

# Application of Bayesian Networks in distribution system risk management

K. Sand, M. D. Catrinu, G.H. Kjølle, *SINTEF Energy Research* and S. Bonnoit, J. Aupied, *EdF R&D*

**Abstract--** There is an increasing trend towards using the concept of risk assessment as an important tool in distribution system asset management. In an ongoing R&D project (RISK DSAM), the main objective is to investigate how information about risk exposure can improve maintenance and reinvestment decisions in an electrical distribution company. As the popularity of Bayesian Networks (BN) is increasing, the capacity of this methodology is being explored in the project. The paper presents the BN framework and the link to risk based asset management of electrical distribution systems. Examples are included to illustrate the concept as well as a case study evaluating the risk of failure of an overhead line and its dependence of risk influencing factors such as maintenance intensity and environment.

**Index Terms--** Decision-making, Power distribution maintenance, Power distribution reliability, Risk analysis.

## I. DISTRIBUTION SYSTEM ASSET MANAGEMENT

THERE is an increasing trend towards using the concept of risk assessment as an important tool in distribution system asset management. Asset management is defined in [1] as “Systematic and coordinated activities and practices through which an organization optimally manages its assets, and their associated performance, risks and expenditures over their lifecycle for the purpose of achieving its organizational strategic plan”. Over the lifecycle, the distribution system assets and the overall distribution system are subjected to a variety of threats and uncertainties. Risk management is a suitable framework for addressing these challenges.

Risk is defined as a combination of the frequency or probability, of occurrence and the consequence of a specified hazardous event [2]. Risk analysis attempts to answer three fundamental questions [3]:

1. What can go wrong?
2. How likely is it to happen?
3. What are the consequences?

The answer to the first question are the risk scenarios (S), describing the considered threats or hazards, while the answer to the second question is a probability statement regarding the different scenarios (p).

The answer to the third question is a qualitative or quantitative description or evaluation of the consequences (C) of the different scenarios related to individuals, professionals, populations, property or the environment.

Each risk scenario might hence be described by three

parameters:  $\langle S, p, C \rangle$  and the total risk picture is given by listing all risk scenarios with their associated probabilities and consequences.

Risk scenarios can include threats or hazards, events and trends. Adverse weather, ageing overhead lines, overloading of components and lack of maintenance, are all examples of risk scenarios that might lead to faults in the system. Hence, understanding component failure mechanisms and how possible barriers (technology, maintenance, protection etc) influence the risk, are of great importance in asset management decision making.

In an ongoing R&D project (RISK DSAM<sup>1</sup>), the main objective is to investigate how information about risk exposure can improve maintenance and reinvestment decisions in a power distribution company. A risk management concept (flowchart) is developed to support the overall working process, and Bayesian Networks are explored in the project as a generic decision support methodology.

The risk management concept and the application of BN are described in the following sections. The paper also includes some small examples and case studies that illustrate the possible application of BN to provide risk relevant parameters.

## II. THE RISK DSAM RISK MANAGEMENT CONCEPT

A generic flowchart for risk based distribution system management is given in Fig.1. The concept is largely based on combining distribution system planning concepts used by utilities for years and general risk management concepts described in standards like [1], [2] and [16].

The flowchart consists of four main activities:

1. Risk Study Planning
2. Risk Scenario Identification
3. Risk Modeling, Analysis and Decision Making
4. Risk Communication

The content of each activity is briefly described below. For further details see [4].

<sup>1</sup> RISK DSAM is an R&D project funded by the Research Council of Norway, EdF and network companies.

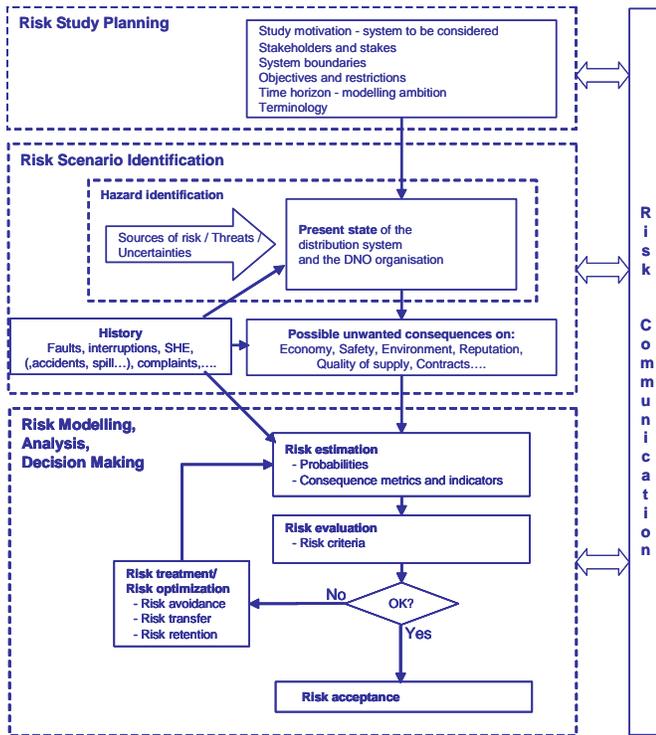


Fig. 1. A risk based distribution system management approach

#### A. Risk Study Planning

This is largely a project planning activity identifying the motivation for the study, identifying the system to be studied, stakeholders involved, objectives and restrictions, and to decide on modeling ambition needed to fulfill the overall goal of the study in a cost effective way.

#### B. Risk Scenario Identification

To perform risk management, a core activity is to identify the sources of risk, threats and uncertainties that might have harmful consequences. These hazards, when present, influence the distribution system and unwanted events with potential unwanted consequences. The hazard identification is supported by the historic track record and experience of the distribution system operator (DNO) as indicated in Fig.1.

To evaluate the consequence of this unwanted event, it is necessary to identify the impact for all relevant stakeholders.

The consequences might be classified in several ways. In the presented concept the following consequence categories are used:

- Economy: Economic impact for the stakeholders involved
- Safety: Occupational and public safety
- Environment: Pollution, leakage etc.
- Reputation: Branding and goodwill effects
- Quality of supply: Interruptions, voltage quality impact
- Contracts: Violation of contracts, regulations etc.

#### C. Risk Modeling, Analysis and Decision Making

This main activity comprises:

- The Risk Assessment i.e the overall process of risk analysis and risk evaluation
- The Risk Treatment i.e. the process of selection and implementation of measures to modify risk
- The Risk Acceptance i.e. the decision to accept the estimated risk being within the acceptable risk level for each risk criterion.

#### D. Risk Communication

Risk communication is a parallel activity supporting the other three activities and serves the purpose of exchanging or sharing of information about risk between decision-makers and other stakeholders.

### III. BAYESIAN NETWORKS SUPPORTING THE RISK MANAGEMENT CONCEPT

Bayesian Networks constitute a modeling framework which has found applications in domains like, e.g., software reliability, fault finding systems, and general reliability modeling [5],[7],[9],[10],[11],[12],[13],[15]. As reliability modeling is an important part of the risk management concept, it is of interest to test the usefulness of BNs in distribution system risk management. BNs are more flexible than the traditional Fault Tree concept both in terms of modeling features and in the calculations scheme. BNs can model more complicated relationships between variables and do not require Boolean variables, but can support several variable types. BNs offer among others a compact presentation of the interactions in a stochastic system by visualizing system state variables and their dependencies.

More specifically, the qualitative part of a BN consists of a, a directed acyclic graph where the nodes mirror the random variables, and the edges of the graph represents the conditional dependence between variables. In a risk management context the sources of risk, the consequences, and the threats, all are random variables which might to some degree be dependent of each other.

Bayesian Networks can be used to model and give decision support for large parts of distribution system asset management given in Fig.1. An example is given in Fig. 2:

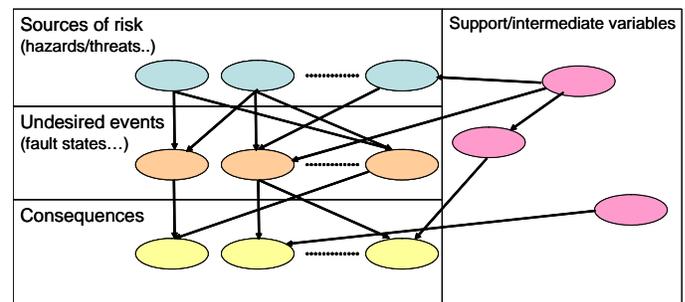


Fig. 2. Risk analysis modeled as a Bayesian Network

The oval shapes in the figure represent state variables and the arrows the causal relationships. The layers in the graph structure the variables in four main classes: 1) sources of risk variables used to describe hazards and threats for a component, group of components or a system; 2) unwanted events variables; 3) consequence variables describing possible unwanted effects of the risk scenarios; and 4) intermediate or support variables not directly fitting into the three other classes.

Bayesian Networks are further described in the following section.

#### IV. BAYESIAN NETWORKS

Bayesian Networks (BN) originated in the field of Artificial Intelligence, where it was used as a robust and efficient framework for reasoning with uncertain knowledge. BN constitute a modeling framework which is particularly easy to use in interaction with domain experts. A BN consists of two main parts:

- A qualitative part – a directed acyclic graph
- A quantitative part – a set of conditional probability functions

Fig.3 gives an example of a small BN:

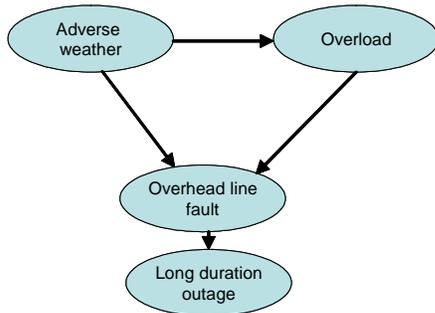


Fig. 3. Risk of overhead line outage. Example of a Bayesian Belief Network

The BN model in Fig. 3 consists of four nodes which are explained below:

1. {Adverse weather} – a state variable giving the probability of adverse weather in the overhead line surroundings (snow, strong wind etc.)
2. {Overload} – a state variable giving the probability of current overload with respect to the ampacity of the overhead line.
3. {Overhead line fault} – a state variable giving the probability of an overhead line fault.
4. {Long duration outage} – a state variable giving the probability of a long duration outage resulting from the overhead line fault.

The arrows in the diagram represent dependencies between nodes – and can be interpreted as causal relationships. Hence the probability of an overhead line fault is dependent of the

two parent states: {Adverse weather} and {Overload}. (The example also indicates a causal relationship or correlation between {Adverse weather} and {Overload} which might be due to the fact that when adverse weather such as snow occurs, the load increases due to increased need for space heating, increasing the probability of line overload.

The parents of any variable  $X_i$  is denoted:  $pa\{X_i\}$ , e.g.:

$$pa\{\text{Overhead line fault}\} = \{\text{Adverse weather, Overload}\}$$

The arrows in the graph represent the assertion that a variable is conditionally independent of its non-descendants given its parents in the graph. Hence, {Long duration outage} is conditionally independent of the variables {Adverse weather} and {Overload} given the variable {Overhead line fault}.

The underlying assumptions of conditional independence encoded in the graph allow us to calculate the joint probability function as in (1):

$$f(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i | pa(x_i)) \quad (1)$$

Hence, the conditional probability might be calculated, e.g. the probability of overhead line fault given the parents {Adverse weather, Overload}:

$$f(\text{Overhead line fault} | (\text{Adverse weather, Overload}))$$

As stated in [5] one of the interesting properties of the BN framework is that it can be extended to represent decisions using so-called influence diagrams. The basis for the representation is utilities, which are quantified measures for preference. That is, a real number is attached to each possible scenario in question. Exploiting the probability updating of the BN framework, it is easy to calculate the expected utility for each decision option in a domain. The example given in Fig. 3 is expanded by introducing redundancy in terms of a second overhead line L2, feeding the load in parallel with the existing line L1. The utility of the expansion is evaluated by the expected costs of energy not supplied (CENS). The expanded BN model is shown in Fig. 4:

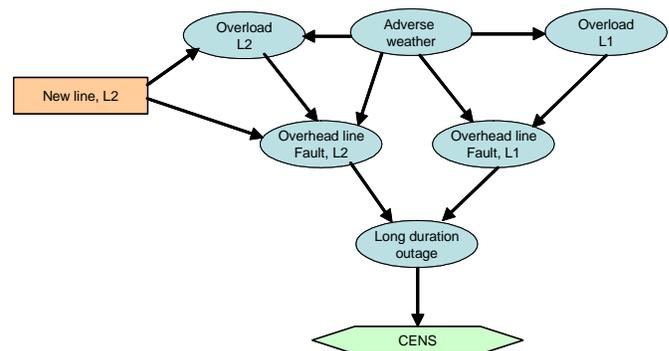


Fig.4. Risk of overhead line outage. Example of a BN including decision node "New Line L2" and utility node "CENS". Existing line is denoted L1.

The model given in Fig. 4 is entered into a BN tool<sup>2</sup> with some sample data. For simplicity, all variables in the example are discrete two-state variables, i.e. yes/no-variable. In general BN allow both discrete multi state variables and to some extent continuous variables as shown in the case study presented later in the paper.

The results based on the sample data are shown in Fig.5:

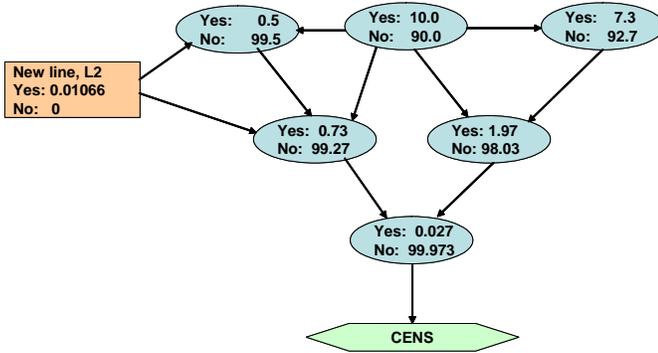


Fig. 5. Risk of overhead line outage. Example of a BN with sample data

The probabilities of the different variables are shown in the colored boxes, e.g. the probability of adverse weather is 10%. As {Adverse weather} has no parents, the probabilities shown for this node are the user input data. The utility of introducing a second overhead line is 0.01066 units compared to not introducing the new line. One unit is equal to 1 mill. NOK per year – the CENS reduction due to the introduction of a new line is approx. 11.000 NOK/year.

## V. CASE STUDY

### A. Introduction

The study focuses on the calculation of the failure probability for a MV overhead line feeder, to assess the impact of different risk influencing factors, such as vegetation, age of the feeder, wind and maintenance strategy. For this purpose, a risk analysis is carried out using Bayesian Networks. The variables are modeled based on available knowledge of the probability distribution for each random variable from historic fault statistics combined with expert judgment. The case study covers part of the overall concept presented in Fig.1 and illustrates some of the features of the BN method.

### B. Modeling

#### 1) The Bayesian Belief Network:

The Bayesian Belief Network studied is shown in Fig.6 :

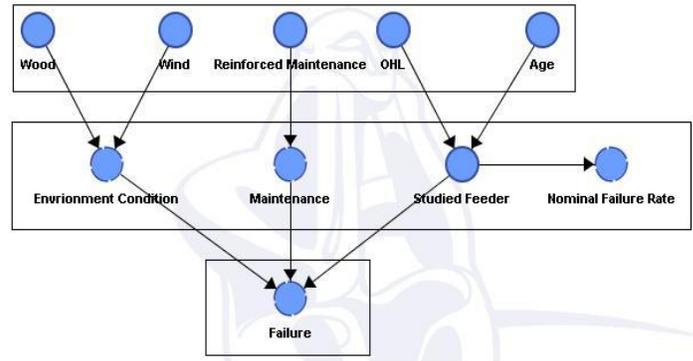


Fig.6. Overhead line model

The top level represents the input variables (sources of risk variables). The mid level gives the intermediate variables of the network and the bottom level gives the output variable (the consequence variable).

The values taken by the input variables are entered by the user depending on the type of overhead line (OHL), its age, its environment and if a reinforced maintenance strategy is carried out or not.

The intermediate random variables depend on the values of the input variables and are the parents of the consequence variable - the failure probability density of the considered feeder in its environment.

#### 2) Random Variables Description:

This section provides a discussion about the variables and the input information used in the case study. The random variables used are characterized in terms of: type (D=Discrete or C=Continuous). The features of variables used in the case study are summarized in Table I.

TABLE I  
DESCRIPTION OF THE RANDOM VARIABLES

	Random Variables	Type	Possible Values
Input Var.	Wood	D	Yes, No
	Wind	D	$\leq 30$ km/h $>30$ km/h
	OHL	D	Family 1, Family 2
	Age	D	Young, $\leq 37$ yrs Old $> 37$ yrs
	Reinforced Maintenance	D	Yes, No
Inter. Var.	Environ. Conditions	C	Index with possible values [0,6]
	Maintenance	C	Index with possible values [0,0.5]
	Studied feeder	D	Discrete table
	Nominal Failure Rate	C	Range [0,1.3] faults/100km/yr
Output Var.	Failure	C	Range [0,10] faults/100km/yr

Further details are given in the following tables:

<sup>2</sup> Netica see. www.norsys.com

a) *Input random variables:*

The input variables, are set by the user and they are specific to the studied feeder (OHL & age), its location (wooded area & wind) and if a reinforced maintenance strategy is carried out or not.

TABLE II  
FEATURES OF THE INPUT RANDOM VARIABLES

Random Variables	Remarks
Wooded area	Wood is set to Yes or No respectively if the feeder's location is in a wooded area or not. Example : 30% of the OHL feeder is in a wooded area : $P(\text{Wood} = \text{Yes}) = 30\%$ $P(\text{Wood} = \text{No}) = 70\%$
Wind	Depending on the yearly mean wind speed during one hour, if it is lower than 30 km/h or higher than 30 km/h the variable takes the value Inf 30 or Sup 30. Example. Localisation in a windy area – 40% of the length subjected to wind > 30 km/s : $P(\text{Wind} = \text{Inf } 30) = 60\%$ $P(\text{Wind} = \text{Sup } 30) = 40\%$
OHL	MV Overhead Lines construction : Example: If 55% of the length of the feeder is of family 1 and the rest is of type family 2: $P(\text{OHL} = \text{Family1}) = 55\%$ $P(\text{OHL} = \text{Family2}) = 45\%$
Age	If the feeder is older than 37 years then the variable is set to Old if not, it is set to Young. Example of a MV feeder where 80% is younger than 37 years: $P(\text{Age} = \text{Young}) = 80\% \Rightarrow P(\text{Age} = \text{Old}) = 20\%$
Reinforced Maintenance	The variable is set to Yes, if Reinforced Maintenance has been carried out on the feeder else it is set to No. Example : Reinforced maintenance used at 30% length : $P(\text{Rein.Maint} = \text{Yes}) = 30\%$ $P(\text{Rein.Maint} = \text{No}) = 70\%$

b) *Intermediate random variables:*

The intermediate random variables {Environment Conditions} and {Maintenance} correspond to a numerical coefficient (index) that impacts on the output variable's probability density function.

{System} acts as a reminder of the considered type of OHL and age whereas the {Nominal Failure Rate} gives the probability density function of the considered feeder without any external influence except ageing.

As shown in the Table III, the parameters for the probability density function of the intermediate variables depend on the possible values of the parents nodes:

TABLE III  
FEATURES OF THE INTERMEDIATE RANDOM VARIABLES

Random Variables	Parents Node	Conditional probability density function $P(\text{Intermediate R.V.}/\text{Parents' Node Value})$
System	Age, OHL	Discrete table
Nominal Failure Rate	System	$Normal(M_{\text{System}}, S_{\text{System}})$
Environment	Wood, Wind	$Normal(M_{\text{Wood, Wind}}, S_{\text{Wood, Wind}})$
Maintenance	Maint.	$Normal(M_{\text{R. Maint.}}, S_{\text{R. Maint.}})$

The intermediate variables {Environment Conditions} and {Nominal Failure Rate} have been estimated based on experience feedback. For example, to model the “nominal failure rate” the following assumptions have been made:

- a “material” failure rate has been defined as a function of the age of the component, based on recorded data-experience feedback. The failure rate probability density function is modeled by a normal distribution shown in the Fig. 7 below. This ‘material’ failure rate corresponds to the intrinsic or nominal failure rate excluding external impact.
- two categories of OHL have been considered based on failure rate recordings: Young OHL (< 37 years) and Old OHL (> 37 years) – see Fig.7

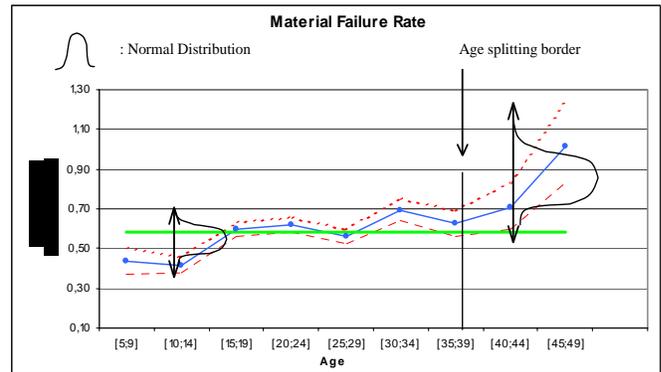


Fig.7. MV feeder material failure rate versus age

The last intermediate random variable “Maintenance” is a continuous variable (see Table III) and is modeled based on expert judgment and experience.

c) *Output random variable:*

The output of the BN model is a continuous random variable which gives the failure rate probability density function of the OHL. This output variable depends on the values of the variables of its parents nodes. The conditional probability density function is assumed to be normal with the following mean value:

$$\text{Mean}_{\text{Failure}} = \text{Mean}_{\text{System}} \times \text{Value}(\text{Env.Cond.}) \times [1 - \text{Value}(\text{Main.})] \quad (2)$$

- $\text{Mean}_{\text{System}}$  is a static value which depends on the feeder, each feeder has its own.
- $\text{Value}(\text{Env.Cond.})$  and  $\text{Value}(\text{Main.})$  are the index values of the random variables {Environment} and {Maintenance}.

The result is a normal distribution whose mean value depends on normal distributed random variables.

The output random variable reflects the combination of all the risk influencing factors for the MV overhead line feeder: external factors (wind and wood), the intrinsic vulnerability of the components and the effect of reinforced maintenance.



low level tool in the overall risk management concept or at a high level including decisions and costs/utilities directly in the BN model.

Scenario analysis is a very popular decision support tool in many organizations. Using BN, scenarios with probabilities can be defined by varying for example the input probabilities of critical risk factors for a system under study.

BNs can also be used in maintenance strategy making for groups of components. For example, when developing a maintenance strategy for coastal overhead lines, 'standardized' BNs can be constructed to account for the risk factors common to this type of components: salt pollution, wind, age of the overhead line, etc.

In addition, the BN approach can be used to improve the analysis of how combinations of different undesirable events (e.g. poor asset condition/asset ageing, severe weather events, fire, human error, geographical factors, etc), can contribute to unwanted consequences on the economics, environment, safety, security, quality of supply or on the reputation of a company.

Some if not all of these criteria are inherently uncertain, being influenced by risk factors that often cannot be controlled by the decision maker. BN is an important tool here because it offers means for handling this uncertainty. Future work will focus on applications of the BN approach together with other decision support techniques such as those in the category of Multi-Criteria Decision Aid (MCDA).

There are however several drawbacks concerning the use of BN in practice, by asset managers. Building BNs using statistical data or expert opinion can be both difficult and time consuming [5]. A BN model building comprises several steps: 1) decide what to model; 2) define variables; 3) define probabilities and the conditional relationships between the model components; 4) verify the model. This is typically an exercise that may require the intervention of a BN expert that guides the model building, asks relevant questions, and explains the assumptions that are encoded in the model to asset managers.

Particularly difficult when building up BNs is to decide what probability data to use. In the example in section V probability density functions have been used which implied:

- a choice of the probability distributions fitted to each physical phenomena: Normal, Weibull.
- Include probability information from expert judgment by choosing the appropriate probability distributions even if the information is not complete.

## VII. CONCLUSION

This paper discusses the use of BN as a relevant tool for risk assessment in distribution systems asset management. The method gives decision makers a good visualization of relationships between risk influencing factors, system states and risk consequences (variables in the graph) as well as offering statistical estimation of the variables included. Hence, in general it contributes to estimation of the answers of the

three fundamental questions of the risk analysis:

1. What can go wrong?
2. How likely is it to happen?
3. What are the consequences?

The examples and case study included, illustrates the relevance of the method in distribution system asset management when risk is involved.

## VIII. ACKNOWLEDGMENT

This paper is written as part of the reporting from the project *Risk Based Distribution System Asset Management (RISK DSAM)*. The project is sponsored by network companies in Norway and France, and the Norwegian Research Council. More information about the project can be found on the web-page: [www.energy.sintef.no/Prosjekt/RISKDSAM/](http://www.energy.sintef.no/Prosjekt/RISKDSAM/).

## IX. REFERENCES

- [1] PAS 55-1 Asset management Part 1:Specification for the optimized management of physical infrastructure assets, British Standard Institution, April 2004
- [2] IEC 60300-3-9 Dependability management Part 3 Application guide Section 9 Risk analysis of technological systems
- [3] Kaplan, S.,Garrick,B.J.: On the quantitative definition of risk, Risk Analysis, Vol.1, No. 1, 1981, pp. 11-27
- [4] Kjell Sand, Oddbjørn Gjerde, Dag Eirik Nordgård: Current risk exposure in the distribution sector -Initial study. Technical report no.TR A6576, SINTEF Energy Research, October 2007
- [5] Langseth, H: Bayesian Networks in Reliability Analysis,5th International Conference on Mathematical Methods in Reliability (MMR'07), Glasgow, 2007
- [6] C. Spelleman and L. Gauthier, "Impact of ageing on the MV underground asset reliability", in 2007 CIRED, paper 0152
- [7] G. Deleuze, H. Bertin, A. Dutfoy, S. Pierlot, O. Pourret, "Use of Bayesian Belief Networks for risk management in energy distribution", in 2004 PSAM, paper 0013
- [8] M. Bouissou and N. Thuy, "Decision making based on expert assessments: use of belief networks to take into account uncertainty, bias and weak signals", 2002, ESREL
- [9] F. Jensen, V. Finn, An introduction to Bayesian Networks, New York: Springer Verlag, 1996.
- [10] P. Naïm , A Becker, Les réseaux bayésiens, Paris: Eyrolles, 1999.
- [11] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, San Mateo, California, Morgan-Kaufmann, 1988.
- [12] M. Modarres, Reliability and risk analysis, Ed. Marcel Dekker, 1993
- [13] C.P. Robert , The bayesian choice, Springer, 2001
- [14] M. Bouissou, F. Martin, A. Ourghanlian, "Assessment of a safety critical system including software: a bayesian belief network for evidence sources", RAMS Congres, Washington, 1999.
- [15] D. Heckerman, M.P. Wellman, "Real world applications of Bayesian Networks", Communication of ACM - Association forComputing machinery, special issue, 38 (3), pp. 24-57, 1995.
- [16] ISO/IEC Guide 73 Risk Management – Vocabulary – Guidelines for use in standards. IEC/ISO 2002
- [17] Netica (software and documentation), [www.norsys.com](http://www.norsys.com)
- [18] Bayesia (software), [www.bayesia.com](http://www.bayesia.com)

## X. BIOGRAPHIES



**Kjell Sand** was born in Bodø, Norway, on October 24, 1952. He graduated as Electrical Engineer (M.Sc), 1976 from Norwegian University of Science and Technology (NTNU, former NTH) and received his Ph.D degree from NTNU in 1987. He has been with SINTEF Energy Research since 1978 working mainly with power system planning and analysis, quality of supply (reliability, voltage quality), monopoly regulation and benchmarking. Sand holds a position as associate professor at the Department of Electric Power Engineering, NTNU.



**Maria. D. Catrinu** was born on December 24, 1975 in Campulung, Romania. She received her BSc and MSc degrees in Electrical Power Engineering and Power Systems Management from 'Politehnica' University of Bucharest in 1999 and 2000. She holds a PhD degree in Power Engineering from the Norwegian University of Science and Technology (NTNU), from 2006. She has been with SINTEF Energy Research, Department of Energy Systems since 2006, working mainly with energy systems planning, risk analyses and decision support.



**Gerd H. Kjølle** was born on February 25 1958 in Oslo, Norway. She received her MSc and PhD degrees in electrical engineering from the Norwegian University of Science and Technology (NTNU, former NTH), in 1984 and 1996 respectively. She has been with SINTEF Energy Research since 1985. She is presently a Senior Research Scientist, working mainly with reliability and interruption cost assessment, energy systems planning and risk analyses.



**Sylvie Bonnoit** was born in Grenoble, France, in 1975. She graduated as Electrical Engineer from the National Polytechnic Institute of Grenoble (INPG) in 2000. She entered EDF R&D in 2000 to work about the protection plans in the transmission networks. And now, she works on maintenance, asset management and risk studies in the electrical networks.



**Jean Aupied** is a senior engineer at EDF R&D. He was born in France in 1951, and received a M.S. degree from engineering school "Ecole Nationale Supérieure de Arts et Métiers" in 1975. He joined Electricité de France (EDF) R&D Division in 1976 to perform probabilistic safety assessments (PSA) for nuclear power plants, and dependability, maintenance and risk studies in electrical network field. He has published a book on reliability of equipment and gives RAMS lectures in several universities.