

TOWARDS THE EVALUATION OF HUMAN FACTORS IN AVALANCHE EARLY WARNING SYSTEMS

Martina Sättele^{1*}, Michael Bründl¹ and Kilian Zwirgmaier²

¹ WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland

² Technische Universität München, Munich, Germany

ABSTRACT: Early warning systems (EWS) play a crucial role in managing the risks posed by avalanches and other mass-movement processes to settlements and infrastructures. Although guidelines to evaluate the effectiveness of structural protection measures, such as avalanche defense structures, are commonly used in practice, this is not the case for EWS. Recently, a novel framework to evaluate their effectiveness as a function of their reliability was proposed (Sättele et al., 2016). In this framework, the reliability of automated EWS parts is modeled probabilistically in Bayesian Networks (BN) and expressed in terms of the probability of detection and the probability of false alarms. It could be shown that the optimal trade-off between those two strongly depends on the strategy applied to monitor the hazard and the thresholds applied to deliver early warning. Besides this, the results revealed that human factors, such as the risk attitude of decision-makers, are decisive in an EWS evaluation. In this paper, we demonstrate how the influence of human factors on the reliability of avalanche EWS can be assessed quantitatively. In line with established human reliability analysis (HRA) methods, which are frequently applied in industries, such as nuclear power (Kirwan and Ainsworth, 1992), we start with a detailed task analysis. To identify the cognitive tasks required of avalanche experts we conducted structured interviews. From these interviews a generic mental model is derived. The interviews also allowed us to identify the personal and external factors that influence if a task is conducted successfully. To demonstrate a qualitative assessment, nodes of the mental model and personal and external human factors are comprised in an extended BN. Finally, we discuss the potential of, and limitations regarding the integration of human factors into our framework for the evaluation of EWS.

KEYWORDS: early warning system, human factors, reliability, decision-making

1. INTRODUCTION

In the last decade, early warning systems (EWS) are increasingly applied as mitigation measures to prevent avalanche damage and casualties on roads, railways and in settlements (Gubler 2000; Bründl et al. 2004; Sättele 2015). EWS have become widely accepted as flexible and cost effective mitigation measures, although little is known about their effectiveness (degree of risk reduction they achieve). In the context of an integrated risk management approach the effectiveness and the costs of alternative mitigation measures are compared to identify optimal risk mitigation strategies (Bründl et al. 2009). When evaluating the effectiveness of structural mitigation measures, clear guidelines are in place; e.g. in Switzerland (Margreth and Romang 2010).

Facilitating the evaluation of complex EWS, Sättele et al. (2016) recently proposed a novel framework generically applicable for EWS applied to mass movement processes. A classification, distinguishing EWS according to their degree of automation forms the basis of this framework. The reliability of automated EWS is expressed in terms of the probability of detection (POD) and the probability of false alarms (PFA), which both influence the EWS effectiveness. EWS commonly reduce risk by decreasing the presence probability of persons and mobile objects in endangered areas. To be effective, they have to detect hazardous events (POD) and lead to protective actions with which affected persons comply. A large number of false alarms can reduce the probability of compliance (POC), which is known as cry wolf syndrome (Breznitz 1989).

In this framework approach the reliability is modeled in a Bayesian network (BN). Essentially, a BN represents a probabilistic modeling tool, which allows efficient representation of a joint probability distribution of several random variables. BNs consist of a qualitative part, represented through a

* *Corresponding author address:*

Martina Sättele, SLF,
Davos Dorf, Switzerland;
tel: +41-81-4170-261;
email: saettele@slf.ch

directed acyclic graph (DAG), and a set of local conditional probability distributions, which quantitatively represent the dependencies between random variables. Over the past decades BNs have been increasingly applied in the fields of engineering risk analysis and reliability (Langseth and Portinale 2007; Straub and Der Kiureghian 2010; Weber et al. 2012) and for natural risk (Straub 2005; Aguilera et al. 2011). For a more in depth introduction we refer to the relevant textbooks, e.g. Jensen and Nielson (2007).

The reliability assessment of non-automated EWS is treated in less detail in this recently published framework, although it is well known that human factors strongly influence the EWS effectiveness (Sorensen and Mileti 1987). This is because human tasks conducted in EWS are very complex and primarily cognitive e.g. to evaluate the danger and decide on protective actions (Stoffel and Schweizer 2008). Similar to naturalistic decision-making approaches described by Klein (1998), tasks in EWS are conducted under time pressure, deal with high stakes (death/life), with incomplete information, poorly defined procedures, dynamic conditions and involve a large degree of uncertainty (see also Stewart et al. 1997; Doswell III 2004; Downton et al. 2005; Morss et al. 2005; Guzzetti 2015; Morss et al. 2015).

The influence of human factors in the field of avalanche risk management has received little attention in the past and existing studies focused on recreational decision-making (Zweifel and Haegeli 2014). Decisions made by avalanche experts are highly dynamic and influenced by iterative/evolutionary tasks in which information is assembled cumulatively over time (McClung 2002a; McClung 2002b). To evaluate avalanche danger, experts judge the snow instability in space and time relative to a given trigger level. To do so, they analyze snow pack data, cues on current and future snow and weather conditions and triggers such as new snow, skiing, explosives, temperature change. The performance of the experts depends on their personal traits and perception, which again depend e.g. on experiences, risk propensity and biases. An extensive study on organizational biases was recently published by Johnson et al., (2016). Their interviews with 392 avalanche experts revealed that organizational factors, such as a strong safety culture, regular trainings and clear procedures, lead to good decisions; while operational pressure, weak management and decision-making structures have a negative influence. However, compared to other industries, decisions made by avalanche experts involve a large degree

of freedom. For example, thresholds defined in safety concepts for critical amounts of new snow are mainly indications that can be adapted for individual situations, rather than fixed target values (Schweizer and Föhn 1996). By using their expertise, experts can often compensate for incomplete and uncertain information and make better decisions. For example, new snow data measured by automated stations can vary strongly within miles of critical release areas. Furthermore, weather forecasts, besides being inherently uncertain, typically provide lower spatial resolution than required.

These characteristics of avalanche EWS hinder the application of existing quantitative human reliability analysis (HRA) methods developed for safety critical industries, such as nuclear power, aerospace and railway. Unlike in the field of avalanche risk, procedures and decision-rules are clearly defined in these industries, and thus make the concept of human error applicable. Despite the difficulties of evaluating human error in the assessment of avalanche EWS, the basic ideas and steps of HRA can be adapted for our needs.

Within the context of a probabilistic risk assessment for a man-machine system, HRA is concerned with the influence of human failure events (HFEs). After their identification HFEs are quantified using human error probabilities (HEPs) (Kirwan and Ainsworth, 1992). Typically HEPs are assigned conditionally on a specific context, which is described through performance shaping factors (PSFs). Both internal and external factors such as stress, education level, and management culture can be considered (Groth and Mosleh 2012; VDI 2015). Currently more than 50 HRA methods exist and new methods are continuously developed (Groth and Swiler 2013). For comprehensive summaries on HRA methods see (Kirwan 1994; Bell and Holroyd 2009; Spurgin 2010; Boring 2012; Di Pasquale et al. 2013). To overcome the weaknesses of traditional quantification frameworks and to increase the traceability of HRA methods, BNs have become increasingly popular in HRA (Zwirgmaier et al. submitted). A review of BN based HRA methods is provided by Mkrtchyan et al. (2015).

In the field of natural hazard risk management mental models to graphically capture relevant cognitive tasks have been proposed in the context of flash flood or hurricane warning processes (Bostrom et al. 2015; Morss et al. 2015). In line with Morgen (2002) these models are elicited on structured interviews. Such mental models can be represented in BNs. In the present contribution we

demonstrate how human factors can be considered in BNs when evaluating EWS.

2. RESEARCH APPROACH

Our approach is based on the outcomes of structured interviews with five avalanche experts. All experts work in the field of avalanche safety in Switzerland and have been assessing avalanche danger for between 13 and 25 years. Each interview was conducted following the same procedure and took approximately 2 h. Initially interviewees were explained the background of the project and the overall aim. It was then illustrated how we want to achieve this goal and how they could contribute (10 min). Next, we asked them predefined questions regarding their person (responsibility, experience, risk perception) and the organizational context of their work (procedures, safety concepts, documentation, and training) (20min).

The majority of the time (~1 h) was used to conduct a detailed task analysis and graphically summarize results in a mental model. The model comprises all tasks conducted by the experts when they evaluate the avalanche danger. To ensure that all mental models become comparable and could be integrated into one generic mental model, we asked similar questions (details see Section 3.1). Expected answers were preprinted on sticky notes and blank notes were available to add new thoughts/ideas. Whenever an expert mentioned a task, it was added to the mental models and rated according to a scale: very important, important and less important. By picking those tasks which had been selected by all five experts and rated by at least three experts as very important, the generic mental model was arranged.

The interviews were also used to identify the human factors that strongly influence the EWS reliability. In the last 30 min we stimulated an open-ended discussion by asking them to talk about their experience, past incidents and to name factors that influence their performance. We summarized all answers in a list to identify factors mentioned several times. In doing so, we were able to identify EWS specific PSFs (personal and organizational). The interview results provided a powerful basis to establish a BN representing the mental model and factors that have major influence on the EWS reliability. We use an extended BN to demonstrate how human factors can be incorporated in a quantitative reliability assessment of EWS.

3. RESULTS

3.1 *Task analysis: Mental model approach*

From the structured interviews we could identify and rate activities and cognitive tasks, which are regularly conducted by avalanche experts and should be integrated in our generic mental model. In Tbls.1-5 we summarize the answers selected by the experts for each question (no. answer) and indicate how often an answer is rated as very important (+++) important (++) or less important (+).

Question 1: Which information sources do you use to evaluate the avalanche danger in extreme situations? As expected, all experts consider information from weather forecasts, data from weather radars and data regularly measured by automated snow and weather stations as very important (Tbl. 1). Besides that, they strongly rely on regular measurements from local observers. Some also profit from information derived from an informal network. Information from the national avalanche bulletin is important for evaluating the general situation. While being in touch with experts from the national warning service is considered to be important by three experts, exchange with experts from the national meteorological warning service as well as online available snow profiles are considered less important.

Tbl. 1: Data sources used

	No.	+++	++	+
Weather forecast	5	5		
Automated stations	5	5		
Radar data	5	4	1	
Observer measurements	5	3	1	1
Informal network	5	1	3	1
Bulletin	5		5	
Expert SLF	3	2	1	
Expert MeteoSwiss	3		2	1
Available snow profiles	3		2	1

Question 2: Which activities do you conduct in order to evaluate the avalanche danger? Interestingly activities or own measurements are rated comparatively low (Tbl. 2). Experts justify lower ratings because they are often not able to conduct own measurements or observations in extreme situations.

Tbl. 2: Own measurements/actions conducted

	No.	+++	++	+
Own observations	4	2	2	
Amount new snow (top)	2	2		
Amount new snow (valley)	3	1	1	1
Make snow profile	1		1	

Question 3: Please name important hints given by avalanches. Experts consider information on recent avalanche activity as especially important (Tbl. 3). The importance of results from explosive control varies depending on locations and the availability of such temporary avalanche control measures. Some also consider the avalanche history and indicator avalanches as an important source of information.

Tbl. 3: Hints given by avalanches

	No.	+++	++	+
Recent avalanche activity	5	4	1	
Result explosive control	5	2	2	1
Avalanche history (winter)	4	2	2	
Indicator avalanches	2		2	

Question 4: Which snow and weather parameters do you evaluate? Experts mainly consult the amount of new snow and wind related snowdrift (Tbl. 4). The snow temperature and radiation are only important during spring. The snow height is important, while the air humidity is not of interest to most experts.

Tbl. 4: Snow and weather parameters evaluated

	No.	+++	++	+
Amount fresh snow	5	5		
Wind/snowdrift	5	4	1	
Air temperature	5	1	2	2
Radiation/sunshine	5		4	1
Snow height	5		4	1
Snowpack temperature	4		3	1
Air humidity	1		1	

Question 5: Which other relevant parameters do you evaluate? The stability of the snowpack and the endangered aspect are very important for all (Tbl. 5). Ratings for other parameters, such as avalanche type and the condition of structural avalanche protection measures vary strongly.

Question 6: How do you define the final danger level? All experts assess the likelihood of an avalanche and conditional on this its size. Therefore the avalanche danger is only high, if an expected event has the potential to reach infrastructures.

Tbl. 5: Other parameters evaluated

	No.	+++	++	+
Endangered aspect	5	5		
Stability snowpack	5	4	1	
Surface roughness	5	1	3	1
Avalanche type	5	1	2	2
Similar situation	5	1	2	2
Fracture depth	4	3	1	
Condition perm. structures	4	1	3	
Avalanche time	4		4	
Endangered elevation	3	1	1	1
Steepness release area	3	1	2	
No of prone-locations	3	2		1

Fig. 1 shows our generic mental model, in which all very important tasks (+++) are summarized.



Fig. 1: Generic mental model "evaluate danger".

3.2 Relevant human factors: EWS specific PSF

In the course of discussions at the end of the interviews, the participants shared their opinions on possible human errors in the context of avalanche EWS and personal as well as organizational factors influencing them.

According to the interviewees, deadly incidents occurred not more than once in their area and term of office. In one case an avalanche event was missed, because defense structures failed under heavy snow loads. Here, experts misinterpreted the conditions of snow supporting structures due to missing observations. In another case, experts assumed to have triggered an avalanche using explosive control. Again due to an inability to conduct field observations, experts incorrectly assumed that the measure was successful. Our interviews revealed that omitting certain tasks is not critical in their work but incomplete data sets are often misinterpreted. The ability of the experts to compensate missing or wrong information strongly depends on their personal characteristics. All interviewed experts estimated

that their decision making ability strongly depends on:

- Passion/ fascination for mountains/ snow
- Mountaineering abilities
- Daily contact with avalanches
- Familiarity with the area
- Knowledge of physical snow processes
- Experience

Moreover, experts mentioned certain characteristics, which they think help them to make good decisions:

- Risk type
- Gut feeling/intuition,
- Self-confidence
- Current physical condition
- Stress tolerance

All experts also stated that the organizational/external context is just as important as the personal background. Important external factors are:

- Data availability and uncertainty
- Team composition/ two man rule
- Decision-making power/ competencies
- Management structures/ failure culture
- Bias/ pressure
- Training facilities
- Work hours (24 h service)
- Procedures documented and trained
- Availability of decision aids/ checklists
- Temporal flexibility to set up measures
- Explosive control available

3.3 Human Reliability EWS: Quantification in BN

The influence of human factors on the EWS's reliability is modeled quantitatively in a decision graph (DG), which is a BN augmented with decision (grey, rectangular) and utility nodes (grey, diamond-shaped); see Fig. 2.

Green nodes in the DG represent critical parameters, which influence the probability of a "hazard event" (top node). The same parameters build the basis for the expert decision, whether to warn or not (lowest node). Three states (low, medium, high) are assigned to the parameters "danger", "size", "new snow", "snow drift" and "stability", while for the nodes "likelihood" and "aspect" two states (yes, no) are sufficient i.e. an event can either occur or not and the endangered aspect is either identified correctly or not.

In the conditional probability tables (CPTs), which are assigned to the nodes, the probability distribution of the respective node conditional on its par-

ents is defined. In Fig. 3 the CPT of the node "hazard event" conditional on the avalanche "danger" is illustrated.

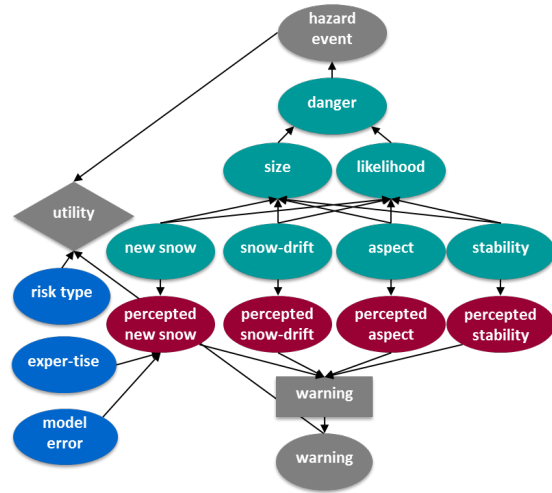


Fig. 2: Decision graph to model the influence of human factors on the EWS reliability.

danger		low	medium	high
hazard event	yes	0.2	0.5	0.8
	no	0.8	0.5	0.2

Fig. 3: CPT of the node "hazard event" includes estimated conditional probabilities (best-guess).

Red nodes in the DG represent the perception of the expert when evaluating critical parameters. To account for varying information sources (bottom line of the generic model in Fig. 1) considered by the experts when evaluating one certain parameter, several red child nodes were added (Fig. 4). For example, when an expert evaluates the amount of "new snow" information (possibly at different points in time) from the weather forecast ("new snow forecasted"), from radar images ("new snow measured") and from the observer network ("new snow observed") is considered.

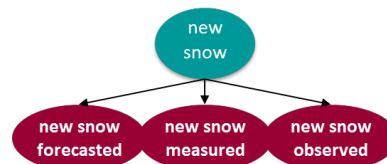


Fig. 4: For the evaluation of the parameter "new snow" experts consider varying information sources.

The perception of the expert and his/her ability to correctly evaluate the actual state of a critical parameter (blue nodes) is strongly influenced by personal and organizational factors, which are added in form of blue nodes in the DG. In Fig. 5 the node

"new snow forecasted" is shown together with its parent nodes; and an extract of the corresponding CPT is displayed below. The red node represents the ability of the expert to forecast the amount of new snow conditional on his/her expertise and the accuracy of the forecast model. If the expertise and the uncertainty involved in the weather forecast model are low, the probability of the expert to select the correct state (low, medium, high) is high (0.7); however, if an expert is inexperienced and the model uncertainty high, the experts will choose the forecasted new snow amount randomly.

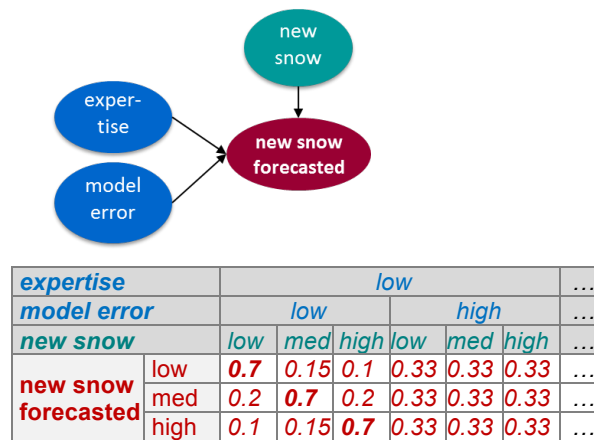
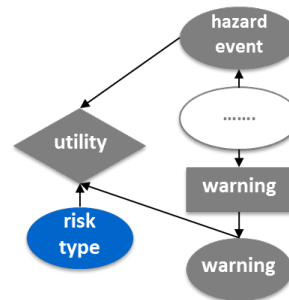


Fig. 5: Parent nodes and CPT of the node "new snow forecasted" with probabilities (best-guess).

The diamond-shaped utility node is added in the DG to represent the subjective costs and not necessarily monetary costs, associated to a decision scenario (Fig. 6). In this context, values for four scenarios are specified: hit (warning: yes; event: yes), false (warning: yes; event: no), miss (warning: no; event: yes) and neutral situation (warning: no; event: no). Moreover, the risk perception of the avalanche expert has an influence on these subjective costs i.e. a risk taking person considers a false alarm worse than a risk averse person.

By computing the DG one can assess the reliability of an EWS with respect to personal and human factors. The reliability is quantified in terms of the POD and PFA. The average POD (respectively PFA) is thereby defined as the probability of issuing a warning conditional on (not) having a hazard event. The assumption behind this is that the expert makes an optimal decision (warning/no warning) meaning that conditional on his/her state of knowledge at that particular point in time, the decision which maximizes the expected utility is chosen.



risk type	taking				averse			
	yes	no	yes	no	yes	no	yes	no
warning	yes	no	yes	no	yes	no	yes	no
event	yes	no	yes	no	yes	no	yes	no
utility	0	-10	-1000	0	0	-5	-1000	0

Fig. 6: The utility is defined conditionally on the risk type and the decision scenario.

4. DISCUSSION

In this contribution we present a structured approach to include human factors in the reliability assessment of EWS. First we develop a mental model in which all relevant tasks are summarized. Then we identify the factors determining successful completion of a task; finally we summarize both results in a DG to demonstrate the reliability assessment. Accounting for human factors is essential for a comprehensive EWS evaluation. In the majority of cases, the final warning decision is made by experts, who are increasingly exposed to external pressure. The reliability of EWS not only depends on the detection of hazard events, but also on the avoidance of false alarms.

Our contribution is a first step towards a comprehensive evaluation of EWS and the proposed DG has been proven to be a powerful tool for a structured assessment of human factors. We can combine the DG with our existing approach, see introduction, which can be applied to address both qualitative and quantitative of an EWS evaluation.

From a qualitative perspective, the DG clearly shows the role of experts in the EWS as well as the interaction of human factors and between humans and technical system components. The DG is derived from a mental model developed from five structured interviews conducted with avalanche experts. To provide a more sophisticated model a greater number of interviews with international experts should be conducted. The Swiss experts presented herein are closely connected and have all undergone a similar education, thus a degree of unity in the results can be anticipated. However, the framework recently published by Statham et al. (in review) reveals that Canadian

and US avalanche experts assess similar parameters to determine the avalanche danger. Furthermore, our model is flexible enough to account for additional needs.

Additionally the presented DG facilitates a quantitative reliability assessment, wherein the influence of varying human and organizational factors on the reliability of the EWS and optimal decision strategies can be identified. In order to demonstrate opportunities associated with a DG, we used numbers based on best guesses. To be able to identify more sophisticated data for the CPTs, e.g. to quantify the value of an experienced avalanche expert, further studies need to be conducted. As suggested by Rheinberger (2013), a collective database, in which decisions (warning yes/no) and final outcomes (event yes/no) are stored and evaluated, should be made available in the future. A more efficient way to gain quantitative data regarding personal and external factors could be to test both experts and less experienced forecasters in simulators. Such an approach is common in HRA, whereby experts receive selected data, and are confronted with varying scenarios; during which their decisions are recorded. Results from databases or simulators could then be used to enhance our DG.

5. CONCLUSION

This contribution presents a first step towards a comprehensive method to evaluate the EWS effectiveness. We show that BNs are a powerful tool to quantitatively assess the human influence on the reliability of EWS. In the BN the reliability (POD/PFA) can be modeled conditional on personal and external influence factors. This is the expert's expertise, their risk propensity and the underlying model uncertainties. Our results look promising and we believe that further studies should be initiated to facilitate a comprehensive evaluation of EWS in the context of an integrated risk management approach.

ACKNOWLEDGEMENTS

We thank all avalanche experts (Jon Andri Bisaz, Peder Caviezel, Gian Darms, Hanspeter Hefti and Hans Martin Henny) who actively participated in the interviews and provided the basis for our results. Moreover, we thank Lukas Stoffel (SLF) and Pascal Hägeli (Simon Fraser University) for their valuable inputs. Finally, we acknowledge Daniel Straub (ERA group/ TUM) as well as Luca Podofilini and Vinh Dang (PSI) for sharing their knowledge on HRA and BNs.

REFERENCES

- Aguilera, P. A., A. Fernández, R. Fernández, R. Rumí, and A. Salmerán, 2011: Bayesian networks in environmental modelling. *Environmental Modelling and Software*, **26**, 1376-1388.
- Bell, J., and J. Holroyd, 2009: *Review of human reliability assessment methods*. Health and Safety Execution.
- Boring, R. L., 2012: Fifty years of THERP and human reliability analysis. *Proceedings of the Probabilistic Safety Assessment and Management and European Safety and Reliability Conference (PSAM 11 & ESREL 2012)*.
- Bostrom, A., R. E. Morss, J. K. Lazo, J. L. Demuth, H. Lazrus, and R. Hudson, 2015: A mental models study of hurricane forecast and warning production, communication, and decision making. *Weather, Climate, and Society*, 111-129.
- Breznitz, S., 1989: *Cry wolf: The psychology of false alarms*. Erlbaum.
- Bründl, M., H. Romang, N. Bischof, and C. Rheinberger, 2009: The risk concept and its application in natural hazard risk management in Switzerland. *Natural Hazards and Earth System Sciences*, **9**, 801-813.
- Bründl, M., H.-J. Etter, M. Steiniger, C. Klingler, J. Rhyner, and W. Ammann, 2004: IFKIS-a basis for managing avalanche risk in settlements and on roads in Switzerland. *Natural Hazards and Earth System Science*, **4**, 257-262.
- Di Pasquale, V., R. Iannone, S. Miranda, and S. Riemma, 2013: *An overview of human reliability analysis techniques in manufacturing operations*. INTECH Open Access Publisher.
- Doswell III, C. A., 2004: Weather forecasting by humans-Heuristics and decision making. *Weather and Forecasting*, **19**, 1115-1126.
- Downton, M. W., R. E. Morss, O. V. Wilhelmi, E. Grunfest, and M. L. Higgins, 2005: Interactions between scientific uncertainty and flood management decisions: Two case studies in Colorado. *Global Environmental Change Part B: Environmental Hazards*, **6**, 134-146.
- Groth, K. M., and A. Mosleh, 2012: A data-informed PIF hierarchy for model-based Human Reliability Analysis. *Reliability Engineering & System Safety*, **108**, 154-174.
- Groth, K. M., and L. P. Swiler, 2013: Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H. *Reliability Engineering & System Safety*, **115**, 33-42.
- Gubler, H. U., 2000: Five years experience with avalanche-, mudflow-, and rockfall-alarm systems in Switzerland. *Proceedings of the international snow science workshop (ISSW), Big Sky, Montana*.
- Guzzetti, F., 2015: Forecasting natural hazards, performance of scientists, ethics, and the need for transparency. *Toxicological & Environmental Chemistry*, 1-17.
- Jensen, F. V., and T. D. Nielsen, 2007: *Bayesian networks and decision graphs*. Springer Science + Business Media.
- Johnson, J., P. Haegeli, J. Hendrikx, and S. Savage, 2016: Accident causes and organizational culture among avalanche professionals. *Journal of Outdoor Recreation and Tourism*, **13**, 49-56.
- Kirwan, B. and L. K. Ainsworth, 1992. A guide to task analysis: the task analysis working group. Taylor & Francis Ltd, London.

- Kirwan, B., 1994: A guide to practical human reliability assessment. Taylor & Francis Ltd, London.
- Klein, G.A., 1998: Sources of Power: How People Make Decisions. MIT Press.
- Langseth, H., and L. Portinale, 2007: Bayesian networks in reliability. *Reliability Engineering & System Safety*, **92**, 92-108.
- Margreth, S., and H. Romang, 2010: Effectiveness of mitigation measures against natural hazards. *Cold Regions Science and Technology*, **64**, 199-207.
- McClung, D., 2002a: The elements of applied avalanche forecasting, Part I: The human issues. *Natural Hazards*, **26**, 111-129.
- McClung, D., 2002b: The elements of applied avalanche forecasting, Part II: the physical issues and the rules of applied avalanche forecasting. *Natural Hazards*, **26**, 131-146.
- Mkrtchyan, L., L. Podofillini, and V. N. Dang, 2015: Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability Engineering & System Safety*, **139**, 1-16.
- Morgan, M. G., 2002: *Risk communication: A mental models approach*. Cambridge University Press.
- Morss, R. E., O. V. Wilhelmi, M. W. Downton, and E. Grunfest, 2005: Flood risk, uncertainty, and scientific information for decision making: lessons from an interdisciplinary project. *Bulletin of the American Meteorological Society*, **86**, 1593.
- Morss, R. E., J. L. Demuth, A. Bostrom, J. K. Lazo, and H. Lazrus, 2015: Flash flood risks and warning decisions: a mental models study of forecasters, public officials, and media broadcasters in Boulder, Colorado. *Risk analysis*, **35**, 2009-2028.
- Rheinberger, C. M., 2013: Learning from the past: statistical performance measures for avalanche warning services. *Natural Hazards*, **65**, 1519-1533.
- Sättele, M., 2015: Quantifying the Reliability and Effectiveness of Early Warning Systems for Natural Hazards, Technische Universität München TUM.
- Sättele, M., M. Bründl, and D. Straub, 2016: Quantifying the effectiveness of early warning systems for natural hazards. *Natural Hazards and Earth System Sciences*, **16**, 149-166.
- Schweizer, J., and P. M. Föhn, 1996: Avalanche forecasting: an expert system approach. *Journal of Glaciology*, **42**, 318-332.
- Sorensen, J. H., and D. Mileti, 1987: Decision making uncertainties in emergency warning system organizations. *International Journal of Mass Emergencies and Disasters*, **5**, 33-61.
- Spurgin, A. J., 2010: *Human reliability assessment theory and practice*. CRC Press.
- Statham, G., P. Haegeli, E. Greene, K. Birkeland, C. Israelson, B. Tremper, C. Stethem, B. McMahon, B. White, J. Kelly, in review: The Conceptual Model of Avalanche Hazard. *Natural Hazards*.
- Stewart, T. R., P. J. Roebber, and L. F. Bosart, 1997: The importance of the task in analyzing expert judgment. *Organizational Behavior and Human Decision Processes*, **69**, 205-219.
- Stoffel, L., and J. Schweizer, 2008: Guidelines for avalanche control services: organization, hazard assessment and documentation—an example from Switzerland. *Proceedings ISSW 2008. International Snow Science Workshop, Whistler BC, Canada, 21-26 September 2008*.
- Straub, D., 2005: Natural hazards risk assessment using Bayesian networks. *9th International Conference on Structural Safety and Reliability ICOSSAR'05*, G. Augusti, G. I. Schuëller, and M. Ciampoli, Eds., Millpress.
- Straub, D., and A. Der Kiureghian, 2010: Bayesian network enhanced with structural reliability methods: Methodology. *Journal of engineering mechanics*, **136**, 1248-1258.
- VDI, 2015: *Human reliability - Methods for event analysis regarding human behaviour*. VDI-Gesellschaft Produkt- und Prozessgestaltung (GPP) Fachbereich Zuverlässigkeit.
- Weber, P., G. Medina-Oliva, C. Simon, and B. lung, 2012: Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*, **25**, 671-682.
- Zweifel, B., and P. Haegeli, 2014: A qualitative analysis of group formation, leadership and decision making in recreation groups traveling in avalanche terrain. *Journal of Outdoor Recreation and Tourism*, **5**, 17-26.
- Zwirgmaier, K., D. Straub, and K. Groth, submitted: Capturing cognitive causal paths in human reliability analysis with Bayesian network models. *Reliability Engineering and System Safety*.