}68)$$

$[..]_{\mu_i}$ means evaluate partial derivatives at mean of X_i. $[..]_\mu^2$ means evaluate the partial derivative of τ at mean of X_i.and square it.

For known best-fit distributions of the random variables (u_n, ϕ' , γ, c'), their moments (mean and variance) can be calculated using Eqs.(9-25) and (9-26) (see section ii). The best-fit distributions of the four random variables and their respective parameters are given in Table 6.1 and equations of commonly used pdfs and cdfs in engineering are provided in appendix 9.1.2. However, for FOSM and SOSM methods it is not mandatory to have best-fit distributions for the calculation of moments. An alternative way for getting the moments (mean and variance) of random variables, such as the (u_n, ϕ' ,γ, c') in this case, is to calculate the moments directly from observed data using the equations provided in section i.

Therefore, from the pdfs of respective random variable or equally from their observed data series the following first and second moments of the four random variables are determined:

Table 9.6: Moments of random variables in shear strength-stress equations (sample slice).

Parameter (unit)	mean	variance
u_n (kN/m^2)	$\mu u_n = 359.39$	$\nu u_n = 1021.83$
ϕ' (degrees)	$\mu_{\phi'} = 42.13$	$\nu_{\phi'} = 9$
γ (kN/m^3)	$\mu_\gamma = 17.8$	$\nu_\gamma = 7.08$
c' (kN/m^2)	$\mu_{c'} = 5.4$	$\nu_{c'} = 57.95$

In practice the effect of *c'* is only significant in cohesive materials. As a result, it could safely be neglected in calculations for slices residing in shell zone of dams.

The first and second partial derivatives of τ with respect to the different random variables can be calculated from Eq.(6-18) and is given as:

$$\left[\frac{\partial \tau}{\partial u_n}\right] = -3.28 \cdot tan\,\phi' \qquad \left[\frac{\partial^2 \tau}{\partial u_n^{\,2}}\right] = 0$$

$$\left[\frac{\partial \tau}{\partial \phi'}\right] = (724.037 - 3.28 \cdot U + 55.7 \cdot \gamma) \cdot \sec^2 \phi' \qquad \left[\frac{\partial^2 \tau}{\partial \phi'^2}\right] = 0.0006(724.037 - 3.28 \cdot U + 55.7 \cdot \gamma) \cdot \sec^2 \phi' \cdot \tan \phi'$$

$$\left[\frac{\partial \tau}{\partial \gamma}\right] = 55.7 \cdot \tan \phi' \qquad \left[\frac{\partial^2 \tau}{\partial \gamma^2}\right] = 0$$

$$\left[\frac{\partial \tau}{\partial c'}\right] = 3.28 \qquad \left[\frac{\partial^2 \tau}{\partial c'^2}\right] = 0$$

Evaluating the above derivatives at the respective means of the random variables; i.e. taking $u_n = \mu u_n, \phi' = \mu_{\phi'}, \gamma = \mu_\gamma$ and $c' = \mu_{c'}$ (see Table 9.6)yields:

$$\left[\frac{\partial \tau}{\partial u_n}\right]_\mu = -3.28 \cdot tan\,\phi' = -2.97 \qquad \left[\frac{\partial^2 \tau}{\partial u_n{}^2}\right]_\mu = 0$$

$$\left[\frac{\partial \tau}{\partial \phi'}\right]_\mu = \frac{\pi}{1800}\left(724.037 - 3.28 \cdot U + 55.7 \cdot \gamma\right) \cdot sec^2\,\phi' = 17.08 \qquad \left[\frac{\partial^2 \tau}{\partial \phi'^2}\right]_\mu = 0.0006(724.037 - 3.28 \cdot U + 55.7 \cdot \gamma) \cdot \sec^2 \phi' \cdot \tan \phi' = 0.539$$

$$\left[\frac{\partial \tau}{\partial \gamma}\right]_\mu = 55.7 \cdot \tan \phi = 50.42 \qquad \left[\frac{\partial^2 \tau}{\partial \gamma^2}\right]_\mu = 0$$

$$\left[\frac{\partial \tau}{\partial c'}\right] = 3.28 \qquad \left[\frac{\partial^2 \tau}{\partial c'^2}\right] = 0$$

Evaluating τ, Eq.(6-18) at the mean values of the random variables gives:

$$\begin{aligned}\tau\left(\mu u_n, \mu_{\phi'}, \mu_\gamma, \mu_{c'}\right) &= 3.28 \cdot \mu_{c'} + \left(724.037 - 3.28 \cdot \mu u_n + 55.695 \cdot \mu_\gamma\right) \cdot tan\,\mu_{\phi'} \\ &= 3.28 \cdot 5.4 + \left(724.0377 - 3.28 \cdot 359.39 + 55.695 \cdot 17.8\right) \cdot tan\,42.13 \\ &= 503 kN/m\end{aligned}$$

Substituting the above results in Eq.(9-67) yields the SOSM approximation for the mean of τ:

$$\mu\tau_{SOSM} = 503 + \frac{1}{2} \cdot \left(1021.83 \cdot 0 + \ 9 \cdot 0.539 + 17.8 \cdot 0 + 57.95 \cdot 0\right) = 508\,kN/m$$

Similarly, substituting in Eq. (9-68) yields the SOSM approximation for variance of τ:

$$\begin{aligned}\nu\tau_{SOSM} &= \left(1021.83 \cdot \left(-2.97\right)^2 + 9 \cdot \left(17.08\right)^2 + 7.08 \cdot \left(50.42\right)^2 + 57.95 \cdot \left(3.28\right)^2\right) \\ &\quad - \frac{1}{4} \cdot \left(1021.83 \cdot 0 + 9 \cdot 5.4 + 7.08 \cdot 0 + 57.95 \cdot 0\right)^2 \\ &= 29{,}673 (kN/m)^2\end{aligned}$$

Consequently, the standard deviation for τ will be:

$$\sigma\tau_{SOSM} = \sqrt{29{,}673} = 173\,kN/m$$

Note that there are minor rounding errors when doing the calculation using conventional hand calculators than Mathematica 5.2. Mathematica does calculations with over 100-digit precision.

9.4.2 FOSM approximation for moments and uncertainty of shear strength (τ)

General relations for FOSM approximations of the mean value (μ_Y) and variance (v_Y) of a multivariate FRV $Y = h(\boldsymbol{X})$, where $X = (X_1, X_2, \ldots, X_n)$ is continuous random vector, is given by Eqs.(3-22) and (3-23), respectively. Similarly, replacing τ for h and u_n, ϕ' , γ, c' in place of X_1, X_2, X_3 and X_4 in the general equations yields the expression for the FOSM approximation of the mean and variance of τ as follows:

$$\mu\tau_{FOSM} = \tau\left(\mu u_n, \mu_{\phi'}, \mu_\gamma, \mu_{c'}\right) \qquad (9\text{-}69)$$

$$v\tau_{FOSM} = vu_n \cdot \left[\frac{\partial\tau}{\partial u_n}\right]_\mu^2 + v_{\phi'} \cdot \left[\frac{\partial\tau}{\partial\phi'}\right]_\mu^2 + v_\gamma \cdot \left[\frac{\partial\tau}{\partial\gamma}\right]_\mu^2 + v_{c'} \cdot \left[\frac{\partial\tau}{\partial c'}\right]_\mu^2 \qquad (9\text{-}70)$$

Following the same steps as in the case of SOSM approximation above and Evaluating τ, Eq.(6-18), at the mean values of the random variables, and replacing the moments of the random variables in Eq.(9-70) we reach at:

$$\mu_{\tau\,FOSM} = 503\,kN/m$$

$$\sigma\tau_{FOSM} = 173\,kN/m$$

Using either of the computed FOSM or SOSM mean and standard deviations of τ ($\mu_{\tau\,FOSM}$ and $\sigma_{\tau,FOSM}$ or $\mu_{\tau\,SOSM}$ and $\sigma_{\tau,SOSM}$) and by assuming central limit theorem the randomness in shear strength at the slice base can be estimated by a normal distribution having the respective computed mean and standard deviations as its parameters.

$$f_\tau(\tau)_{SOSM} = \frac{1}{173\cdot\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\tau-508}{173}\right)^2}, \quad -\infty < x < +\infty \qquad (9\text{-}71)$$

$$f_\tau(\tau)_{FOSM} = \frac{1}{173\cdot\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\tau-503}{173}\right)^2}, \quad -\infty < x < +\infty \qquad (9\text{-}72)$$

Plots of these pdfs are given in Figure 6-3 and Figure 6-4. Refer Eq.(9-43)and (9-44) in appendix 9.1.1 for pdf and cdf equations of a normal distribution.

9.4.3 SOSM and FOSM approximation of moments and uncertainty of shear stress (*G*)

Effective stress (gravitational driving stress) at the base of a sample slice of length l_n ($l_n = b_n/\cos\alpha$) that has a unit width along the dam axis direction can be given by Eq.(6-6). Replacing unit weight of water and relevant geometric values yields Eq.(6-7).

If γ is taken to be random variables then Eq. (6-7) is a function of random variable (FRV) with $n = 1$ random components $\boldsymbol{X} = (\gamma)$. This implies that G is a random variable and will have a cdf $F_G(G)$ describing its randomness, which can be estimated using the approximate moment methods FOSM and SOSM.

General relations for SOSM approximations of the mean value (μ_Y) and variance (ν_Y) of a multivariate FRV $Y = h(\boldsymbol{X})$, where X = (X_1, X_2, … , X_n) is continuous random vector, is given by Eq.(3-14) and Eq.(3-21), respectively. Again, replacing G for h and γ' in place of X in Eq.(3-14) and Eq.(3-21) yields the expression for the SOSM approximation for the mean and variance of G:

$$\mu G_{SOSM} = G(\mu_\gamma) + \frac{1}{2} \cdot \left(\nu_\gamma \cdot \left[\frac{\partial^2 G}{\partial \gamma^2} \right]_\mu \right) \qquad (9\text{-}73)$$

$$\nu G_{SOSM} = \left(\nu_\gamma \cdot \left[\frac{\partial G}{\partial \gamma} \right]_\mu^2 \right) - \frac{1}{4} \cdot \left(\nu_\gamma \cdot \left[\frac{\partial^2 G}{\partial \gamma^2} \right]_\mu \right)^2 \qquad (9\text{-}74)$$

$[..]_{\mu_i}$ means evaluate partial derivatives at mean of X_i. $[..]_\mu^2$ means evaluate the partial derivative of G at mean of X_i.and square it.

The first and second partial derivatives of G with respect to the random variable γ can be calculated from Eq.(6-7) and is given as:

$$\left[\frac{\partial G}{\partial \gamma} \right] = 6 \qquad \left[\frac{\partial^2 G}{\partial \gamma^2} \right] = 0$$

Substituting the above results and the moments of γ (μ_γ, ν_γ) in Eq.(9-73) yields the SOSM approximation for the mean of G:

$$\begin{aligned} \mu G_{SOSM} &= G(\mu_\gamma) + \frac{1}{2} \cdot \left(\nu_\gamma \cdot \left[\frac{\partial^2 \tau}{\partial \gamma} \right]_\mu \right) \\ &= 6.0 \cdot 17.8 + 78.03 + \frac{1}{2} \cdot (7.08 \cdot 0) \\ &= 180\, kN/m \end{aligned}$$

Similarly, substituting the results for the derivatives of G and moments of γ in Eq.(9-74) yields the SOSM approximation for the variance of G:

$$\nu G_{SOSM} = \left(\nu_\gamma \cdot \left[\frac{\partial G}{\partial \gamma} \right]_\mu^2 \right) - \frac{1}{4} \cdot \left(\nu_\gamma \cdot \left[\frac{\partial^2 G}{\partial \gamma^2} \right]_\mu \right)^2$$

$$= \left(7.08 \cdot (6)^2\right) - \frac{1}{4} \cdot (7.08 \cdot 0)^2$$

$$= 255\,(kN/m)^2$$

Consequently, the standard deviation for τ will be:

$$\sigma G_{SOSM} = \sqrt{255} = 16\,kN/m\,.$$

General relations for FOSM approximations of the mean value (μ_Y) and variance (ν_Y) of a multivariate FRV $Y = h(\boldsymbol{X})$, where $X = (X_1, X_2, \ldots, X_n)$ is continuous random vector, is given by Eqs.(3-22) and (3-23), respectively. Here also, replacing G for h and γ in place of X in Eqs.(3-22) and (3-23), the expression for the FOSM approximation μG_{SOSM} and νG_{SOSM} follows as:

$$\mu G_{FOSM} = G(\mu_\gamma) \qquad (9\text{-}75)$$

$$\nu G_{FOSM} = \nu_\gamma \cdot \left[\frac{\partial G}{\partial \gamma} \right]_\mu^2 \qquad (9\text{-}76)$$

Substituting the appropriate values for the derivative and moments of γ yields in Eq. (6-7):

$$\mu G_{FOSM} = \mu_\gamma \cdot \left(17.2 + \left(9.81 \cdot 22.8\right)\right) \cdot 3.26 \cdot sin\left(6.16\right)$$

$$= 6.0 \cdot 17.8 + 78.03$$

$$= 185\,kN/m$$

$$\nu G_{FOSM} = 7.08 \cdot (6)^2$$

$$= 254\,(kN/m)^2$$

Consequently,

$$\sigma G_{FOSM} = \sqrt{254} = 16\,kN/m$$

Similarly, using either of the computed FOSM or SOSM mean and standard deviations of G and by assuming central limit theorem the randomness in shear stress at the slice base can be

estimated by a normal distribution having the respective computed mean and standard deviations as its parameters.

9.4.4 SOSM and FOSM approximation of reliability (Z_s), sliding factor of safety ($F_{s,\,s}$) and sliding failure probability ($P_{f,\,s}$)

The governing equation for sliding reliability (Z_s) is given by Eq.(6-3). Eq.(6-4) gives sliding factor of safety ($F_{s,\,s}$) and sliding failure probability ($P_{f,\,s}$) is given by Eq.(6-5).

Thus, $P_{f,\,s}$ can be determined from cumulative distribution function (cdf) or probability density function (pdf) of Z_s or $F_{s,\,s}$. But, Z_s and $F_{s,\,s}$ are FRV. Because, the load (G) and strength (τ) mechanisms are functions of uncertain design parameters. Hence, the task of determining $P_{f,\,s}$ deals with computations with Z_s or $F_{s,\,s}$ FRVs and defining cdf or pdf of Z_s or $F_{s,\,s}$. FOSM or SOSM approximation methods can be used for estimating the moments of Z_s and $F_{s,\,s}$. Once the moments (mean and standard deviations of Z_s and $F_{s,\,s}$) are estimated then a normal distribution that describe the randomness in Z_s and $F_{s,\,s}$ can be constructed based on these moments and by assuming central limit theorem (CLT). However, for the $P_{f,\,s}$ estimation using FOSM and SOSM the use of Z_s equation gives relatively more accurate result than that found from using $F_{s,\,s}$ equation. That is because Z_s is a linear function while $F_{s,\,s}$ is rational function and the accuracy of results from FOSM and SOSM methods is dependent up on the degree of linearity of the FRV of interest near the mean values of its random variables (discussed in section 3.2.4 and 8.2). Therefore, here under, Z_s equation Eq.(6-3) is used for the computation of $P_{f,\,s}$.

The estimation of moments of Z_s and $F_{s,\,s}$ using FOSM and SOSM can follow two approaches. The first alternative is, to consider Z_s and $F_{s,\,s}$ equations as FRVs with $n = 2$ random components $\boldsymbol{X} = (\tau, G)$ and to use Eq.(6-3) and Eq.(6-4) for the computation. In this approach the moments of τ and G computed in sections 9.4.1 to 9.4.3 can be used in the corresponding FOSM and SOSM estimation of moments for Z_s and $F_{s,\,s}$. The second alternative is, to consider or Z_s and $F_{s,\,s}$ equations as FRVs with $n = 4$ random components $\boldsymbol{X} = (u_n,\ \phi'\ ,\ \gamma,\ c')$ and to use moments of the four random variables given in Table 9.6. The first alternative gives less accurate result than that of the second alternative. This is because the first approach involves two levels of approximations. Consequently, here under, the second approach is implemented for the computation of $P_{f,\,s}$ moments of Z_s and $F_{s,\,s}$.

Thus, substituting Eqs.(6-18) and (6-7) in Eq.(6-3) we can write:

$$Z_s = \left(3.28 \cdot c' + \left(724.037 - 3.28 \cdot u_n + 55.695 \cdot \gamma\right) \cdot tan\,\phi'\right) - \left(6.0 \cdot \gamma + 78.03\right) \qquad (9\text{-}77)$$

Replacing Z_s for h and u_u, ϕ' , γ, and c' in place of X_1, X_2, X_3 and X_4 in Eq.(3-14) and Eq.(3-21) yields the expression for the SOSM approximation for the mean and variance of Z_s:

$$\mu_{Zs,\,SOSM} = Z_s\left(\mu u_n, \mu_{\phi'}, \mu_\gamma, \mu_{c'}\right) + \frac{1}{2} \cdot \left(\nu u_n \cdot \left[\frac{\partial^2 Z_s}{\partial u_n^{\,2}}\right]_\mu + \nu_{\phi'} \cdot \left[\frac{\partial^2 Z_s}{\partial \phi'^2}\right]_\mu + \nu_\gamma \cdot \left[\frac{\partial^2 Z_s}{\partial \gamma^2}\right]_\mu + \nu_{c'} \cdot \left[\frac{\partial^2 Z_s}{\partial c'^2}\right]_\mu \right) \qquad (9\text{-}78)$$

$$v_{Z_{S,SOSM}} = \left(vu_n \cdot \left[\frac{\partial Z_s}{\partial u_n}\right]_\mu^2 + v_{\phi'} \cdot \left[\frac{\partial Z_s}{\partial \phi'}\right]_\mu^2 + v_\gamma \cdot \left[\frac{\partial Z_s}{\partial \gamma}\right]_\mu^2 + v_{c'} \cdot \left[\frac{\partial Z_s}{\partial c'}\right]_\mu^2 \right)$$

$$-\frac{1}{4}\cdot\left(vu_n \cdot \left[\frac{\partial^2 Z_s}{\partial u_n^{\,2}}\right]_\mu + v_{\phi'} \cdot \left[\frac{\partial^2 Z_s}{\partial \phi'^2}\right]_\mu + v_\gamma \cdot \left[\frac{\partial^2 Z_s}{\partial \gamma^2}\right]_\mu + v_{c'} \cdot \left[\frac{\partial^2 Z_s}{\partial c'^2}\right]_\mu \right)^2 \qquad (9\text{-}79)$$

The first and second partial derivatives of Z_s with respect to the different random variables can be calculated from Eq.(9-77) and they are given as:

$$\left[\frac{\partial Z_s}{\partial u_n}\right] = -3.28 \cdot tan\,\phi' \qquad \left[\frac{\partial^2 Z_s}{\partial u_n^{\,2}}\right] = 0$$

$$\left[\frac{\partial Z_s}{\partial \phi'}\right] = (724.037 - 3.28 \cdot U + 55.7 \cdot \gamma) \cdot \sec^2 \phi' \qquad \left[\frac{\partial^2 Z_s}{\partial \phi'^2}\right] = 0.0006\left(724.037 - 3.28 \cdot U + 55.7 \cdot \gamma\right) \cdot sec^2\,\phi' \cdot tan\,\phi'$$

$$\left[\frac{\partial Z_s}{\partial \gamma}\right] = 55.7 \cdot \tan \phi' - 6 \qquad \left[\frac{\partial^2 Z_s}{\partial \gamma^2}\right] = 0$$

$$\left[\frac{\partial Z_s}{\partial c'}\right] = 3.28 \qquad \left[\frac{\partial^2 Z_s}{\partial c'^2}\right] = 0$$

Evaluating the above derivatives at the respective means of the random variables; i.e. taking $u = \mu u_n$, $\phi' = \mu_{\phi'}$, $\gamma = \mu_\gamma$ and $c' = \mu_{c'}$ yields:

$$\left[\frac{\partial Z_s}{\partial u_n}\right]_\mu = -3.28 \cdot tan\,\phi' = -2.97 \qquad \left[\frac{\partial^2 Z_s}{\partial u_n^{\,2}}\right]_\mu = 0$$

$$\left[\frac{\partial Z_s}{\partial \phi'}\right]_\mu = (724.037 - 3.28 \cdot U + 55.7 \cdot \gamma) \cdot \sec^2 \phi' = 17.08 \qquad \left[\frac{\partial^2 Z_s}{\partial \phi'^2}\right]_\mu = 0.0006(724.037 - 3.28 \cdot U + 55.7 \cdot \gamma) \cdot \sec^2 \phi' \cdot \tan \phi' = 0.539$$

$$\left[\frac{\partial Z_s}{\partial \gamma}\right]_\mu = 55.7 \cdot \tan \phi - 6 = 44.4 \qquad \left[\frac{\partial^2 Z_s}{\partial \gamma^2}\right]_\mu = 0$$

$$\left[\frac{\partial Z_s}{\partial c'}\right] = 5.42 \qquad \left[\frac{\partial^2 Z_s}{\partial c'^2}\right] = 0$$

Evaluating Eq.(9-77) at the mean values of the random variables yields:

$$Z_s(\mu u_n, \mu_{\phi'}, \mu_\gamma, \mu_{c'}) = \left(3.28 \cdot \mu_{c'} + \left(724.037 - 3.28 \cdot \mu u_n + 55.695 \cdot \mu_\gamma\right) \cdot tan\,\mu_{\phi'}\right) - \left(6.0 \cdot \gamma + 78.03\right)$$
$$= \left(3.28 \cdot 5.4 + \left(724.037 - 3.28 \cdot 359.39 + 55.695 \cdot 17.8\right) \cdot tan\,42.13\right) - \left(6.0 \cdot 17.8 + 78.03\right)$$
$$= 318\,kN/m$$

Substituting the above results in Eq.(9-78) yields the SOSM approximation for the mean of Z_s:

$$\mu_{Z_s\ SOSM} = 318 + \frac{1}{2} \cdot \left(1021.83 \cdot 0 + \ 9 \cdot 0.539 + 17.8 \cdot 0 + 57.95 \cdot 0\right) = 314\,kN/m$$

Similarly, substituting the above results in Eq.(9-79) yields the SOSM approximation for the variance of Z_s:

$$\nu_{Zs\ SOSM} = \left(1021.83 \cdot (-2.97)^2 + 9 \cdot (17.08)^2 + 7.08 \cdot (44.42)^2 + 57.95 \cdot (3.28)^2\right)$$
$$- \frac{1}{4} \cdot \left(1021.83 \cdot 0 + 9 \cdot 5.4 + 7.08 \cdot 0 + 57.95 \cdot 0\right)^2$$
$$= 25{,}645\,(kN/m)^2$$

Consequently, the standard deviation for τ will be:

$$\sigma_{Zs\ SOSM} = \sqrt{25{,}645} = 160\,kN/m$$

As discussed in section 3.2.4, SOSM and FOSM methods do not provide a distribution of known standard type for characterizing the randomness of outputs of FRV. Rather results from FOSM and SOSM are mostly presented using normal distribution. The normal distribution is constructed from computed first two moments and by assuming central limit theorem.

Therefore, using the computed SOSM mean and standard deviations of Z_s ($\mu_{Zs\ SOSM}$ and $\sigma_{Z\,SOSM}$) the sliding reliability randomness at the slice base can be simulated by a normal distribution with mean $\mu_{Zs\ SOSM} = 318$ and $\sigma_{Zs\ SOSM} = 160$ (refer Eq.(9-43) and (9-44) in appendix 9.1.1 for pdf and cdf equations of normal distribution):

$$f_{Zs}(Z_s) = \frac{1}{160 \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{Z_s - 318}{160}\right)^2} \quad , -\infty < x < +\infty \qquad (9\text{-}80)$$

$$F_{Z,s}(Z_s) = \frac{1}{2}\left[1 + erf\left(\frac{Z_s - 318}{160\sqrt{2}}\right)\right] \qquad (9\text{-}81)$$

The sliding failure probability ($P_{f,\,s}$), which is defined in Eq.(6-5), can thus be computed from either (9-80) or (9-81) as:

$$P_{f,s} = P(Z_s \le 0) = \int_{-\infty}^{0} f_{Zs}(Z_s)dZ_s = F_{Zs}(0) = 0.024$$

Similarly, substituting Eqs.(6-18) and (6-7) in (6-4) we can write the equation for the factor of safety against sliding ($F_{s,\,s}$), as a FRV with $n = 4$ random components $\boldsymbol{X} = (u_u,\ \phi'\ ,\ \gamma,\ c')$:

$$F_{s,s} = \frac{3.38 \cdot c' + (724.037 - 3.28 \cdot U + 55.695 \cdot \gamma) \cdot \tan\phi'}{6.0 \cdot \gamma + 78.03} \qquad (9\text{-}82)$$

Replacing $F_{s,s}$ for h and $u_u,\ \phi'\ ,\ \gamma,$ and c' in place of X_1, X_2, X_3 and X_4 in Eq.(3-14) and Eq.(3-21) respectively yields the expression for the SOSM approximation for the mean and variance of $F_{s,\,s}$:

$$\mu_{FsS,\ SOSM} = F_{s,s}\left(\mu_U, \mu_{\phi'}, \mu_\gamma, \mu_{c'}\right) + \frac{1}{2} \cdot \left(\nu_U \cdot \left[\frac{\partial^2 F_{s,s}}{\partial U^2}\right]_\mu + \nu_{\phi'} \cdot \left[\frac{\partial^2 F_{s,s}}{\partial_{\phi'}{}^2}\right]_\mu + \nu_\gamma \cdot \left[\frac{\partial^2 F_{s,s}}{\partial \gamma^2}\right]_\mu + \nu_{c'} \cdot \left[\frac{\partial^2 F_{s,s}}{\partial c'^2}\right]_\mu \right) \qquad (9\text{-}83)$$

$$\nu_{F\,s,s\,,SOSM} = \left(\nu u_u \cdot \left[\frac{\partial F_{s,s}}{\partial u_u}\right]_\mu^2 + \nu_{\phi'} \cdot \left[\frac{\partial F_{s,s}}{\partial \phi'}\right]_\mu^2 + \nu_\gamma \cdot \left[\frac{\partial F_{s,s}}{\partial \gamma}\right]_\mu^2 + \nu_{c'} \cdot \left[\frac{\partial F_{s,s}}{\partial c'}\right]_\mu^2 \right)$$

$$-\frac{1}{4} \cdot \left(\nu_U \cdot \left[\frac{\partial^2 F_{s,s}}{\partial {u_u}^2}\right]_\mu + \nu_{\phi'} \cdot \left[\frac{\partial^2 F_{s,s}}{\partial \phi'^2}\right]_\mu + \nu_\gamma \cdot \left[\frac{\partial^2 F_{s,s}}{\partial \gamma^2}\right]_\mu + \nu_{c'} \cdot \left[\frac{\partial^2 F_{s,s}}{\partial c'^2}\right]_\mu \right)^2 \qquad (9\text{-}84)$$

Again, the first and second partial derivatives of $F_{s,\,s}$ with respect to the different random variables can be calculated from Eq.(9-82) and is given as:

$$\left[\frac{\partial F_{s,s}}{\partial u_n}\right] = \frac{9.376 \cdot tan\,\phi'}{223.6 + 17.2 \cdot \gamma} \qquad \left[\frac{\partial^2 F_{s,s}}{\partial {u_n}^2}\right] = 0$$

$$\left[\frac{\partial F_{s,s}}{\partial \phi'}\right] = \frac{0.0499(724.726 - 3.28 \cdot U + 55.748 \cdot \gamma) \cdot \sec^2\phi'}{223.6 + 17.2 \cdot \gamma} \qquad \left[\frac{\partial^2 F_{s,s}}{\partial \phi'^2}\right] = \frac{0.00174(724.726 - 3.28 \cdot U + 55.748 \cdot \gamma) \cdot \sec^2\phi' \cdot \tan^2\phi'}{223.6 + 17.2 \cdot \gamma}$$

$$\left[\frac{\partial F_{s,s}}{\partial \gamma}\right] = \frac{-161.274 \cdot C' + (-7.275 \cdot 10^{-12} + 161.27 \cdot U) \cdot \tan\phi'}{(223.6 + 17.2 \cdot \gamma)^2} \qquad \left[\frac{\partial^2 F_{s,s}}{\partial \gamma^2}\right] = \frac{5547.82 \cdot c' + (-5547.82 \cdot U - 1.455 \cdot 10^{-11} \cdot \gamma) \cdot \tan\phi'}{(223.6 + 17.2 \cdot \gamma)^3}$$

$$\left[\frac{\partial F_{s,s}}{\partial c'}\right] = \frac{9.376}{223.6 + 17.2 \cdot \gamma} \qquad \left[\frac{\partial^2 F_{s,s}}{\partial c'^2}\right] = 0$$

Evaluating Eq.(9-77) at the mean values of the random variables yields:

$$F_{s,s}(\mu u_u, \mu_{\phi'}, \mu_\gamma, \mu_{c'}) = \frac{3.28 \cdot \mu_{c'} + \left(724.037 - 3.28 \cdot \mu u_n + 55.695 \cdot \mu_\gamma\right) \cdot \tan \mu_{\phi'}}{6.0 \cdot \gamma + 78.03}$$

$$= \frac{3.28 \cdot 5.4 + \left(724.037 - 3.28 \cdot 359.39 + 55.695 \cdot 17.8\right) \cdot \tan 42.13}{6.0 \cdot 17.8 + 78.03}$$

$$= 2.75 kN/m$$

Following the same routine procedures of substituting moments of random variables, i.e. taking $u = \mu u_n$, $\phi' = \mu_{\phi'}$, $\gamma = \mu_\gamma$, and $c' = \mu_{c'}$ and evaluating the derivatives at the respective means of the random variables yield:

$$\mu_{FsS,\ SOSM} = 2.75 \qquad\qquad \nu_{F\,s,s,SOSM} = 0.62$$

These are the SOSM approximations for the mean and standard deviations of the factor of safety against sliding $F_{s,s}$. Consequently,

$$\sigma_{F\,s,s\ SOSM} = \sqrt{0.62} = 0.79\,.$$

Here again, a normal distribution bounding the uncertainty of the prevailing factor of safety value can be fitted using the computed SOSM mean and standard deviations of $F_{s,\,s}$.

$$f_{Fs,s}(F_{s,s}) = \frac{1}{0.79 \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{F_{S,S} - 2.75}{0.79}\right)^2}, \quad -\infty < x < +\infty \qquad (9\text{-}85)$$

The plot of this equation is given in Figure 6-6. $P_{f,.s}$ can be calculated also from (9-85) for the case $P(F_{s,s} \le 1)$. However, as discussed above the $P_{f,.s}$ estimated using $F_{s,s}$ equation is less accurate when compared to the one estimated using Z_s equation. This is because $F_{s,\,s}$ is rational function and the accuracy of results from FOSM and SOSM methods is dependent up on the degree of linearity of the FRV of interest near the mean values of its random variables.

FOSM approximations for the moments of Z_s, $F_{s,\,s}$ and thus that of $P_{f,.s}$ can be completed following the same procedures as demonstrated in 9.4.2 and 9.4.3. Results of the SOSM and FOSM computations made in sections9.4.1 to 9.4.4 are summarized in Table 6.2 and plots of the pdfs is given in Eq.(6-3).

10 REFERENCES

Adams, R. A. (2006). "Calculus: A Complete Course." *Prentical Hall*, Canada.

Ang, A. H. S. and Ellingwood, B. R. (1971). "Critical analysis of reliability principles relative to design." *Proc. of the 1st international conference on application of statistics and probability to soil structural engineering*, Hong Kong university press.

Ang, A. H. S. and Cornell, C. A. (1974). "Reliability bases of structural safety and design." *J. Struct. Div.*, ASCE, 100(ST9), 1755-1769.

Ang, A. H.S. and Tang, W. H. (1975). "Probabilistic concepts in engineering planning and design. Vol. I – basic principles." *John Wiley and Sons*, New York.

Ang, A. H.S. and Tang, W. H. (1984). "Probabilistic concepts in engineering planning and design. Vol. II – Decisions, Risk and Reliability." *John Wiley and Sons*, New York.

Ansell, J. I. and M. J. Phillips. (1994). "Practical methods for reliability data analysis." *Clarendon Press, Oxford University Press Inc.*, New York.

Arrora, Jasbir s. (ed.) (1997). "Guide to structural optimization." *Structural Engineering Institute Technical committee on Optimal Structure*, *American society of civil engineers*. 211-231.

Association of State Dam Safety Officials (ASDSO) (2009). "Dam failures and incidents" <http://www.damsafety.org> (5. Jan. 2009).

Ayyub, B.M (ed.) (1998). "Uncertainty modeling and analysis in civil engineering." *CRC Press*, Boston.

Baecher, G. B. and Christian, J. T. (2003). " Reliability and statistics in geotechnical engineering, *John Wiley and Sons,* London.

Benjamin, J. R, Cornell, C. (1970). "Probability, statistics and decision for civil engineers." *McGrew-Hill*, NY.

Bhunya, P.k, R. Berndtsson, C.S.P. Ojha, S.K. Mishra (2006). "Suitability of gamma, chi-square, weibull, and beta distributions as synthetic UH." J. *of Hydro.*, 334, 28-38.

Bishop, A. W. (1955). "The use of slip circle in stability analysis of slopes." *Geotechnique*, 5(1), 7-17.

Bishop, A. W. and Bjerrum, L. (1960). "The relevance of traxial test to the solution of stability problems." *In proceedings of the Conference on Shear Strength of Cohesive Soils*, Boulder, CO, American Society of Civil Engineers, New York, 437-501.

Bishop, A. W. and Morgenstern N.(1960). "Stability coefficient for earth slopes." *Geotechnique*, 10(4), 129-50.

Cheung-Bin, Lee, Ju-Won park (1997). "Reliability analysis based on fuzzy-Bayesian approach." *7th Int. Conf. on Computing in Civil and Building Engineering*. Seoul, Korea, pp 589-593.

Chou, Karen C. and Yuan, Jie (1993). "Fuzzy-Bayesian approach to reliability of existing structures." ASCE, *J. structural engineering*, 119(11), 3276-3290.

Chow, V.T. (1951). "A general formula for hydrologic frequency analysis." *Tans. American Geophysical Union*, 32, 231-237.

Chow, V.T., R. Maidment, and Mays, Larry. (1988). "Applied hydrology." *McGraw-Hill*, NY.

CNN(a) (2005). "The latest on Katrina's aftermath." <http://us.cnn.com/2005/US/09/22/katrina. impact/index.html> (29 September 2005).

Colubi, Ana, Domínguez-Mencheroa, J. Santos, López-Díaza, Miguel and Ralescu, Dan A. (2001). "On the formalization of fuzzy random variables." *Information sciences*, *Elsevier*, 133 (1-2), 3-3.

Cornell, C. A. (1967). "Bounds on reliability of structural systems." J. struct. Div., ASCE, 93(ST1), 171-200.

Cornell, C. A. (1969). "A probability-based structural code." *J. Amer. Concrete Inst.*, ACI, 66(12), 974-985.

Crowder, M. J., A. C. Kimber, R. L. Smith, and Sweeting, T. J. (1991). Statistical analysis of reliability data. *Chapman and Hall*, London.

Crystal Ball (2005a). "Crystal Ball 7.1 Reference Manual." *Decisioneering Inc*. Denver.

Crystal Ball (2005b). "Crystal Ball 7.1 User Manual." *Decisioneering Inc*. Denver.

D'Agostino, Ralph. B. and Stephens, Michael. A. (eds.) (1986) Goodness-of-fit techniques. *Marcel Dekker*, New York.

DeGroot, Morris H. (1986). "Probability and statistics." *Addison-Wesley*.

Devore, J. and Peck, R. (1986). "Statistics- the exploration and analysis of data." *West Publishing*.

Ehrenberg, Andrew. S. C. (1986). "A primer in data reduction- an introductory statistics textbook." *John Wiley and Sons,* .

Elishakoff, Isaac (ed.) (1999). "Whys and hows in uncertainty modeling probability, fuzziness and anti-optimization." *Springer*, New York.

El-Ramly, H., Morgenstern, N. R. and Cruden, D. M. (2002). "Probabilistic slope stability analysis for practice." *Canadian Geothechnical J.*, 39, 665-683.

Feller , W. (1967). "An Introduction to probability theory and its applications." Wiley, NY.

Frankel, Ernst. G. (1988). "Systems reliability and risk analysis." *Kluwer Academic Publishers Group*, The Netherlands.

Gibb, Sir Alexander and partners (1975). "Feasibility study of the lower Awash valley, final report part II, Anex 2, climate and hydrology." *Sir Alexander Gibbs in association with Hunting technical services ltd*., London.

Golberg, M. A. (1984). "An introduction to probability theory with statistical approach." *Plenum Press*, NY.

Greenberg, Michael (1998). "Advanced engineering mathematics, *Prentice Hall,* 629-641.

Gringorten, Irving. I. (1963). "A plotting rule for extreme probability paper." *J. of Geophysical Research*, 68(3), 813-814.

Grosh, D. L. (1989). "A primer of reliability theory." *Wiley*, New York.

Gumbel, E. J. (1941). "Probability interpretation of the observed return periods of floods." T*rans. American Geophysical Union*, 21, 836-850.

Gumbel, E. J. (1942). "Statistical control curves for flood discharge." *Trans. American Geophysical Union*, 23, 489-500.

Gumbel, E. J. (1943). "On the plotting of flood discharge." *Trans. American Geophysical Union*, 24, 699-719.

Gumbel, E. J. (1958). "Statistics of extremes." *Columbia University Press*, N.Y.

Haktanir, T. and Sezen, N. (1999). "Suitability of two-parameter gamma and there parameter beta distributions as synthetic UH in Anatolia." *Journal of Hydrology*. Sc., 35(2). 167-184.

Halcrow (1989). "Master plan for development of surface water resources in the awash basin-Climate and Hydrology." *Halcrow*, 4(A), Addis Ababa.

Halcrow, SOGREAH, and Metaferia Consult (2006). "Awash River basin flood control and watershed management study-Hydrological Studies." *Halcrow*, 2(B), Addis Ababa.

Halcrow (2006). "Awash River basin flood control and watershed management study project-sediment sources and control measures." *Halcrow*, Base line report (6), Addis Ababa.

Harr, M. E. (1989). "Probabilities estimates for multivariate analysis." *Applied mathematical modeling*, 13, 313-318.

Harris, Bernard (1966). "Theory of probability." *Addison-Wesley*, Massachusetts, USA.

London.

Hartford, N. Desmond and Baecher, Gregory B. (2004). "Risk and uncertainty in dam safety." *Thomas Telford*, London.

Hasofer, Abraham and Lind, Niels (1974). "Exact and invariant second moment code format." *J. of the engineering mechanics division*, 100 (1), 111-120.

Horlacher, H.–B., and Negede, Abate Kassa (2008). "Dam risk and safety analysis - comparison of probabilistic methods." *Proc. of the 76th International Commission on Large Dams (ICOLD) Annual Meeting and Symp. in Operation, Rehab. & Up-grading of Dams*, (CD-ROM), ICOLD, Sofia, Bulgaria.

ICOLD (1995). "Dam failure statistical analysis." *Bulletin 99, ICOLD*, Paris.

ICOLD (1992)."Selection of design flood." *Bulletin 82, ICOLD*, Paris.

IKSE. (2004). "Documentation des Hochwasser vom August 2002 im Einzugsgebiet der Elbe." *Internationale Kommission Zum Schutz der Elbe*, Magdeburg.

Iman, R. L., Helton, J. C. and Campbell, J. E. (1981). "An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment." *J. quality technology*, 13(3): 174-183.

Institute of Civil Engineers(ICE) (1996). "Floods and reservoir safety." *Thomas Teleford*, UK.

Jansen, Robert B., John Lowe III, Richard W. Kramer, Steve J. Poulos (1988). "Earthfill Dam Design and Analysis" in Advanced Dam Engineering for Design: Construction, and Rehabilitation (ed. R. B. Jansen)." *Van Nostrand Reinhold*, New York, 256-320.

Jansen, Robert B. (ed.) (1988). "Advanced dam engineering for design: construction, and rehabilitation." *Van Nostrand Reinhold*, New York.

Johnson, N. L. and Kolz, S. (1969). "Discrete distributions." *Houghton Mifflin*, Boston.

Johnson, N. L. and Kolz, S. (1970). "Continuous uni-variate distributions." *Houghton Mifflin*, New York.

Johnson, N. L. and Kolz, S. (1972). "Distributions in statistics: continuous multivariate distributions." *John Wiley*, New York.

Johnson, N. L., Kolz, S. and Kemp, A. W. (1992). "Uni-variate Discrete distributions." *Wiley*, New York.

Joos, B., J. Darakhani, Mouvet, L., and Mehinrad, A. (2005). "An integrated probabilistic approach for determining the effect of extreme hydrological events on flood evacuation systems." *Proceedings ICOLD 73rd Annual Meeting*, Tehran, Iran, 147-S2.

Klein, Bastian (2006). "Stochastic generation of hydrographs for the flood design of dams." *Proc., 3rd Int. Symp. in Integrated WRM*, Ruhr-University, Bochum, Germany.

Krätschmer, V. (2001). "A unified approach to fuzzy random variables." *Fuzzy sets and systems*, *Elsevier*, 123(11), 1-9.

Kuo, Jan-T., Yen, Ben-C., Hsu, Yung-C., and Lin, Huei-F. (2007). "Risk analysis for dam overtopping-Feitsui reservoir as a case study." *J. Hydraulic Eng., ASCE*, 133(8), 955-963.

Kwakernaak, H. (1978). "Fuzzy random variables-I. definitions and theorems." *Information sciences, Elsevier*, 15, 1-29.

Kwakernaak, H. (1979). "Fuzzy random variables-II. algorithms and examples for the discrete case." *Information sciences*, Elsevier, 17, 253-278.

Lawless, Jerald F. (1982). Statistical models and methods for lifetime data. *Wiley and Son*, NY.

Lecornu, J (1998). "Dam safety: from the engineer's duty to risk management." *Int. J. on Hydropower and Dams.*, 2, 53-56.

Lange, S. J. de (1997). "Statistics 1 – IHE lecture note series." *IHE*, Delft Then Netherlands.

Leitch, Roger. D. (1995). "Reliability analysis for engineers." *Oxford Univ. Press.*, New York.

Lian, Yanqing, M. ASCE, Ben Chie Yen, and FASCE (2003). "Comparison of risk calculation methods for culverts." *J.of Hydraulic Engineering*, 129(2):140-152.

Linsley, K., M.Kohler, and J.Paulhus(1962). "Hydrology for engineers.", *McGrawHill*, NY.

Lumb, P. (1966). "The variability of natural soils." *Canadian geotechnical journal*, 3(2), 74-97.

Lumb, P. (1974). "Application of statistics in soil mechanics." *Chapter 3 in Soil-mechanics- new horizon, ed. By I. K. Lee*, Newnes Butterworth, London.

Manache, G. and Melching, C. S. (2004). "Sensitivity analysis of a water-quality model using Latin hypercube sampling." *J. Water Resources planning and management, ASCE*, 130(3), 232-242.

Martin, Helmut and Reinhard Pohl (1998). "New aspects of freeboard design at dams." *Dam safety- procededings of the international symposium on new trends and guidelines on dam safety*, L.Berga (ed.), A.A. Balkema, Rotterdam.

Mays, Larry W. (1999). "Hydraulic design handbook." *McGraw-Hill*, New York. PP 7.16-7.38.

McKay, M. D. (1988). "Sensitivity and uncertainty analysis using a statistical sample of input values." *In Uncertainty Analysis, chapter 4, Ronen, Y., ed.*, *CRC press, Boca Rata*, Florida. Page 145-186.

McKay, M.D., Beckman, R. J., Conver, W. J. (1979). "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code." *Technometrics (American Statistical Association)*, 21(2): 239-245.

Melchers, Robert E. (1987). "Structural reliability analysis and prediction." *John Wiley and Sons*, New York.

Melching, Charles S., M.ASCE and Bauwens, Willy (2001). "Uncertainty in coupled nonpoint source and stream water quality models." *J. Water Resources planning and management*, ASCE, 127(6), 403-413.

Melih, A. Yanmaz and Beser, Resat (2004). "On the reliability-based safety analysis of the Porsuk Dam." *Turkish Journal of Engineering and Environmental Science*. 29:309-320.

Möller, Bernd and Beer, Michael (2004). "Fuzzy Randomness: uncertainty in civil engineering and computational mechanics." *Springer*, Belin.

Möller, Bernd 2006. "Lectures on risk analysis." Dresden University of Technology, Faculty of Civil Engineering, Institute for Statics and Dynamics, Rehabilitation Engineering Program.

Möller, Bernd and Reuter, Uwe (2007). "Uncertainty forecasting in engineering." *Springer*, Berlin.

Nadarajah, Saralees (2007). "Probability models for unit hydrograph derivation." *J. of Hydrology.*, 344,185-189.

National Environmental Research Council (NIC) (1975). "Flood studies report.", *NIC*, 1, London.

Negede, A. Kassa, H.–B. Horlacher (2009). "Approaches for dam overtopping probability evaluation." *Proc. of the 33th Int. Ass. for Hyd. Eng. and Research (IAHR) congress*, *American Society of Civil Engineers (ASCE)*, Vancouver, Canada.

Negede, Abate K., Horlacher, H.–B. Horlacher (2008a). "Analytical solutions for assessing randomness in dam design, risk and safety analysis I-theory." *Proc. of the 16th Int. Ass. for Hyd. Eng. and Research (IAHR)-Asia Pacific Division (APD) and 3rd IAHR Int. Symp. on Hyd. Stru., Tsinghua Uni. Press*, Nanjing. 5, China 1874-1880.

Negede, Abate Kassa, and Horlacher, H.–B. (2008b). "Analytical solutions for assessing randomness in dam design, risk and safety analysis II-case study." *Proc. of the 16th IAHR-APD and 3rd IAHR-ISHS, Tsinghua Uni. Press*, Nanjing, China, 5, 1881-1887.

Negede, Abate Kassa (2002). "Investment-maintenance-risk optimization using probabilistic design methods." M.Sc. Thesis, International Institute For Infrastructural Hydraulic and Environmental Engineering (IHE), Delft, The Netherlands.

Novak, P. , Moffat, A. I. B., Nalluri, C. and Narayanan, R. (2003). "Hydraulic structures." *Taylor and Francies Group*, London.

O'Connor, Patric D.T. (1991). "Practical reliability engineering.", *John Wiley and Sons*, Chichester.

Pohl, R. (1999). "Estimation of the probability of hydraulic-hydrological failure of dams." *Proc. of Int. Workshop on Risk Analysis in Dam Safety Assessment*, Taipei, Republic of China.

Pohl, R. (2000). "Failure frequency of gates and valves at dams and weirs". *Int. J. on Hydropower and Dams*, (6), 77-82.

Ponce, V. M. (1989). "Engineering hydrology." *Prentice-Hall*. New Jersey.

Rosenblueth, Emilio (1975). "Point estimates of probability moments." *Proc. , National academy of science*, Washingeton, D.C., 3812-3814.

Rosenblueth, Emilio (1976). "Towards optimum design through building codes." *ASCE J. Struct. Div*. 102, 591-607.

Rosenblueth, Emilio (1981). "Two point estimates in probabilities." *Applied mathematical modeling*, 5, 309-335.

Rosenblueth, E. and Esteva L. (1972). "Reliability basis for some Mexican codes." *Probabilistic design of reinforced concrete buildings, ACI publication*, SP-31, Detroit.

Satoh, H. and Y. Yamaguchi (2005). "Probabilistic evaluation of under seepage of dam rock foundation considering spatial correlation in permeability." *Proceedings ICOLD 73rd Annual Meeting*, Tehran, Iran, 087-S3.

Senturk, Fuat (1994). "Hydraulics of dams and reservoirs." *WRS Publications*, Colorado.

Şen, Zekâİ (1978). "A mathematical model of monthly flow sequences." *Hydrological sciences*, 23(2), 223-229.

Sherard, James L., Richard J. Woodward, Stanley F. Gizienski and William A. Clevenger (1963). "Earth-Rock Dams: Engineering problems, design and construction." *John Wiley and Sons, Inc.*, New York.

Sherman, L. K. (1932). "Stream flow from rainfall by unit-graph method." *Eng. News-Record*. 108, 501-505.

Sillman, S., Schmucker, D., Sack, E. and Spencer, B. (1990). "Reliability theory as applied to estimation in converging versus non-converging flow fields." *In ModelCARE 90: calibration and reliability in ground water modeling. International Association of Hydrological Sciences (IAHS) publication no 195*, IAHS press, Wallingford

Sogreah-FAO (1965). "Report on survey of the Awash River basin, Volume II, climatology and hydrology", *Food and Agriculture Organization of the United Nations (FAO)*.

Subramanya, K. (1984). "Engineering hydrology." *McGraw-Hill*, New Delhi.

Thoft-Christensen, P. and Backer, M. J. (1982). "Structural reliability theory and its application." *Springer-verlag*, Berlin.

Toft-Christensen and Yoshisada Murotsu (1986). "Application of structural system theory." *Springer-verlag*, Berlin.

Thomas, H. A. and Fiering, M. B. (1962). "Mathematical synthesis of stream flow sequences for analysis of river basins by simulation. In: Design of water resources systems, eds. A. mass et al. chapter 12, Harvard university press, Cambridge.

Treiber, Bertold and Plate, Erich J. (1997). "A stochastic model for simulation of daily flows." Hydrological sciences, XXII(1), 175-192.

US Army Copps of Engineers (USACE) (2006). "USACE hydrologic safety assessment."<http://www.ferc.gov/eventcalendar/files/20060315114002-USACE%20hydrology%20assesemnet.pdf>(7 Sept. 2008) published on Northwest Dam Safety Regional Forum February 14-15, 2006, Jerry W. Webb. USACE, Portland, USA.

UZBEK (1985). "Feasibility study, proposal and estimates of cotton development on the area of 60,000 ha in the lower Awash valley, Ethiopia."

Vrijling, J. K. and Verhagen (2000). "Probabilistic design of hydraulic structures." *IHE Lecture Note Series*, Delft, The Netherlands.

Webb, J. W. (2006). "US Army of Corps of Engineers (USACE) Hydrologic safety assessment. 2006 Northwest Dam Safety Regional Forum, Portland." <http://www.ferc.gov> (10 Feb. 2009).

Wikipedia (2009). "Hurricane Katrina." <http://en.wikipedia.org/wiki/Hurricane_Katrina> (29 April 2009), Wikipedia the free encyclopedia.

World Meteorological Organization (WMO) (1973). Manual for estimation of probable maximum precipitation, *Secretariat of the World Meteorological Organization*, Geneva.

WWDSE (Ethiopian Water Works Design and Supervision Enterprise) (2005a). "Hydrology of Tendaho Dam.", *WWDSE*, Addis Ababa.

WWDSE (2005b). "Dam and appurtenant works." *WWDSE*, Addis Ababa.

Yoon, Gil L. and O'Neill, Michael W. (2002). "Investigation of the load model effects on resistance factor determination for pile foundation." I*n Honjo, Yusuke, Osamu Kusakabe, Kenji Matsui, Masayuki Kouda and Gyaneswor Pokharel (ed.). Foundation design codes and soil investigation in view of international harmonization and performance. Swets and Zeitlinger*, Lisse. 283-288.