A Bayesian Belief Networks Approach to Risk Control in Construction Projects

Ibsen Chivatá Cárdenas, Saad SH Al-jibouri & Johannes IM Halman
University of Twente, The Netherlands

Abstract

Although risk control is a key step in risk management of construction projects, very often risk measures used are based merely on personal experience and engineering judgement rather than analysis of comprehensive information relating to a specific risk. This paper deals with an approach to provide better information to derive relevant and effective risk measures for specific risks. The approach relies on developing risk models to represent interactions between risk factors and carrying out analysis to identify critical factors on which risk measures must focus. To ameliorate the problem related to the scarcity of risks information often encountered in construction projects, Bayesian Belief Networks are used and expert knowledge is elicited to augment available information. The paper describes proposed modifications to the standard methods used to develop Bayesian Belief Networks in order to deal with divergent information originated from epistemic uncertainty of risks. The capacity of the proposed approach to provide better information to support risk related decision making is verified by means of an illustrative application to risk factors involved in the construction of cross passages between tunnels tubes in soft soils.

Keywords: Risk control, risk modelling, reliability modelling, risk-related knowledge modelling, relevant information, epistemic uncertainty, Bayesian Belief Networks.

1 Introduction

Construction projects are risky undertakings. This is particularly true in underground construction projects. According to the International Association of Engineering Insurers, IAEI (Landrin et al., 2006), the estimated economic losses as a consequence of failures in some 18 underground construction projects worldwide from 1994 to 2005 were more than 570 million Euro. Additionally the IAEI's report indicates that the average delay in the completion of the projects was 19 months. Also the damage to property due to those failures was estimated to be a considerable 600 million Euro. The specific failures were events in which significant parts or all of the works either collapsed or saw excessive deformation. Abundant evidence has highlighted that a large proportion of failures in ground-related construction projects are a result of shortcomings in using the available knowledge rather than unknown factors such as unexpected ground conditions (e.g. Muir Wood 2000, Bea 2006, Van Tol 2007, Wearne 2008, Mitchell 2009). The shortcomings were seen mainly in situations in which tacit or explicit information, such as risks and uncertainties, technological knowledge, design assumptions, monitoring records, thresholds and tolerances, was either ignored, improperly used, rejected or not passed on by someone in the project.

The above situation can be made worse due to inadequate standard approaches and practices to assess and analyse construction risks comprehensively. There are a number of reasons for that. Tunnel works are very complex projects that involve many tasks and people. The risks are also numerous and, consequently, much information has to be gathered to support risk analysis and management.

However historical data on risks and failures are usually scarce, confidential, and often not available until several years after a failure event. These factors make assessment and use of knowledge on risks in such projects major and difficult task.

The problem is usually exasperated due to the uncertain nature of construction. For instance many negative impacts on a project due to uncertainties related to ground conditions often materialise in the construction stage. Most of these unexpected factors and their influences on the project could not have been foreseen earlier because they are either partially known or unknown to engineering practice. Project-specific factors can also make the process even harder due to, mainly, the wide ground conditions and project features variability. Even when averaged statistics existed of incidents with particular combinations of the several factors, their application to a specific project would have little significance (Muir Wood, 2000). In addition, a construction project, its environment and therefore its risks are continuously evolving. The project risks should thus be continuously assessed and modelled throughout the project. New risks have to be continually identified and analysed. This is barely done in current practice in a comprehensive way.

Given the challenges to underground construction projects described above, most of the conventional methods will fall short of providing reliable support to risk decision-making. For example, many of the traditional methods evaluate risks independently and incomprehensively. Even the few more comprehensive methods, such as, failure mode and effects analysis, hazard analysis, top-level event tree, fault tree analyses, among others (Sousa, 2010), would struggle to deal with a further problem, the epistemic uncertainty of risks caused by incomplete of information, which is an usual condition in the case of construction risks.

The approach proposed in this paper provides a possibility to overcome these difficulties to a certain degree. It offers the chance of incorporating expert judgement about the risk factors associated with critical risks to bridge the gaps in the available historical data. Furthermore, the approach enables information on new risk factors and causal relations to be analysed based on Bayesian Belief Networks (BBNs) approach. BBNs method is a suitable way of representing complex and uncertain relationships among many factors that contribute to the occurrence of risks (Sigurdsson et al., 2001; Daniels et al., 2008; Weber et al., 2010). There are other methods for representing risks such as Markov Chains, Petri Nets, Artificial Neural Networks (ANNs), the Analytical Hierarchy Process (AHP), Systems Dynamics, and Fuzzy Systems (Sousa 2010; Taroun 2011). However, some of these are regarded as being too complicated to be used by practitioners, notably Markov Chains and Petri Nets (Simon et al., 2007). Others, such as ANNs, the AHP, and Systems Dynamics require abundant data which, as mentioned, are usually unavailable for construction risks. Also whilst Fuzzy Systems modelling is a realistic alternative to the BBN approach, the latter makes use of mathematically sounder and simpler rules for drawing inferences in order to execute efficient analysis and provide information on the most critical factors risk measures should focus on.

Bles et al. (2003) were the first to demonstrate the application of Bayesian Belief Networks (BBNs) for representing underground construction risks while Sousa (2010) had demonstrated the application of Bayesian Networks (BNs) for tunnelling. With a different approach, using only hard data, from a tunnelling project in Porto, this author developed a geologic prediction model. Dynamic BNs were also used by Špačková and Straub (2011) to model the excavation performance of a road tunnel built using the New Austrian Tunnelling Method. Despite the enormous contribution of these studies, none of them addressed the problem of dealing with differences in certainty between experts judgements caused by incomplete information of the risk data. Expert judgement naturally diverges (Pate-Cornell, 1996) and thus, there is no certainty about the estimated values assigned to a variable. The epistemic uncertainty must be informed to avoid decision making to be misled (Aven, 2009).

BNNs is a powerful tool to model incomplete information of high dimensional phenomena, but some modifications in their construction are required to deal with epistemic uncertainties of risks. Most of the developed BNNs reported can be regarded as causal models having a fixed structure with constant conditional probabilities representing single interactions amongst variables. This 'deterministic' configuration of standard BNNs impedes to model epistemic uncertainty directly reflected in differences in information provided by experts. Using the unique capabilities of BNNs to model risks, this paper shows the modifications to the standard ways of constructing BNNs and a particular example using a Bayesian Networks risk model is presented showing how they provide information to derive risk measures for the construction of cross passages between tunnels tubes in soft soils. The construction of cross passages is regarded as one of the most risky activities of tunnel works in the Netherlands. A single improper operation in construction process can impact the whole construction integrity of the tunnel.

2 Risk modelling

The BBN approach is essentially a framework for modelling the relationships between variables, and for capturing the uncertainty in the dependencies between these variables using conditional probabilities (van der Gaag, 1996). The probability of a value of a factor in the BBN occurring is determined by the occurrence of change in other interrelated factors (Onisko et al., 2001). The inference mechanism used in a BBN is the Bayes theorem which makes it possible to compute the probability of an effect on any variable in the model from the probability of a given cause. With two directly related variables, the probabilities can be computed as follows (Vick, 2002):

$$P[\text{effect}]=[P[\text{effect}/\text{cause}].P[\text{cause}]]/ P[\text{cause}/\text{effect}] \quad (1)$$

where:

P[cause] = probability that the cause occurs,
P[effect] = probability that the effect occurs,
P[effect/cause] = conditional probability of the effect, given the cause,
P[cause/effect] = conditional probability of the cause, given the effect.

The posterior probability of the cause from the effect can similarly be derived as:

$$P[\text{cause}/\text{effect}] = [P[\text{effect}/\text{cause}].P[\text{cause}]]/P[\text{effect}] \quad (2)$$

BBNs can be used to construct risk models composed of scenarios based on a set of known possible risk factors associated with the risks being analysed. These possible scenarios must be structured as a set of mutually exclusive and collectively exhaustive elements to which a probability distribution can be attributed. The probabilities of the risk factors are usually encoded based on expert judgement. Judgements are elicited from experts by means of a structured method which is specified to minimise bias in the estimates provided.

In a BBN, the interrelationships between variables are expressed graphically in the form of diagrams. Variables are represented by nodes. Diagram nodes that have interdependencies are connected by arcs, whereas independent nodes are not connected. The direction attached to an arc reflects the direction of causal influence, which might be indicated by an expert, or scientifically proven. Figure 1 shows a fragment of a BBN model produced in this research to represent interactions between some factors leading to the event 'excessive deformation in the soil' associated with the 'collapse of cross-passage excavation' which is depicted in Figure 2. Information on conditional probabilities attached to the causal influences of the risk factors is not indicated on the diagram but is stored in the model and accessible to the user. An example illustrating how conditional probability information is stored in a BBN is provided below.

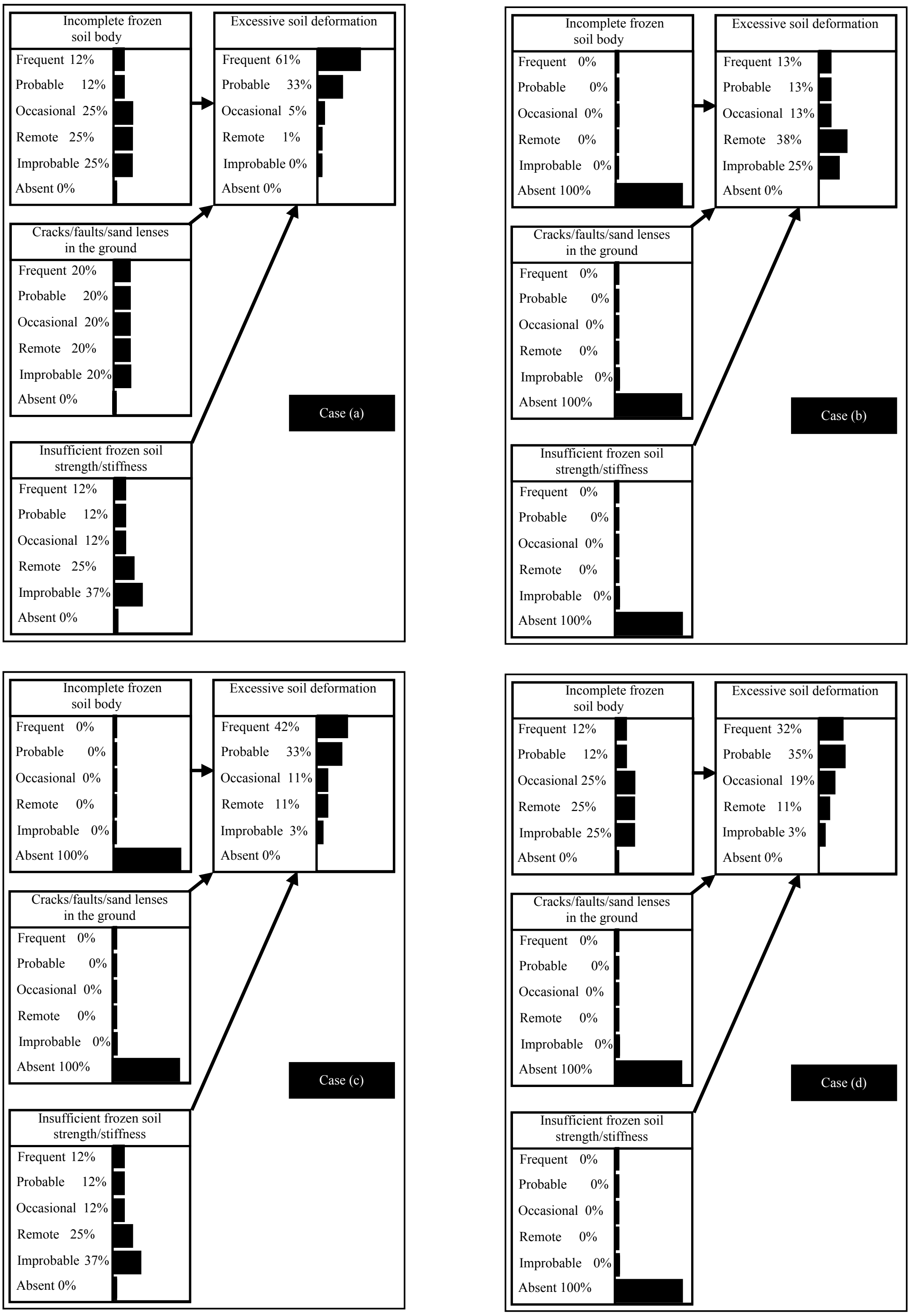

Figure 1. Example distributions representing evidence supporting the variables in the model

To address the uncertainties of risk information, modifications to the standard methods to construct Bayesian Networks were implemented. The modifications are related to the way uncertainties in the variables and in the relationships are represented, and involve the use of specific methods to process that information as described later in this paper.

The first modification involves representing the variables in the models by means of 'evidence probability distributions' rather than with point conditional probabilities. These evidence probability distributions reflect the strengths of the evidence obtained from the experts judgements.

Figure 1 shows a set of variables with their evidence probability distributions derived on the basis of experts judgements. In this setting, each expert's assessment is treated as a piece of evidence and each evidence probability distribution is obtained from the relative counts of expert estimates indicating a particular chance interval. Figure 1, case (a), shows an evidence distribution for each variable and a joint distribution for the 'excessive deformation in the soil' factor. In Figure 1, case (b), the likelihood of the variable 'excessive deformation in the soil' can be associated mainly with the category 'Remote' since this category is supported by the best part of evidence incorporated into the model.

Each variable in the model is regarded as an event or condition representing a fault event, state of failure, or unfavourable condition. Fault events or states of failure associated with a variable can be events in which a risk factor exceeds a predefined threshold. Accordingly, most of the variables have two possible states: 'absent' or 'present'. A variable is regarded as having the status 'absent' when it is not active under the particular conditions being analysed. As an example, under the particular conditions analysed in Figure 1, case (d), only one of the three conditioning variables (incomplete frozen body) is affecting the joint evidence distribution of the conditioned variable 'excessive deformation in the soil'.

Although variables can have more than two possible states, few variables in our models required expressing in such a way. The 'present' status was further discretized into five chance categories. In line with this, experts were asked to provide estimates of chance regarding the occurrence of risk factors in terms of qualitative probabilities using the scale of five categories shown in Figure 1.

The second modification to the standard BBN method used in this work relates to the relationships in the models. To capture the relationships among the factors, conditional probabilities were elicited from experts or established from proven scientific relationships. Such information is conventionally stored in Conditional Probability Tables (CPTs). A CPT can model any kind of probabilistic interactions between variables. To enable epistemic uncertainties to be captured and then processed, a modification to the standard approach to developing CPTs was applied. Extending the work of Galán and Díez (2000), a CPT for a pair of binary variables, A and B, can be specified as shown in Table 1:

Table 1. Modified Conditional Probability Table

<table>
<tr><td rowspan="3"></td><td colspan="2" rowspan="3">B/A</td><td colspan="3">A</td><td></td></tr>
<tr><td colspan="3">Present</td><td rowspan="2">Absent</td></tr>
<tr><td>a_1</td><td>$...a_j...$</td><td>a_n</td></tr>
<tr><td rowspan="5">B</td><td rowspan="5">Present</td><td>b_1</td><td>$m_{(b1/a1)}$</td><td>...</td><td>...</td><td>0</td></tr>
<tr><td>...</td><td>...</td><td>...</td><td>...</td><td>0</td></tr>
<tr><td>b_j</td><td>...</td><td>$m_{(bj/aj)}$</td><td>...</td><td>0</td></tr>
<tr><td>...</td><td>...</td><td>...</td><td>...</td><td>0</td></tr>
<tr><td>b_m</td><td>...</td><td>...</td><td>$m_{(bm/an)}$</td><td>0</td></tr>
<tr><td></td><td colspan="2">Absent</td><td>$1-\sum_{b=1}^{m} m_{(bj/a1)}$</td><td>...</td><td>$1-\sum_{b=1}^{m} m_{(bj/an)}$</td><td>1</td></tr>
</table>

In Table 1, a_j and b_j are respectively the chance categories for events A and B. Categories of chance can be discretized, for example into a vector {Frequent, Probable, Occasional, Remote, Improbable}. In our situation, a_j is the chance of A according to a given expert. Then, b_j is the chance of B occurring when A has happened, and $m(b_j/a_j)$ is the available evidence supporting (b_j /a_j). The latter is obtained from the relative counts of expert estimates indicating the given combination of b_j and a_j. Using the term $1-\sum_{b=1}^{m} m_{(bj/an)}$ ensures that the sum of the relative counts equals unity, in agreement with the law of total probability and accounts for the evidence against bj /aj occurring. Table 2 provides a numerical example of a modified conditional probability table.

The CPT shown in Table 2 reflects that variable A has a chance ranging from 'Remote' to 'Improbable'. Similarly, the available evidence indicates that, even if A occurs, B has little chance of occurring and only covers the 'Remote' to 'Improbable' range of chance categories – that is, there is stronger evidence that B will not occur.

Table 2. Example of a modified Conditional Probability Table

			A					
		B/A	Present					Absent
			Frequent	Probable	Occasional	Remote	Improbable	
B	Present	Frequent	0	0	0	0.01	0.05	0
		Probable	0	0	0	0.01	0.05	0
		Occasional	0	0	0	0.03	0.05	0
		Remote	0	0	0	0.04	0.10	0
		Improbable	0	0	0	0.07	0.34	0
		Absent	1	1	1	0.84	0.41	1

When the number of variables and the number of states describing those variables, becomes large, specifying the probabilistic relationships is very time-consuming and the outcome often lacks reliability. To reduce the model building effort significantly, the Conditional Probability Tables can be approximated to Noisy-MAX and Noisy-AND gates. Noisy-MAX and Noisy-AND distributions require fewer conditional probability estimates in approximating the CPTs. A comprehensive framework on the use of this approximation and its constraints were provided by Díez and Druzdzel (2007).

For this research, both the Netica (Norsys Corporation, Canada) and the Genie (Decisions Systems Laboratory, University of Pittsburgh) software packages were used to construct the networks and to perform the analyses. The models developed were compiled in both software packages in order to verify the correctness of the computations when propagating the data incorporated into the models.

3 Analysis

The analysis of the models is intended to obtain different indicators informing on the relative influence of the risk factors on the major risk under analysis. By determining those factors that increase most the chance of occurrence of a selected target risk in the model a project manager increases his understanding of risks and this information can support the planning of control actions.

To determine the relative influence on the occurrence of the major risk under analysis, the Bayesian theorem explained above is used, which is powered and automated by Bayesian Networks. The necessary data was directly elicited from experts by means of a structured judgement elicitation method which is reported in another paper.

For the proposed analysis additional experimental data was used. This experimental data consist of information of the marginal chance of occurrence of the risk factors. Information on the chance of the relevant factors depends on the particular setting of each project. However, for the purpose of this analysis, experts were asked to give their best estimation of the risk factors occurring. For the case of the variables related to ground conditions, which depend entirely on the particular conditions of each project; they were modelled using uniform distributions. Once information on the marginal probability risk factors is incorporated into the model the most critical risk factors can be determined. The most influential factors are determined by a set of experiments consisting of either altering the probability distributions of each of the risk factors or removing them one-by-one and observing the related changes in the posterior probabilities of the joint evidence probability distribution of the target factor being analysed. In this way one variable is changed or removed from the original set of input variables and the changes in the output compared with those of other input variables (Deng et al., 2010). If a change of a variable modify the output distribution significantly, then the input will be registered as a very influencing one. Respectively, if a factor is non-influencing to the model output, then little or no change will be observed in the output distribution and the input variable will be associated with a low importance. An example in the next section shows how the above analysis can be used to identify relevant risk factors from the risk model developed.

4 Results

In Figure 2, the components of the developed risk model for the risk factors involved in the construction of cross-passages in soft soils are displayed (variable states are not shown for reasons of clarity). According to the experts consulted, more than fifty risk factors were identified as relevant to the occurrence of excessive deformation of soil and water inflow, potentially leading to the collapse of the excavation of cross-passages. The model included issues limited to soft soils similar to Dutch ground conditions. Dutch ground conditions are characterised by saturated, low stiffness sandy soils with medium-fine size particles and a high groundwater table. The developed risk model refers only to bored cross-passages using ground freezing technologies in combination with outer struts protecting the main tunnel tubes as the temporal support and concrete linings cast in situ as the definitive support.

To illustrate how the model can provide relevant information to support risk control, a set of risk factors are analysed. The procedure described in section 3 is employed to analyse the fragment of the model displayed above in Figure 1. According to the estimates gathered from the experts, the event ‘excessive deformation in the soil’ has a marginal probability of occurrence that falls mainly into the categories ‘remote’ and ‘improbable’ (Case b). This case corresponds to the condition in which none of the factors directly related to the event ‘excessive deformation in the soil’ are present. In the case (a), the distribution attached to the event ‘excessive deformation in the soil’ is the joint distribution of evidence for the condition in which three influencing factors are present. In this case the factor ‘excessive deformation in the soil’ exhibits a high a chance of occurrence falling into the categories of ‘Frequent’ and ‘Probable’.

Cases (c) and (d) in Figure 1 are used to make a comparison and determine which risk factors directly influencing the occurrence of ‘excessive deformation in the soil’ are the critical ones. For example according to the data incorporated into the model, the factor ‘Insufficient frozen soil strength/stiffness during freezing up/maintenance’ appears to be more critical than the ‘incomplete frozen soil body’ event. Despite there exists a slight difference between the output distributions, the chance of that the ‘insufficient frozen soil strength/stiffness during freezing up/maintenance’ event results in ‘‘excessive deformation in the soil’ is higher according to the evidence available. With this result a practitioner might decide to allocate more resources on guaranteeing the required strength or stiffness of the frozen soil.

Nevertheless, it is necessary to notice, that the similarity in the output distributions does not provide a clear-cut information to decision making. Sensitivity indicators measuring the uncertainty in the output distributions can facilitate making choices. Appropriate sensitivity indicators were investigated in this research. Their evaluation was a subject of another paper.

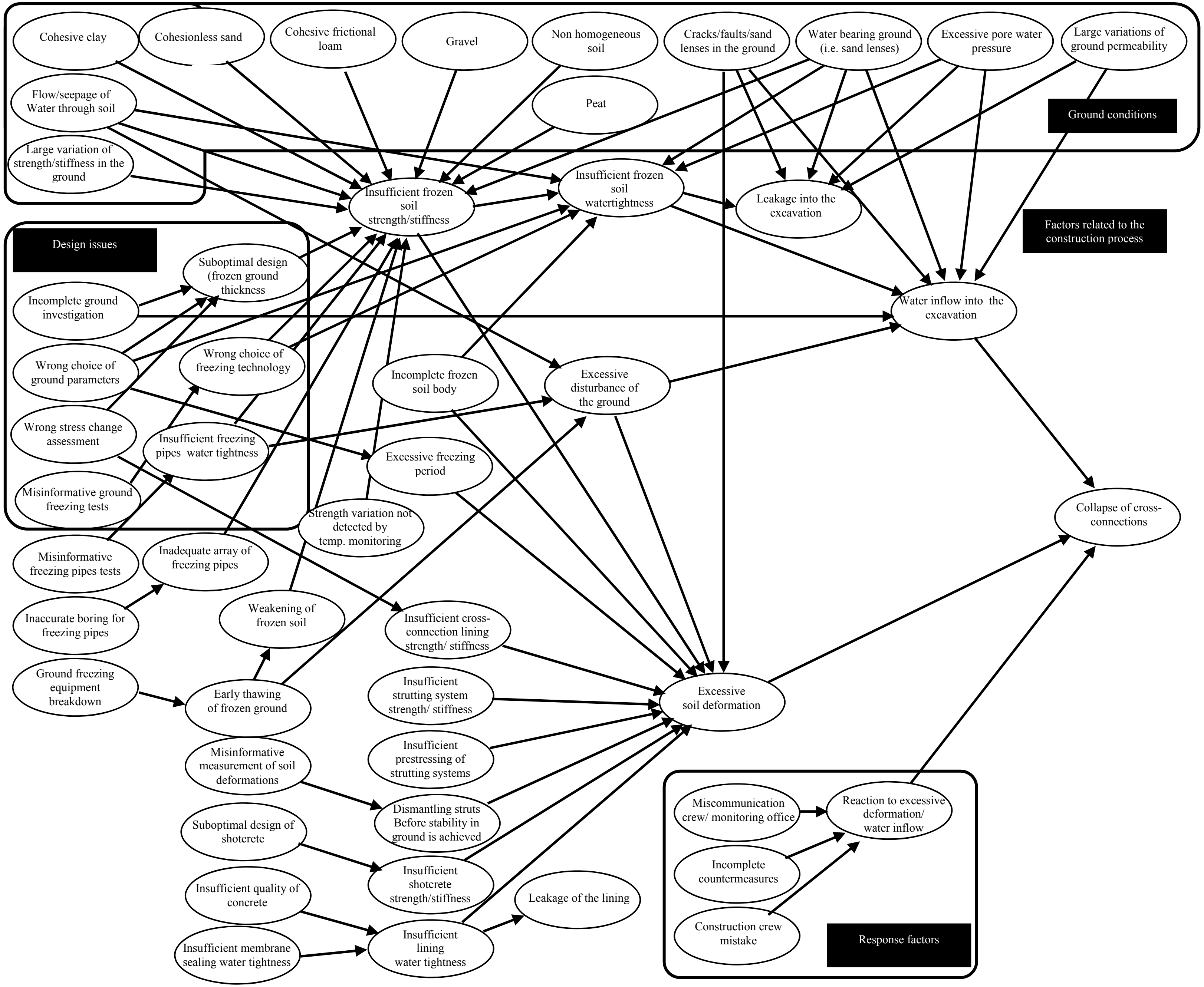

Figure 2. Risk model for construction of cross-passages in soft soils in bored tunnels

The analysis described in this section in combination with the use of sensitivity indicators was repeated for every variable in the model which delivered information on the most relevant risk factors to control for the case of the collapse of the construction of cross-passages in soft soils. The results sampled from the analysis and evaluation of the model reported in this paper confirms the ability of the model (and the proposed approach) to provide information to support risk management decisions.

It is worth mentioning here that the information provided by the model can also be combined with other criteria, such as the cost of the risk measures or the controllability of risk factors enabling better informed decision making.

5 Discussion

This paper has reported on an approach to represent and analyse construction risks. The approach consist of using available historical data in combination with expert judgement to characterise the risk factors and their interactions involved in major construction risks. Expert judgement is used exhaustively to bridge the gaps in the available risk-related information, that frequently is incomplete. The approach enables risk factors associated with major risks to be realistically modelled to account for the epistemic uncertainty within such factors. In the approach information and knowledge on risks is integrated and modelled using Bayesian Belief Networks BBNs. In order to determine specific risk measures and support decision making, a number of analysis are powered by BBN, in cost-effective and friendly terms, allowing information on critical risk factors to be rendered. The elements in the proposed approach contribute to the standard construction practice by evaluating a method to analyse risks in a comprehensive way, which is not a standard practice in construction.

The paper has shown how critical factors can be derived from the developed model using experimental data. This however does not indicate that the approach is not useful to support real projects. Project-specific information, such as previously unidentified risk factors (and their probabilistic information) emerging during the construction of the tunnel works, can be incorporated into the models and analysed on a case-by case basis as demonstrated in another paper.

In this research six tunneling risk-models have been developed using the approach proposed in this paper, thus, showing the feasibility of the approach. To make this approach useful in real projects, knowledge focused on different critical risks has to be captured and integrated to encompass a great proportion of the major risks usually identified in underground construction projects.

Future studies are required to determine whether there is a need for improvements and refinements, as well as, their limitations, concerning the applicability of the approach and the use of knowledge rendered by the model in real projects.

6 Conclusion

This paper has described an approach to represent and analyse risks in construction projects. The approach shows how tunnel risks and their epistemic uncertainty can be modelled and analysed using Bayesian Belief Networks in order to provide more reliable information to aid decision-making. The paper concludes that, despite the complex and uncertain nature of construction risks, the developed approach can produce useful results which could guide the allocation of resources to specific risk remedial measures on a cost-efficient basis.

Acknowledgements

This research has been supported by DELTARES, a Dutch research institute, whose financial support is gratefully acknowledged. The authors would further like to thank the experts who agreed to be involved in this research and without their insights and support the work could not have been achieved.

References

AVEN, T., 2009. Perspectives on risk in a decision-making context – Review and discussion. *Safety Science*, 47(10), 1355-1364.

BEA, R., 2006. Reliability and Human Factors in Geotechnical Engineering. *J Geotech and Geoenvir Engrg*, 132(5), 631-642.

BLES, T., AL-JIBOURI, S., and VAN DE ADEL, J., 2003. A risk model for pile foundations. *In: 20th International Symposium on Automation and Robotics in Construction ISARC2003 The Future Site*, 2003, *Eindhoven*.

DANIELS, D.C., HUDSON, L.D., MAHONEY, S.M., WARE, B.S. and WRIGHT, E,J., 2008. Terrorism risk management. In: PUORRET, O., NAIM, P., and MARCOT, B., eds. *Bayesian Networks: A Practical Guide to Applications*. Chichester, West Sussex: John Wiley & Sons Ltd., pp. 239-261

DÍEZ, F.J. and DRUZDZEL, M.J., 2007. *Canonical Probabilistic Models for Knowledge Engineering. Version 0.9.* Madrid: Centro de Investigación sobre Sistemas Inteligentes de Ayuda a la Decisión; 2007. Technical Report CISIAD-06-01.

DENG, X., ZENG, X., VROMAN, P. and KOEHL, L. 2010. Selection of relevant variables for industrial process modeling by combining experimental data sensitivity and human knowledge. *Engineering Applications of Artificial Intelligence,* 23(8), 1368-1379.

GALÁN, S.F. and DÍEZ, F.J., 2000. Modelling Dynamic Causal Interactions with Bayesian Networks: Temporal Noisy Gates. ***In: CANEW' 2000 2nd International Workshop on Causal Networks held in conjunction with ECAI 2000 the 14th European Conference on Artificial Intelligence, 2000, Berlin.***

LANDRIN, H., BLÜCKERT, C., PERRIN, J.P., STACEY, S. and STOLFA, A., 2006. ALOP/DSU coverage for tunnelling risks? ***In: 39th Annual Conference of the International Association of Engineering Insurers, 2006, Boston.***

MUIR WOOD, A.M., 2000. *Tunnelling: Management by design*. London and New York: Taylor & Francis Group.

ONISKO, A., DRUZDZEL, M.J. and WASYLUK, H., 2001. Learning Bayesian Network Parameters from Small DataSets: Application of Noisy-OR Gates. *International Journal of Approximate Reasoning*, 27(2), 165-182.

PATE-CORNELL, M.E., 1996. Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering and System Safety,* 54(2-3), 95-111.

SIGURDSSON, J.H., WALLS, L.A. and QUIGLEY, J.L., 2001. Bayesian belief nets for managing expert judgement and modelling reliability. *Qual Reliab Engng Int*, 17, 181–190.

SIMON, Ch., WEBER, Ph. and LEVRAT, E., 2007. Bayesian Networks and Evidence Theory to Model Complex Systems Reliability. *Journal of Computers,* 2(1), 33-43.

SOUSA, R.L., 2010. *Risk analysis of tunneling projects, PhD Thesis*. Cambridge: Massachusetts Institute of Technology.

TAROUN, A., YANG, J.B., LOWE, D., 2011. Construction Risk Modelling and Assessment: Insights from a Literature Review. *The Built & Human Environmental Review*, 4(1), 87-97.

VAN DER GAAG, L.C., 1996. Bayesian belief networks: Odds and ends. *The Computer Journal*, 39(2), 97–113.

VAN TOL, A.F., 2007. Schadegevallen bij bouwputten. [Damage caused by Deep Excavations]. *Cement*, 59(6), 6-13.

VICK, S., 2002. *Degrees of Belief - Subjective Probability and Engineering Judgment*. Reston: ASCE Press.

WEBER, P., MEDINA-OLIVA, G., SIMON, C. and IUNG, B., 2010. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*. doi:10.1016/j.engappai.2010.06.002.