

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Using dynamic Bayesian networks to model technical risk management efficiency

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1654421> since 2018-02-26T10:33:21Z

Published version:

DOI:10.1002/qre.2186

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

This is the author's final version of the contribution published as:

[Halabi, Anan; Kenett, Ron S.; Sacerdote, Laura, Using dynamic Bayesian networks to model technical risk management efficiency, QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL, vol. 33 (6), 2017, pagg.1179-1196, DOI:10.1002/qre.2186]

The publisher's version is available at:

[<http://onlinelibrary.wiley.com/doi/10.1002/qre.2186/abstract>]

When citing, please refer to the published version.

Link to this full text:

[<http://hdl.handle.net/2318/1654421>]

This full text was downloaded from iris-AperTO: <https://iris.unito.it/>

Using Dynamic Bayesian Networks to Model Technical Risk Management Efficiency

Anan Halabi

Department of Mathematics "G. Peano", University
of Turin, Italy.

Ron S. Kenett

Department of Mathematics, University of Turin,
Italy and Neaman Institute, Technion, Israel

Laura Sacerdote

Department of Mathematics "G. Peano", University
of Turin, Italy.

Abstract

The objective of this paper is to present a mathematical model helping decision makers achieve optimum efficiency in risk management of product development. The optimum we are seeking considers qualitative data derived from expert opinions and quantitative information on project characteristics. The mathematical model proposed here aims at integrating data from these sources to identify opportunities for decreasing product risk. Reduction of overall product risk, before product release to production, is an indicator of the efficiency of the risk management intervention. Acceptable risk targets vary according to industry type, organization characteristics, regulations, etc. In general, the risk management process consists of identification of risks, analysis of risks, risk control and feedback. Here, we propose a mathematical approach to risk management using dynamic Bayesian networks for evaluation of product risks during the development period. The properties of the model are assessed using two validation methods: k-fold cross validation and leave-one-out techniques. Mathematical imputation methods, like multivariate normal imputation are invoked to deal with missing data. In addition, sensitivity analysis

is performed to assess the uncertainty embedded in the parameters derived from the dynamic Bayesian network.

Decision makers should consider the overall risk in product development estimated by this mathematical model. It may help to determine whether to release a product for Beta testing or to conduct additional activities to reduce the overall risk level before customer shipment. In addition, the model may be used for prediction purposes as it provides an estimate of the expected risk at time $t+1$ based on the level of risk at time t .

Key-words Bayesian networks; Dynamic Bayesian network; Risk; Expert Subjective Assessment; What-if scenario.

1 Introduction

Products and processes in product development are becoming increasingly complex. Consequently, errors associated with product development increase. This drives firm to search for appropriate procedural models such as Munich's Procedural Model (MPM) which proposes risk management activities to ensure the achievement of the set of goals in the product development process¹. In the literature, risk management models are defined with different perspectives. While several of them describe the risk management in a context of processes of project development^{2,3} others emphasize the technical aspects. Technical risk management focuses on the risks related to the performance of the product itself⁴. Reliability risk management can be seen as a subset of technical risk management focusing on safety and reliability issues of products^{5,6}. Strategic risk management deals with the question of integrating different risk management activities into an overall enterprise risk management approach and providing central monitoring and early warning capabilities⁷.

Risk management models includes the following phases⁸:

- Risk Identification: Potential risks are detected. This includes collecting the preliminary information available for every potential risk, including the rationale for the identification.
- Qualitative Risk Analysis: A step that further deepens the understanding of a potential risk, without assigning any numerical judgment. This differs from the perception in some of the literature, especially.
- Quantitative Risk Analysis: Numerical values are assigned to a risk's probability of occurrence, magnitude of impact, and its timeframe. It can include mathematically exact models, as well as other types of quantification, for example the assignment to a certain (numerically specified) category based on team discussions.
- Risk Prioritization: The quantitatively described risks are prioritized. The prioritization process can be conducted along a multitude of different measurement or priority systems, taking one or more of the quantified risk attributes into account.
- Execution of Actions: This step is not considered to be "owned" by the risk management process, but by the line organization responsible for the risk.
- Monitoring of Risks: It can be aimed at the risks themselves, or at the performance of the risk management process. The goal is to provide a transparent and current description of the risk situation and to trigger impulses to inform decision makers of significant changes.
- Aggregation of Risks: In the Aggregation step, single risks are aggregated to the next higher level. This step is of central importance if an enterprise-wide integral risk management system is to be established over more than one hierarchical level.

Risk analysis has evolved over the years from addressing only the technical aspect of a system to that which covers additional aspects such as human and organizational factors affecting the system. This evolution is particularly apparent in critical systems such as nuclear power plants, oil rigs and chemical processes where regulations become more and more demanding according to safety rules⁹. In this paper, we propose a systematic model for assessing technical risks during the development cycle of a product or system. This model combines external factors related to the development cycle process. Furthermore, it treats the total accumulated risks of a product instead of managing risks by ranking priority. While the current risk management model used by the company illustrated in the case study, uses top ranking risk method², the proposed model provides assessment of the overall risk of a product which supports decision makers at different check points during the development process. In addition, the new approach provides indicators to control the level of the risks via mitigation plans designed to reduce the level of the recognized risks over the product development.

The technical risk discussed in the model refers to the uncertainty that a product design will not satisfy technical requirements and the consequences thereof. The amount of performance risk associated with any technical performance measurement depends on two factors: 1) the number of possible outcomes, cases, or situations that fail to meet requirements and 2) the consequence or impact of each⁴.

The systematic model proposed here considers qualitative data on risk assessment derived from expert opinions and quantitative information based on a project's characteristics.

The model examines both the probability and the severity of potential risks. We estimate the progression of each risk during product development using a dynamic Bayesian network. Other methods proposed in the literature include ARIMA, ARMAX¹⁰. The dynamic Bayesian network (DBN)^{11,12} is an

expansion of a Bayesian Network (BN)^{13,14}. It enables modeling of temporal relationships among variables at different points of time. In addition, it enables incorporation of external variables such as project characteristics as well as flexibility in generating predictions.¹²

Useful quantities to describe measurement and assessment of risk are¹⁵:

1) The combination of probability and severity of risk impact;

2) The triplet (S_i, P_i, C_i) , where S_i is the i -th scenario, P_i is the probability of that scenario, and C_i the consequence of the i -th scenario, where $i = 1, 2, \dots, N$;

3) The triplet (C', Q, K) , where C' is some specified consequence, Q is a measure of the uncertainty associated with C' (typically probability) and K is the background knowledge supporting C' and Q (which includes a judgement of the strength of this knowledge);

4) Expected consequences (damage, loss) of a risk event:

1. Expected number of fatalities in a specific period or the expected number of fatalities per unit of exposure time;

2. The product of the probability of the hazard occurring and the probability that the relevant object is exposed given the hazard, and the expected damage given that the hazard occurs and the object is exposed to it (the last term is a vulnerability metric);

3. Expected disutility.

The term *consequences* used here refers to the outcome of an event.

The traditional approach to risk analysis is based upon the principles and methods of probability and classical statistics (Nilsen and Aven¹⁶). The probability of an event is defined as the relative frequency of that event when the condition from which it develops is hypothetically repeated an infinite number of times¹⁷. Furthermore, Aven¹⁵ states that the way in which a risk is understood and described, strongly influences the way in which it is analyzed and hence may have serious implications for risk management and decision-making.

In today's world, time-to-market is becoming shorter and shorter. Consequently, development cycles become shorter with little data for estimation of probabilistic inputs of the risks embedded in the product under development. As a result, assumptions that allow the use of available data are forced and supplementary information like expert opinions often substitutes the traditional data-driven analysis¹⁸. Furthermore, risk-data can be assessed in the context of a development process via integrated models of assessing it and its probability when we have limited information coming from various activities of development¹.

However, since the goal of technical risk management is often limited to prioritizing activities of product development, organizations may on one hand miss the global perspective of a product's potential overall risk while ignoring risks of low priority and on the other hand may ignore the effect of external variables on risk assessment. One of the innovations in this work is the combination of mathematical risk modeling and systematic risk management of engineering products, which overcomes these deficiencies. The proposed approach can be used both as an indicator of the efficiency of risk management and as a calibration tool for adjusting and optimizing activities designed to achieve acceptable levels of risk during product development. All this requires a systematic approach with well-defined levels of risk,

¹ Halabi, A., Kenett, R.S. and Sacerdote. L. (2016). Modeling the relationship between reliability assessment and risk predictors using Bayesian networks and multiple logistic regression model, *quality Engineering*, submitted paper

with customers informed on the value of the risks associated with the produced item, in particular upon receiving the first items from the production line.

Section 2 in this paper details the methods used to evaluate the mathematical model; in Section 3 we present the mathematical formulation of the risk model and the model properties; Section 4 illustrates an application of the model to a real-life case; Section 5 is a discussion of the results and conclusions

2 Methods

This section presents the combined mathematical methods used to model risk and efficiency of risk management. The main methods which used to develop the model are Bayesian network, dynamic Bayesian network and 'What-if' scenario. While Bayesian network is designed to model the static relationship between risk variables and project characteristics, dynamic Bayesian network is used to model temporal relationship, specifically between risk variables on different time points.

Furthermore, we use supporting methods to ascertain the properties of the model including Cross validation methods such as K-fold and Leave-one-out¹⁹, sensitivity analysis and imputation methods. For treating missing values in the data set, imputation methods are used²⁰; for purposes of prediction and inference, the methods of a "What if" scenario^{21,22} are used as well as the measure of Euclidean distance²³ to assesses the strength of the relationship among variables and for learning the inaccuracy of the parameters of the model sensitivity analysis is performed.

2.1 Bayesian networks

Bayesian networks (BN) implement a graphic modeling structure known as directed acyclic graphing (DAG) that enables effective representation and computation of joint probability distributions (JPD) over a set of random variables.²⁴ The structure of a DAG is defined by its nodes that represent random variables and its directed arcs that represent direct dependencies among the variables. The nodes are drawn as circles labeled by variable names while the arcs are drawn as arrows between nodes. A particular arc from node X_i to node X_j represents a statistical dependency between the corresponding variables indicating that a value taken on by variable X_j depends upon the value taken on by variable X_i . Node X_i is then referred to as the 'parent' of X_j and, similarly, X_j is referred to as the 'child' of X_i . An extension of these genealogical terms is often used to define the sets of 'descendants', i.e., the set of nodes from which a specific node can be reached via a direct path from the original parent node.

The DAG structure guarantees that no node can be its own ancestor (parent) or its own descendent (child). Such a condition is of vital importance in computing the joint probability of a collection of nodes since it allows the introduction of useful factorization.

Although the arrows represent direct causal connections among variables, the reasoning process can operate on a BN by propagating information in any direction. A BN reflects a simple statement of conditional independence, namely that each variable, given the state of its parents is independent of its non-descendants in the graph. This property can be applied to reduce, sometimes significantly, the number of parameters required to characterize the variables in a JPD. This reduction provides an efficient way to compute posterior probabilities given the evidence in the data ^{24,25}.

The DAG structure is often considered as the qualitative part of the model, but it is also necessary to estimate the quantitative parameters of the model. We perform this task by calculating the local conditional probabilities for each node. For discrete random variables, we list these conditional probabilities in a table reporting the local probability of a child node at each of the feasible values for each combination of values of its parents. These tables uniquely determine the joint distribution of a collection of variables.

Formally, for a BN, we introduce is an annotated graph B , that represents a joint probability distribution over a set of random variables, \mathbf{V} .²⁶ A network is defined by pair, $B = (G, \Theta)$ where G is the DAG whose nodes X_1, X_2, \dots, X_n represent random variables and whose arcs represent direct dependencies among these variables. The graph G encodes assumptions of conditional independence, where variable X_i is conditionally independent from its non-descendants given its parents in G . We denote the set of parents π_i . The second component, Θ denotes the set of parameters of the network. This set contains the parameters, $\theta_{x_i|\pi_i} = P_B(x_i|\pi_i)$ for each realization x_i of X_i conditioned upon π_i , the set of parents of X_i in G . Accordingly, B defines a unique joint probability distribution over \mathbf{V} , namely, the set of parameters of the network. This set contains the parameters for each variable (X_1, X_2, \dots, X_n) in the network:

$$P_B(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_B(X_i|\pi_i) = \prod_{i=1}^n \theta_{x_i|\pi_i}. \quad (1)$$

2.2 Dynamic Bayesian networks

Dynamic Bayesian networks (DBN) model temporal relationships among variables. While a Bayesian network (BN), only represents probabilistic relationships among a set of variables at some point in time, a DBN relates the value of some variable to its value at previous points in time as well as to the values of other variables at previous points in time.

To illustrate the structure of a DBN, let us assume that the variables Z_1, Z_2, \dots, Z_n , defined on a BN, change over time. Let us consider $Z_i[j]$ to be a random variable representing the value Z_i at time j for $j=0, \dots, J$ and let

$$Z[j] = \begin{pmatrix} Z_1[j] \\ \vdots \\ Z_n[j] \end{pmatrix}. \quad (2)$$

For all j , each variable $Z_i[j]$ is defined on a space which depends on index i and is called the space of Z_i (for example, a space for discrete variable correspond to two states: low and high). A Dynamic Bayesian Network is defined through the variables that specify the J random vectors $Z[j]$ determined by the following specifications:

1. The initial BN, consisting of:

- a) DAG G_0 containing the variables in $Z[0]$
- b) an initial probability distribution, P_0 of these variables.

2. A transitional Bayesian BN, consisting of:

- a) transition DAG, G , containing the variables $Z[j]$ and $Z[j+1]$ (for example, level of risk at consecutive months j and $j+1$)

b) a transitional probability distribution P_{\rightarrow} which assigns conditional probabilities to each value of $Z[j+1]$ given every value of $Z[j]$, for each realization $z[j+1]$ of $Z[j+1]$ and $z[j]$ of $Z[j]$ that we specify as:

$$P_{\rightarrow} = (Z[j + 1] = z[j + 1] | Z[j] = z[j]). \quad (3)$$

3. The DBN containing the variables that each of them includes, i.e. the J random vectors consisting of:

a) DAG composed of the DAG G_0 and for $0 \leq j \leq J-1$ the DAG G_{\rightarrow} evaluated at j .

b) Joint probability distribution:

$$P(z[0], \dots, z[J]) = P_0(z[0]) \times \prod_{j=0}^{J-1} P_{\rightarrow}(z[j + 1] | z[j]). \quad (4)$$

2.3 Cross validation (CV) - k-Fold Cross-Validation

This method involves randomly dividing the set of observations into k groups or folds of approximately equal size. Initially, the first fold is treated as the validation set and the method is subsequently applied on the remaining $k - 1$ folds. The Mean Squared Error ($MSE_{(1)}$) is then computed on the observations in the held-out fold. This procedure is repeated k times; each time a different group of observations is treated as the validation set. This process results in k estimates of the test error, $MSE_{(1)}, MSE_{(2)}, \dots, MSE_{(k)}$.

When applying k -fold, the estimated CV is computed by averaging these values:

$$CV_{(k)} = \frac{1}{k} \times \sum_{i=1}^k MSE_i, \quad (5)$$

Where

$$MSE_{(i)} = \frac{1}{n} \times \sum_{i=1}^n (y_i - \hat{f}(x_i))^2, \quad (6)$$

Here n is the number of observations in the sample, y_i is observation i and $\hat{f}(x_i)$ is the prediction that \hat{f} gives for the i -th observation.¹⁹

2.4 Cross validation - Leave-one-out

The method proposed by James et al.¹⁹ involves splitting the set of observations into two parts. However, instead of creating two subsets of comparable size, a single observation (x_1, y_1) is used as the validation set

with the remaining observations $\{(x_2, y_2), \dots, (x_n, y_n)\}$ of the training set. The statistical learning method is applied to the $n - 1$ training observations and a prediction, \hat{y}_1 is made for the excluded observation using its value, x_1 . Since (x_1, y_1) was not used in the fitting process, $MSE_{(1)} = (y_1 - \hat{y}_1)^2$ provides an approximately unbiased estimate for the test error. But even though the $MSE_{(1)}$ is unbiased, due to its being based upon a single observation (x_1, y_1) , it is extremely variable and hence it is a poor estimate. The procedure is repeated by selecting (x_2, y_2) for the validation data and the statistical learning procedure is then applied to the remaining $n - 1$ observations $\{(x_1, y_1), (x_3, y_3), \dots, (x_n, y_n)\}$ to get $MSE_{(2)} = (y_2 - \hat{y}_2)^2$. As in k-fold methodology, this procedure is repeated n times to produce n mean squared errors, $MSE_{(1)}, \dots, MSE_{(n)}$. The LOOCV (Leave-One-Out Cross-Validation) estimate for the test MSE is the average of these n estimates of test error:

$$CV_{(n)} = \frac{1}{n} \times \sum_{i=1}^n MSE_i. \quad (7)$$

2.5 Multivariate normal imputation

Missing values are an issue in a substantial number of statistical analyses, with most deleting or ignoring the observations with missing values. However, it is possible to resolve incomplete cases by using Multivariate Normal Imputation^{27,28}, in which random values from the multivariate normal distribution are substituted for the missing values.

The algorithm uses least squares imputation. Entries in the covariance matrix are computed by using all non-missing values for each variable along the diagonal elements while off-diagonal elements are computed using all non-missing values for both variables. In cases where the pairwise inverse

is singular, the algorithm uses minimum norm least squares imputation based upon the Moore-Penrose pseudo-inverse. (See Appendix A- report of imputation)

2.6 Euclidean distance

Euclidean distance is a measure used to evaluate the distance between two points in Euclidean space. In a DBN, it is used to assess the influence of one variable upon another by measuring the distance between the two discrete probability distributions of these variables.

For this purpose, let us define two discrete probability distributions, P and Q :

$$P \in \{(p_1, p_2, \dots, p_n) | p_i > 0, \sum_{i=1}^n p_i = 1, n > 1\} \quad (8)$$

$$Q \in \{(q_1, q_2, \dots, q_n) | q_i > 0, \sum_{i=1}^n q_i = 1, n > 1\} \quad (9)$$

Then the Euclidean distance is defined as:

$$E(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}. \quad (10)$$

If P and Q are two points in some N -dimensional space, (10) calculates the actual distance between the two points. When used to measure the distance between two discrete probability distributions, its range is $(0, \sqrt{2})$ since the sum of all elements is always equal to one. Euclidean distance is a symmetric measure. It is possible to rescale the Euclidean distance to obtain a quantity in the range $[0, 1]$:

$$E_{norm}(P, Q) = \frac{E(P, Q)}{(\sqrt{2})}. \quad (11)$$

2.7 The "What if" scenario in Bayesian Networks

"What if" scenario is an intervention method used to determine causality in Bayesian networks. Traditionally causality is based on application of randomized trials, where the design of the trial aims to identify the effect of an intervention.²⁹ In general, causality has been studied from both "probabilistic" and "mechanistic" points of view. In the probabilistic view, the causal effect of an intervention is determined by comparing the evolution of the system in the two instances of presence/absence of intervention. The mechanistic view focuses on understanding the mechanisms determining how specific effects arise. Causal inference on Bayesian networks has two types of interventions: "structural" and "parametric". The "structural" intervention makes the intervened variable independent of its causes and therefore changes the causal structure of a system before and after intervention. The "parametric" intervention affects the parameterization of the conditional probability distribution of the intervened variables on its parents, while it still leaves the causal structure intact. For more on sensitivity analysis of BN see Cugnata et al.²⁹

In causal Bayesian networks, a 'do' operator that separate a "structural" intervention from a "parametric" one can define the effect of any intervention. Pearl ³⁰ states that the mathematical operator called 'do' simulates physical interventions by deleting certain functions from the model, replacing them with constant $X=x$, while keeping the rest of the model unchanged. For more details on causal calculus see Pearl ³⁰. The 'do' operator makes it possible to conduct "what-if" scenarios even if counterfactuals cannot be directly tested, as it happens in the presence of nonexperimental data.

2.8 Sensitivity Analysis (parameters inaccuracies) in Bayesian networks

The robustness of the output probabilities of a Bayesian network can be investigated by performing a sensitivity analysis on the network. In Bayesian

network, a sensitivity analysis describes the relationship between the probability parameters of the network and its posterior marginals. For more details of the mathematical formulation see Kjaerulff and van der Gaag^{31,32}. In GeNIe SW we used the algorithm proposed by Kjaerulff and van der Gaag^{31,32} to perform the analysis.

In shortly, the algorithm calculates efficiently a complete set of derivatives of the posterior probability distributions over the target nodes over each of the numerical parameters of the Bayesian network. These derivatives give an indication of importance of precision of network numerical parameters for calculating the posterior probabilities of the targets. If the derivative is large for a parameter, then a small deviation in parameter may lead to a large difference in the posteriors of the targets. If the derivative is small, then even large deviations in the parameter make little difference in the posteriors.

3 The risk model

3.1 Mathematical formulation of the problem

The model refers to risk data were collected over K projects by different teams at different time intervals. The considered products were developed in series. Furthermore, data include project time cycles (in units of days) in order to determine their effect upon efficiency of risk management.

Let us introduce a matrix, R_k of the k -th project (a process of product development) identified by engineers in the frame of technical risk management during product development. Let us assume that there are monthly observations made during J months and that n different risks are under consideration. R_j Indicates variable j in matrix R_k . Elements, $r_{i,j,k}$ of matrix R_k represent the i -th risk in the j -th month of k -th project. Where: $j = 1, \dots, J$; $i = 1, \dots, n$ and $k = 1, \dots, K$.

In real-world development, organizations use technical risk management to prioritize risk mitigation plans. In this paper, we estimate the overall risks entailed in developing a product using data characterized by R_k and we apply a dynamic Bayesian network to obtain the dependencies among variables R_j which indicate the overall risk at month j .^{11,12,33} The probabilistic model for the dependencies embedded in the risk variables is:

$$p(R_1, \dots, R_j) = p_1(R_1) \times \prod_{j=1}^J p_{\rightarrow}(R_{j+1} | R_j), \quad (12)$$

where R_0 is a random variable indicating the risk in the first month and p_0 is its initial probability. The transition probability p_{\rightarrow} assigns a conditional probability to every value of the variable R_{j+1} , given each value of R_j .

In the problem under discussion, the dynamic Bayesian network models only the risk variables while we describe the other variables, like project characteristics, through a regular Bayesian network.^{21,22} Furthermore, we characterize the project through two variables A and Z : A is a random variable indicating the specific technical project and Z is a continuous random variable accounting for the time required for development. Then, the joint probability of project characteristics and risks is:

$$p(A, Z, R_0, R_1, \dots, R_{j-1}) = p(Z) \times p(A|T) \times p_0(R_0|A) \times \prod_{j=0}^{j-1} p_{\rightarrow}(R_{j+1} | R_j, \pi_{j+1}). \quad (13)$$

Two ordinal variables have been defined for the risk dimensions: P accounts for the probability of occurrence and C for the severity of risk scenario. Each ordinal variable has 5 levels: from low (1) to high (5).

3.2 The mathematical model and its properties

As defined earlier, risk assessment is a combination of probability and severity of scenarios. Here we subdivide the presentation of the risk model into two parts: Section 3.2.1 models only the probability of the progression of risks for different products over time, Section 3.2.2 describes the progression of severity of risk over time, and then in Section 3.2.3 we present a risk as one model. The risk data analyzed come from the development cycle (which starts from the conception of a product to its release for production) only. In this model we assume that time censoring does not affect the assessment score.

3.2.1 Dynamic Bayesian Network modeling - uncertainty of risk (probability)

To model the risk evaluation over time, for different products, the DBNs considers the variables of product A , the length of the development cycle Z , and the probability of risk scenario P . We assign nodes to the DBN accordingly to these variables. We use the clustering algorithm of GeNIe SW to build the CPDs for the variables in the model. The output of the process is a global joint probability connecting the (A, Z, P) variables:

$$P(A, Z, P_0, P_1, \dots, P_{J-1}) = p(Z) \times P(A|Z) \times P(P_0|A) \times \prod_{j=1}^{J-1} P(P_j|P_{j-1}, A). \quad (14)$$

We evaluate the properties of the model using validation methods like k-Fold and Leave-one out. For both methods, we compute the mean square error (MSE). Figure 1 describes the GeNIe SW worksheet. It contains the nodes defined in the model, the arcs connecting between them and two plates. One of them describes the static and the other the temporal relationships (a number on the arcs).

Figure 1- DBN for probability of risk - GeNIe SW

The detailed process of building the DBN includes the following steps: at the first step, we define the structure of the model and set the arrows between the nodes (Figure1). The second step entails defining the properties of each node: the number of states and the type of each node (chance, ordinal etc`). The third step extrapolates the parameters from the real data collected. The real data may require treatment such as imputation of missing values and discretization of variables.

3.2.2 Dynamic Bayesian Network modeling - Severity of risk

The structure of the severity of risk scenario model is similar to that developed for the probability of risk scenario discussed in Section 3.2.1, except that here we consider the variables (A, Z, C) where C indicates the severity of the risk. To build the CPDs for the variables in the model we use the Clustering algorithm of GeNIe SW. Equation (14) still holds but with C substituting P .

3.2.3 Dynamic Bayesian Network modeling - risk

We introduce here a new variable, M , describing an expected outcome due to risk occurrence. The value of this variable is the product of the probability score by the severity score. We substitute M for the variable C in the models for probability and severity, respectively. Now the model appears as:

$$P(A, Z, M_0, M_1, \dots, M_{J-1}) = p(Z) \times P(A|Z) \times P(M_0|A) \times \prod_{j=1}^{J-1} P(M_j|M_{j-1}, A). \quad (15)$$

4 Case study results

We illustrate the proposed methodology by applying it to a real-life case. Risk data were collected during product development cycle of an electronic measurement product, which combines high precision mechanics with sophisticated electronic and optical components. The organization uses data for prioritizing activities related to the analyzed risks. In the current case, we augmented this data by applying a DBN both to assess the overall risk and to monitor the efficiency of risk management over time, i.e. including identification of risks, analysis, control and feedback. An engineer from the team involved in product development worked with the optical, mechanical and electronic teams to identify, collect and assess risks. The assessment of risks included evaluation of the probability and severity of each risk. Evaluations of probability and severity were measured on a 1-5 scale where 1 is a very low value and 5 is very high. Mitigation plans were defined for top-ranked risks. Based upon results determined from execution of these plans, risk assessment was updated. This process entails treating risks of high priority while ignoring the overall risk and the efficiency of the risk management process. We employ a dynamic Bayesian network to model overall risk and we use it for assessing the risk management process. In the following sections, we describe the data structure and the combined dynamic model used for monitoring risk and its features.

4.1 Data description

4.1.1 Risk assessment over time:

Data are uploaded onto double-entry spreadsheets. Each column corresponds to Probability risk (P), Severity risk (C) and Risk Loss (Combined Probability and Severity), respectively.

Risk assessment data were collected from six projects in the organization (technical content has been omitted because of its proprietary nature). The product development cycle was 14 months long.

In the next 2 Sub-Sections we summarize the descriptive statistics of risk variables and then in the followed Sub-Sections we build the dynamic Bayesian networks.

4.1.2 Severity of risk

Table 1: Descriptive statistics of severity of risk

The average of severity at the first two slices ($j = 0$ and $j = 1$) are almost medium (value 3.2), its value increases to 3.6 on the following next slices from $j = 2$ up to $j = 8$. Then a trend of decrease is observed on the following slices from $j = 11$ up to $j = 13$. The variance of the severity increases at slice $j = 3$ and a trend of decrease is observed after slice $j = 5$.

4.1.3 Probability of risk

Table 2: Descriptive statistics of probability of risk

The average probability of risk is medium or less on all slices except on slice $j = 4$. The variance at Slice $j = 4$ is higher compared to other slices.

GeNIe SW was used to determine the BDN parameters. We present these results in Sections 4.2 - 4.4. To investigate the inaccuracy model parameters, we perform sensitivity analysis but we illustrate the use of this technique only in the case of the overall risk model (4.4), as it

includes the uncertainty determined by estimating both the parameters of severity and probability.

4.2 Dynamic Bayesian network model for probability of risk

The estimated parameters of the Dynamic Bayesian Network include the following nodes: the project variable A , the length of the development cycle Z and the probability P of risk over 14 months based upon $N = 2,767$ observations. We used the GeNIe SW Clustering algorithm to build the probabilistic model (https://dslpitt.org/genie/wiki/Main_Page).

The first step in building such a model is estimating the parameters of the dynamic Bayesian network via calculating the conditional probability tables for all the nodes in the Network. The following examples illustrate the calculation of two types of conditional probability tables. The first example presents a static conditional probability between project node and probability of risk node at $j=0$. The second example presents temporal conditional probability between probabilities of risk at time j given the probability of risk scenario at $j-1$.

Example 1:

We consider two nodes: the probability at $j=0$, P_0 and the project variable, A . In this case, the project node is the 'parent' and the probability of risk scenario at $j=0$ is the 'child'. The conditional probability expression is:

$$P(P_0|A) = \frac{P(P_0 \cap A)}{P(A)}. \quad (16)$$

To illustrate the computation of the P - probability, we set the probability at $j =0$ equal to zero (low probability) and the project equal to 0 (Project A).

$$P(P_0 = 0|A = 0) = \frac{P(P_0 = 0, A = 0)}{P(A = 0)}. \quad (17)$$

The projects in the sample are: A, B, C, D, E and F. we encode it with the corresponding numeric values: 0,1,2,3,4,5.

To calculate the empirical probability from the data for this realization (Cell 1 in Table 3 with the highlighted rectangle) we use the ratio

$$= \frac{\text{number of observation for which } P_0 = 0 \text{ and } A = 0}{\text{number of observation for which } A = 0}$$

Table 3: Frequency of probability of risk, P_0 conditioned on project, A

The second step in this example, is to calculate the marginal distribution based upon the conditional probability table (table 3) obtained from the data. In Table 4 are reported the values of the marginal distribution for the probability of risk scenario at $j = 0$.

Table 4: Marginal distribution for the node at $j=0$

Example 2:

Table 5 shows the transition probabilities between probability of risk at time $j = 2$ given the probability of risk scenario at $j = 1$. For example, the probability to have risks with high probability at $j = 2$ given risks with low probability at $j = 1$ is about 85.7% on project B while in project A is 100%. We create analogous tables for probability of risk at time j given probability of risk at $j-1$.

Table 5: Transition probabilities between probability of risk at time $j = 2$ given the probability of risk scenario at $j = 1$

Figure 2- DBN structure and parameters expressed by marginal distributions for probability ($j=0...4$)

In Figure 2 we present the probability of risk scenario up to $j = 4$ while the next nine steps are presented in Figure 3. Each probability node has two states: low and high. We group different levels of risk together, putting values 1-3 in State 0 and assigning values 4 and 5 to State 1, which corresponds to a high probability of risk.

Figure 3- DBN structure and parameters expressed by marginal distributions for probability ($j=5...13$)

In this case $N = 216$ and $J = 14$, and the resulting mathematical model is:

$$P(A, Z, P) = p(Z) \times P((A|Z)) \times \prod_{i=1}^{216} P(P_{j=1}^i | A) \times \prod_{j=2}^{14} \prod_{i=1}^{216} (P(P_j^i | Pa(P_j^i))) . \quad (18)$$

The frequency of risks characterized by high probability exhibits an alternation of increasing and decreasing behaviors: it increases up to $j = 4$, then at $j = 5$ it decreases to 6%, but then becomes 11% at $j = 6$ and then decreases up to $j = 11$, but at $j = 12$ it increases once again. The probability of risk scenario fluctuation is explained when creating new knowledge relating to existing problems or detecting new problems. This profile is most realistic in risk analysis⁴.

As indicated in the Introduction, intense competition in the market accelerates companies to seek ways to shorten development cycle time to get to market first. Consequently, we examine the demand for shorter development time upon the model by using the methodology of a "What-if" scenario. For example, setting the length of the development cycle to a value of less than one year shows that the frequency of risks with high probability increases from zero to 6% at $j = 0$. In addition, a sharp increase is noticed in $j = 3$ as compared to the previous time.

Figure 4- Scenario of cycle development equal to 1 year as compared to a regular
DBN

In Figure 4 the marginal distribution of the DBN obtained from data is compared to the projected marginal distribution of the DBN conditioned to a one-year development cycle. Results show higher instability under the 1-year condition as compared to the DBN obtained from the data. Product managers to take balanced decisions relating time of development cycle and level of stability of risk can use such predicted inference. Furthermore, it could be a starting point for learning how to move the risk of the system to a more stable point focusing on the specific activities which most contribute to the instability under this new condition.

For prediction matter, we analyzed the accuracy of the model using k -fold and leave-one-out methods. The total accuracy of the model results to be 94.4%. We report the detailed accuracy of each node in Fig. 7:

Figure 5- Summary of the accuracy of the probability model

The model shows high accuracy in predictions for all nodes. While the model has great capability in predicting cases of low probability, it is "less accurate" to predict risks with high probability. The source of this result is imbalanced data set with low representation of risks with high probability. To solve this problem, we apply importance sampling algorithm called self-importance sampling (SIS)³⁴ and ROSE procedure³⁵ which take in account for this bias. We use the ROSE procedure for treating the target node ($t=13$). A report of the detailed accuracy is presented in Fig. 6. The capability of the model to predict risks with high probability is improved. The accuracy of the target node (at $t=13$) increases from 0% to 87%.

Figure 6- Accuracy of the model when using oversampling method and self-importance sampling algorithm

4.3 Dynamic Bayesian network model for severity of risk

In this section, we initially formulate the mathematical model, and then present inference results derived from it and finally summarize the accuracy of the model.

The mathematical model of the DBN is:

$$P(A, Z, C) = p(Z) \times P(A|Z) \times \prod_{i=1}^{216} P(C_{j=1}^i | A) \times \prod_{j=2}^{14} \prod_{i=1}^{216} (P(C_j^i | Pa(C_j^i))). \quad (1)$$

Figure 7 describes the development of severity over time. The high frequency for low severity risks is observed in the first two steps ($j=0, 1$), then their frequency decreases in $j=3$ and $j=4$. Instead, risks of high severity increase in correspondence to these steps.

Figure 7- DBN structure and parameters expressed by marginal distributions for severity (Steps 0 to 6)

In addition to describing the development of severity over time (figure 7), we determine the amount of influence of one node on the others. It is done via visualizing the thickness of the arcs between the nodes. The thickness of the arcs is determined by calculating the Euclidean distance measure between the distributions of any connected nodes.²⁵ The strength of dependency values between nodes are much higher at the borders than to the center of the DBN (the thickest arcs are circled in Figure 8). The meaning of colors in the graph is the

following: green indicates a positive influence, red a negative and blue that no sign is determined. The influence between nodes in the edges of the DBN are positive.

Figure 8- Strength of influence between severity nodes

Furthermore, we create risk DBN per product via conditioning the DBN each time on different product. The analyzed three products exhibit similar fluctuations for the severity of risk scenario (Figure 11). These fluctuations though very sharp are consistent being similar between different projects.

Figure 9- Comparison of severity values for 3 products out of 6

Product B shows the highest instability as compared to the two others. The measure of instability used is Max minus Min.

Accuracy analysis gives an accuracy of 95.4% which is very high. This result is determined by Leave-one-out and cross validation procedures. Detailed analysis presented in the next figure, shows that seventy per cent of the inaccuracy in the model occurred when predicting risk of high severity.

Figure 10- Summary of model accuracy for severity

4.4 Dynamic Bayesian network model for risk

In this Section, we combine the previous two models in one model. we summarize inference results derived from the model and perform sensitivity analysis for the model parameters to identify the effects of these inaccuracies on its outputs. The discussed sensitivity analysis refers to the effect of each conditional probability parameter on the marginal posterior distribution of the target node.

4.4.1 The mathematical model

We formulate the risk model via the following joint probability:

$$P(A, Z, M) = p(Z) \times P((A|Z)) \times \prod_{i=1}^{216} P(M_{j=1}^i | A) \times \prod_{j=2}^{14} \prod_{i=1}^{216} (P(M_j^i | Pa(M_j^i))) \quad (20)$$

The risk increases from 2% to 88% up to $j = 4$. At $j = 7$ the risk is again high but decreases up to $j = 12$ before the final step, $j = 13$.

Figure 11- BDN for overall risk

The overall picture of the products' risk is reported in Figure 12 is consistent. While overall risk is reduced by the end of this process, it seems that the products' levels of risk at the end are dissimilar, with Project B having a higher level of risk at $j = 13$ as compared to the other two projects.

Figure 12- Example of Overall risk over time of 3 out of 6 different products

4.4.2 Sensitivity analysis of the mathematical model:

Let us consider $j = 13$ as a target node on the mathematical model described in previous sub-section. We apply, the algorithm suggested by by Kjaerulff and van der Gaag³¹ to establish a relation between posterior marginal probability of $j = 13$ and the other parameters in the model. In addition, we identify the parameters which most affect the marginal posterior of $j = 13$. One could choose other node to be a target in the analysis, but in the discussed case we decided testing the final achieved risk ($j = 13$). The overall risk at $j = 13$ has two states: state 0 (risks with low probability or low impact or both) and state 1 (overall risk is high).

Results of sensitivity analysis, where $M[13] =$ state 0 (overall risk is low)

The posterior marginal value of the overall risk at $j = 13$ is: about 92%. The conditional probability parameter which most affect the posterior value of this state is the following: $P(M[13] = 0 | A = state1: M[12] = 0)$, with calculated value of 82.8%. Then, we tested different scenarios of uncertainty of this parameter; Table 6 summarizes the results.

Table 6: Range value of the target function of uncertainty of the most derivative parameter

Such analysis can be repeated for other parameters which affect the posterior distribution of $j = 13$. In the following figure, we present the top 10 parameters affecting the marginal posterior distribution.

Figure 13- Result of sensitivity analysis of the posterior distribution of $M_{13} = 0$

Figure 13 shows the most sensitive parameters for a selected state of the node $M[13]$ sorted from the most to least sensitive. The horizontal axis shows the absolute change in the posterior probability of $M[13] = 0$ when each of the parameters changes by 10% percentage.

Results of sensitivity analysis, where $M[13] = \text{state1}$ (high risks)

The results show that the parameter which most affect the posterior distribution of $M[13] = \text{state1}$ is similar to the parameter which affect the posterior with state0. Its value is 82.8%. In Table 7 we summarize the posterior distribution changes of this state as function of percentage of uncertainty in the parameter.

Table 7: Range value of the target function of uncertainty of the most derivative parameter

5 Discussion and conclusions

In this paper, we establish mathematical risk modeling and systematic risk management in engineering products. The results of the proposed method can be used as feedback in the development process in addition to predicting potential levels of product risk. The risk model combines probability of occurrence of risk, severity, and loss. Providing a descriptive relationship, these models can provide input to decision-making during development.

Using risk management as more than a prioritization tool enables measuring and monitoring risk over the development phase of the product. The global picture yielded reflects the maturity of the product, which can then be communicated to customers. Usually systematic risk analysis is conducted in hazardous industries, but it can be useful in other industries as well.

We use a DBN instead of the classical time series method because the flexibility of the DBN enables estimation of model parameters without assuming the linearity of the coefficients or incorporating prior knowledge. In addition, via the DBN we perform predictive inference and sensitivity analysis to address the uncertainty referring to the model as well as to the parameters.

The general model includes all three parts: identification of risks, analysis of risks, and control and feedback. This provides a complete understanding of the overall risk and its components. The approach of managing the overall risk of a product, instead of limiting the analysis to specific instances of high risk, highlights "small" risks which may play more important roles than initially assumed. Usually, industry treats uncertainty assessment without assessing the knowledge or confidence in the assessment probabilities. Our model provides this.

The paper does not discuss how to determine the acceptable level of risk. This is product and time specific and varies across companies. It is nevertheless essential to compare the achieved risk to agreed-upon risk standards. The acceptable level of risk should be defined at the beginning of product development.

Finally, risk behavior can examine using additional risk predictors in order to understand the relationship between risk predictors and technical risk assessment. A comprehensive study of this will expand the work presented here.

The paper presents and applies new mathematical models. These models are a basis for a systematic approach when managing risks in product development. Product development processes are characterized by uncertainty and these models are suitable to deal with this challenge.

References

1. Lindemann U. Methodische entwicklung technischer produkte. Springer: Berlin,2005.
2. Hall E. Managing risk- methods for Software systems development. Addison Wesley Longman: Reading, MA,1998
3. Smith P, Merritt G. Proactive risk management - controlling uncertainty in product development. Productivity Press: New York,2002.
4. Browning TR, Deyst JJ, Eppinger SD. Adding value in product development by creating information and reducing risk. *IEEE Trans. On Engineering Management* 2002;49:443-458.
5. Birolini A. Reliability engineering, theory and practice, Fourth Edition. Springer: Berlin,2004.
6. Stamatelatos M(Ed.). Probabilistic risk assessment procedures guide for NASA managers and practitioners, Version 1.1 NASA Headquarters, Office of Safety and Mission Assurance. Washington,2002.
7. Lessing G, Sierpinski C, Beckman D. Integrales risk management, total risk management Presentation, Lessing IRM GmbH. Altenburg, Germany,2005.
8. Oehmen J, Dick B, Lindemann U, Seering W. Risk management in product development - current methods. 9th International Design Conference. Dubrovnik, Croatia,2006.
9. Léger A, Weber p, Levrat E, Duval C, Farret R.,Iung B. Methodological developments for probabilistic risk analyses of socio-technical systems. Proceedings of the Institution of Mechanical Engineers, *Journal of Risk and Reliability* 2009; 223 (4):313-332.
10. Hamilton J. Time series analysis. Princeton University Press: Princeton, N.J.,1994.
11. Neapolitan ER. Learning Bayesian networks. Prentice-Hall: Upper Saddle River, N.J, 2003.
12. Murphy KP. Dynamic Bayesian networks: representation, inference and learning. *Doctoral dissertation*, University of California, Berkeley, 2002.
13. Jensen FV. An introduction to Bayesian networks. Taylor and Francis: London, 1996.
14. Jensen FV. Bayesian networks and decision graphs. Springer: New York, 2001.
15. Aven T. Risk assessment and risk management: Review of recent advances on their foundation. *European journal of Operational Research* 2016; 253:1-13.

16. Nilsen T, Aven T. Models and model uncertainty in the context of risk analysis. *Reliability Engineering & Systems Safety* 2003; 79:309-317. DOI: 10.1016/S0951-8320(02)00239-9
17. Aven T. On when to base event trees and fault trees on probability models and frequentist probabilities in quantitative risk assessments. *International Journal of Performability Engineering* 2012; 8:311-320.
18. Aven, T. On the need for restricting the probabilistic analysis in risk assessments to variability. *Risk Analysis* 2010; 30: 354-360 with discussion 381-384.
19. James, G., Witten, D., Hastie, T. and Tibshiriani, R. An introduction to statistical learning, Springer New York Heidelberg Dordrecht London, 2014. DOI 10.1007/978-1-4614-7138-7
20. SAS Institute Inc. JMP 12 Basic Analysis. Cary, NC: SAS Institute Inc, 2015.
21. Kenett RS. On Generating High InfoQ with Bayesian Networks. *Quality Technology and Quantitative Management* 2016; 13(3):309-332.
22. Kenett RS. Applications of Bayesian networks, SSRN Tech. Rep. <http://ssrn.com/abstract=2172713> [2012]
23. Koiter JR. Visualizing inference in Bayesian networks. PhD thesis, Delft University of Technology, Delft, The Netherlands, 2006.
24. Pearl J. Causality: Models, reasoning, and inference, Cambridge University Press: UK, 2000.
25. Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society* 1988; Series B 50(2): 157-224.
26. Ben-Gal I. Bayesian networks. Encyclopedia of statistics in quality and reliability. Ruggeri, Kenett and Faltin Editors, Wiley, 2007.
27. Lee JK, Carlin BJ. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *American Journal of Epidemiology* 2010; 171(5): 624-632.
28. Schafer JL. Analysis of incomplete multivariate data. Chapman and Hall: London, United Kingdom, 1977.
29. Cugnata F, Kenett RS, Salini S. Bayesian networks in survey data: Robustness and sensitivity issues. *Journal of Quality Technology* 2016; 48(3): 253-264.
30. Pearl J. Trygve Haavelmo and the Emergence of Causal Calculus. *Economic Theory* 2016; 31:152-179.
31. Uffe K, Van der Gaag CL. Making sensitivity analysis computationally efficient. *Uncertainty in Artificial Intelligence proceedings*, 2000.
32. Castillo EF, Gutierrez JM, Hadi AS. Sensitivity analysis in discrete Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics-Part S: Systems and Humans* 1997; 27(4):412-423.
33. Dean T, Kanazawa K. A model for reasoning about persistence and

- causation. *Artificial Intelligence* 1989; 93(1-2):1-27.
34. Druzdzel JM, Yuan CH. Importance sampling algorithms for Bayesian networks: Principles and Performance, *Mathematical and Computer Modelling* 2006; 43:1189-1207.
35. Menardi G, Torelli N. Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery* 2014;28(1):92-122. DOI:10.1007/s10618-012-0295-5.

Appendix

Explore Missing Values

Commands

Imputation Report

Undo

2680 missing values were replaced by least squares imputation. A shrinkage estimate was used, with off-diagonals scaled by a factor of 0. 211 rows and 15 columns were affected. There were 34 missing value patterns across columns. Imputed values colored light blue.

Missing Columns

Show only columns with missing
Close

Select columns and choose an action.

Select Rows Color Cells
Exclude Rows Color Rows

Column	Number Missing
Development duration [days]	0
4	4
5	4
6	4
7	23
8	23
9	35
10	23
11	32
12	32
13	22
14	45
15	13
16	24
18	89

Table 1: Descriptive statistics of severity of risk

Variable	Average	Variance	Standard deviation	Number of observation
Severity at ($j = 0$)	3.20	0.004	0.06	207
Severity at ($j = 1$)	3.20	0.004	0.06	207
Severity at ($j = 2$)	3.49	0.32	0.56	207
Severity at ($j = 3$)	4.05	0.13	0.36	163
Severity at ($j = 4$)	3.48	0.45	0.67	188
Severity at ($j = 5$)	3.14	0.06	0.25	199
Severity at ($j = 6$)	4.05	0.18	0.43	163
Severity at ($j = 7$)	3.48	0.11	0.33	179
Severity at ($j = 8$)	3.96	0.15	0.38	156
Severity at ($j = 9$)	2.88	0.14	0.38	158
Severity at ($j = 10$)	3.73	0.04	0.19	166
Severity at ($j = 11$)	2.83	0.01	0.12	197
Severity at	2.54	0.05	0.22	153

Variable	Average	Variance	Standard deviation	Number of observation
($j = 12$)				
Severity at ($j = 13$)	2.81	0.16	0.40	123

Table 1: Descriptive statistics of severity of risk

Table 2: Descriptive statistics of probability of risk

Variable	Average	Variance	Standard deviation	Number of observation
Probability at ($j = 1$)	2.60	0.02	0.12	207
Probability at ($j = 2$)	2.60	0.02	0.12	207
Probability at ($j = 3$)	2.96	0.29	0.54	207
Probability at ($j = 4$)	3.79	0.15	0.39	187
Probability at ($j = 5$)	3.03	0.47	0.69	187
Probability at ($j = 6$)	2.45	0.09	0.29	174
Probability at ($j = 7$)	2.39	0.22	0.47	187
Probability at ($j = 8$)	2.76	0.22	0.47	177
Probability at ($j = 9$)	2.51	0.24	0.49	187
Probability at	2.21	0.14	0.37	187

Variable	Average	Variance	Standard deviation	Number of observation
(j =10)				
Probability at (j =11)	1.82	0.03	0.18	166
Probability at (j =12)	2.50	0.01	0.12	197
Probability at (j =13)	2.84	0.03	0.16	184
Probability at (j =14)	2.76	0.19	0.43	106

Table 2: Descriptive statistics of probability of risk

	Project A	Project B	Project C	Project D	Project E	Project F
	State 0	State 1	State 2	State 3	State 4	State 5
Low probability State 0	0.9411692	0.99999664	0.99997723	0.99999219	0.9999746	0.99998148
High probability State 1	0.058830796	3.3634616e-06	2.2771347e-05	7.8125e-06	2.5402632e-05	1.8518519e-05

Table 3: Frequency of probability of risk, P_0 conditioned on project, A

	Marginal probability distribution
Low probability State 0	0.9899
High probability State 1	0.0100

Table 4: Marginal distribution for the node at $j=0$

	State 0- project A		State 1- project B		State 2- project C		State 3- project D		State 4- project E		State 5- project F	
	$j=1$											
$j=2$	low probability	high probability	low probability	high probability	low probability	high probability	low probability	high probability	low probability	high probability	low probability	high probability
low probability	0.0000	0.4985	0.1433	0.0022	0.0000	0.0076	0.0000	0.5000	0.0002	0.0002	0.0000	0.0015
high probability	1.0000	0.5015	0.8567	0.9978	1.0000	0.9924	1.0000	0.5000	0.9998	0.9998	1.0000	0.9985

Table 5: Transition probabilities between probability of risk at time $j = 2$ given the probability of risk scenario at $j = 1$

Percentage of uncertainty on the conditional parameter	Range value of the parameter	Range value of the target ($M_{13} = 0$)
10%	74.5%-91.08%	89.9%-94.13%
20%	66.2%-99.3%	87.8%-96.2%
30%	57.9%-100%	85.8%-96.3%
40%	49.6%-100%	83.7%-96.3%

Table 6: Range value of the target function of uncertainty of the most derivative parameter

Percentage of uncertainty	Range value of the parameter	Range value of the target, ($M_{13} = 1$)
10%	74.5%-91.08%	5.8%-10%
20%	66.2%-99.3%	3.7%-12%
30%	57.9%-100%	3.6%-14%
40%	49.6%-100%	3.6%-16%

Table 7: Range value of the target function of uncertainty of the most derivative parameter