Bayesian Networks for Assessing the Reliability of a Glacier Lake Warning System in Switzerland

Rouven A. STURNY[1,2] and Michael BRÜNDL[1*]

[1] WSL Institute for Snow and Avalanche Research SLF (Davos, Switzerland)
2 Silvaplus Sàrl (Martigny, Switzerland)
*Corresponding author. E-mail: bruendl@slf.ch

Warning and alarm systems have become important preventive measures for the management of risks caused by Alpine natural hazards. However, little is known about how to determine the reliability of these systems. We present a reliability analysis in which Bayesian Networks (BN) are used to model a warning system that is operated to detect a glacier lake outburst flood in Switzerland. Our study revealed that during the summer months of 2009, the alerting unit 'visual acoustic signal' (VAS) has a reliability of 0.94 and the unit 'alerting of the intervention entities' (AIE) has a reliability of 0.83. During the summer months, the probability for a major flood event is 0.0055 per day. Consequently the probability for a working VAS is 0.99967 and 0.99906 for the AIE. Individual system components which are able to cause a sure failure in one of the alerting units have a reliability of approximately 1. The investigated warning system is highly reliable in detecting dangerous floods due to numerous redundancies.

Key words: warning system, reliability, alerting, glacier lake outburst, Bayesian Networks

1. INTRODUCTION

Together with structural mitigation measures, land-use planning and biological measures (e.g. protection forests), warning and alarm systems have become very important in the management of risks caused by Alpine natural hazards such as snow avalanches, debris flows, rock fall and landslides [*Romang et al.*, 2009*; Sättele et al.*, 2012*; Intrieri et al.*, 2012]. Safety in mountain regions can be improved by well-functioning warning and alarm systems. However, there is a lack of experience in evaluating the reliability of warning and alarm systems. In this paper we present a reliability analysis of a warning system, which is installed for a timely detection of glacier lake outburst floods (GLOF) in the Swiss Alps.

2. WARNING SYSTEM GRINDELWALD

The considered glacier lake is situated below the Lower Grindelwald Glacier in the Bernese Alps in Switzerland. It hit the headlines in both 2008 and 2009 due to the potential for large floods. In the summer months the level of the lake can drain rapidly and can cause floods which are potentially affecting residents and tourists in the valley.

To mitigate the risk, a warning system has been installed (**Fig. 1**). It consists of several components, which are located in different areas and can be subdivided into the following units [*Sättele et al.*, 2013]: monitoring, data management, warning unit, power supply and diagnostic units. The monitoring unit includes five level meters: mainly two in the lake, two in the gorge downstream and another one at a bridge at the gorge's exit. The data management unit consists of different servers that communicate with the monitoring unit via mobile and fixed network and via radio. The warning unit is divided in two main elements: two automated visual-acoustic signals (VAS) in the gorge and the alerting of intervention units (AIU), *i.e.* the fire brigade. The power supply for most system components is redundant. Diagnostic units check battery voltage and radio connection. They are able to detect irregularities and prevent unnecessary maintenance and repair actions.

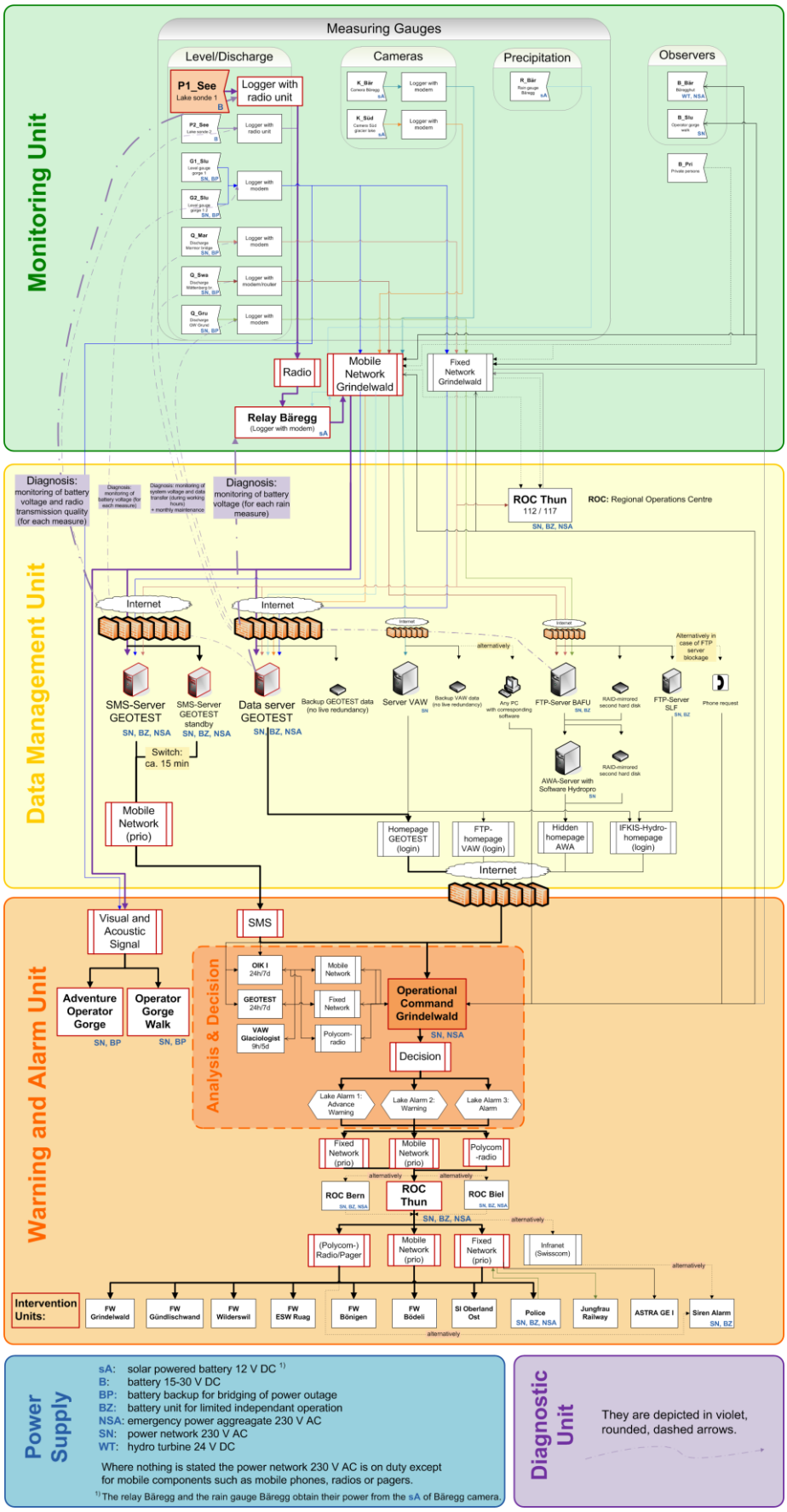


Fig. 1 Overview on the warning system with all components and the corresponding data flow. The red boxes show the important components from the level meter "P1_See" to the alerting units. A graph with higher resolution can be provided by the corresponding authors upon request. Based on Hählen [2012, pers. communication] and Meier [2012, pers. communication].

3. APPLYING BN ON WARNING SYSTEMS

For the assessment of the reliability of the warning system we have chosen Bayesian Networks (BN). A BN is a graphical method that allows the quantification of the reliability of technical systems by combining sub-system reliabilities. In a BN, uncertain components are represented by nodes and informational or causal dependencies between the nodes by arcs [*Pearl,* 1988]. The relation of nodes and child-nodes are expressed in conditional probability tables (CPT) [*Jensen and Nielsen,* 2007]. They were recently used for assessing natural hazard risks and the reliability of warning systems [*Straub* 2005*; Grêt-Regamey and Straub,* 2006*; Sättele et al.,* 2013].

We have chosen BN for the following reasons. Incomplete knowledge can be included in a straightforward way: the relations between different nodes and the state of the nodes without parent nodes must be known [*Jensen and Nielsen,* 2007]. Additionally, realistic values for all Conditional Probability Tables must be identified (CPT; **Fig. 2**). Another strength of BN is that common cause failures can be integrated which is hardly possible with other quantifiable methods such as fault tree analysis [*Bobbio et al.,* 2001]. In addition, BN would allow encompassing human factors [*Gregoriades et al.,* 2010].

There are different approaches to design a BN. In the present study, the bottom-up approach [*Langseth and Portinale,* 2006] has been used. It enabled us to stick close to the system overview (**Fig. 1**). First, we defined the nodes, then we determined the relations and the states of individual components to complete the CPTs in the BN.

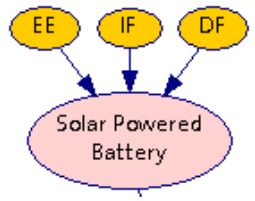
The BN has been developed by using the software GeNIe 2.0 [*University of Pittsburgh,* 2012]. It comes with a freeware license and has proven to be clearly arranged and efficient in use.

3.1 Definition of nodes, states and CPT

Before the nodes and the corresponding states are defined, a standardized schema for integrating all possible kind of failures is developed. We assume that technical components can have three types of parent nodes (**Fig. 2**): external event (EE), internal failure (IF) and diagnostic system failure (DF). Depending on the location of a device, external events such as storms, lightning, landslides, avalanches, sabotage etc. are taken into account. IF are device internal failures that can be derived from Mean-Time-Between-Failure (MTBF) values. DF are particularly related to systems that monitor battery voltages and radio communication quality.

For all nodes of the BN (**Fig. 5**), we define two states. EE, IF and DF can be in the states "yes"/"no" whereas all the other nodes can be either "working"/"faulty" or "correct"/"incorrect".

Relationships between nodes that represent technical system components and their failure causes can be described by 0/1 in the CPT. The CPTs of *e.g.* a solar powered battery and its parent nodes (EE, IF, DF) are illustrated in **Fig. 2**.



yes	0.0215
no	0.9785

EE	yes				no			
IF	yes		no		yes		no	
DF	yes	no	yes	no	yes	no	yes	no
working	0	0	0	0	0	0	0	1
faulty	1	1	1	1	1	1	1	0

Fig. 2 *Top left:* Sample node with the three failure causing parent nodes 'external event' (EE), 'internal failure' (IF) and 'diagnostic system failure' (DF); *top right:* CPT for 'EE' in the left image – derived from recorded past events of the warning system; *bottom:* corresponding CPT for the sample node "Solar Powered Battery".

The CPTs of nodes that represent non-technical components in the BN require a more elaborate proceeding because the parent nodes must be weighted. For the generation of a warning, the information on the water level in the lake is much more important than the information on the water level in the river further down the gorge due to the time it takes the water to flow from the lake downwards.

3.2 Determining the reliability of components

In order to complete the CPTs of the first level nodes EE, IF and DF, information on the corresponding failures is necessary. All data were derived from experience for one summer season when the warning system was in use. **Table 1** provides an overview on the sources.

4. RESULTS

4.1 Failure rate of components

In **Table 2** we exemplarily show the failure rates which we calculated for the parent nodes of different components, which show (1) that

external events for the radio and battery unit are taken into account in the base station where they are situated; and (2) that diagnostic system failures are only present in the radio and battery unit.

Table 1 Main sources for the determination of the reliability of components.

External events (EE)	Events that actually happened in the investigation area, mainly in 2009: reports, analysis of pictures from observation camera
Internal failures (IF)	Manufacturer data (MTBF)
Diagnostic system failure (DF)	Assumptions
Power and communication networks	Operator data based on past events

Table 2 Failure rates for the level meter P1 in the lake for a time frame of one summer season (6 months).

Sonde (incl. cable to base station)	EE: 0.009804 (1.5/153d)	1.5 days of failure due to heavy snow fall and buoyancy of an ice floe [Hählen, 2012, pers. communication]
	IF: 0.002457 (4/407/2y/2)	In 2 years, 4/407 pieces suffered from failure [Stürm, 2012, pers. communication]
Base station with logger	EE: 0.006536 (1/153d)	1 day of failure due to a mobile communication problem [Hählen, 2012, pers. communication]
	IF: 0.000514 (1/975y/2)	MTBF: 1 failure per 975 y [Campbell Scientific, 2012]
Radio	IF: 0.043920 (1/(100000h/24h)*183d)	MTBF: 1 failure per 100'000 h [openPR, 2008]
	DF: 0.004	Assumption: radio quality measurement is misleading or failure cannot be repaired immediately due to var. reasons
Battery	IF: 0.000003	Failure probability for this type of component [Moss, 2005 in Franke (2011)]
	DF: 0.004	Assumption: voltage measurement is misleading or failure cannot be repaired immediately due to var. reasons

4.2 Weighting of nodes

The weighting of nodes has been carried out by assessing the importance of the corresponding parent nodes. The visual-acoustic signal with its six parent nodes *e.g.* receives four signals from two lake and two gorge level meters. In addition, failures caused by external events and internal failures must be considered. Due to the fact that the sensors in the lake are located further upstream, chances are very high that people visiting the gorge can be warned by the visual-acoustic signal. The lead time from a possible lake outburst until the flood causes damage is approximately 20 minutes. For a successful alert the following probabilities can be assigned to the CPT of the visual-acoustic signal node:

(a) at least one signal from a lake level meter is received: 1.0
(b) the upper gorge level meter signal is received: 0.9
(c) the lower gorge level meter signal is received: 0.7
(d) no signal is received: 0.0

4.3 Resulting values

After completing the CPT for all nodes the updated BN can be used to calculate the reliability values for all components (**Fig. 3** and **Fig. 4**). The full BN is shown in **Fig. 5**.

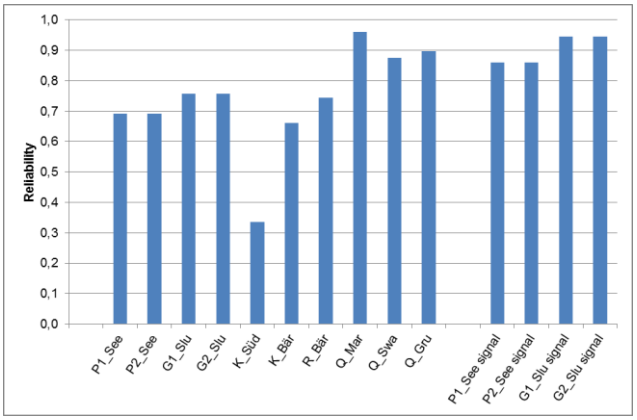


Fig. 3 Reliability values calculated by the BN - part 1. The elements on the x-axis correspond to the monitoring components, cf. **Fig. 1** and **Fig. 5**.

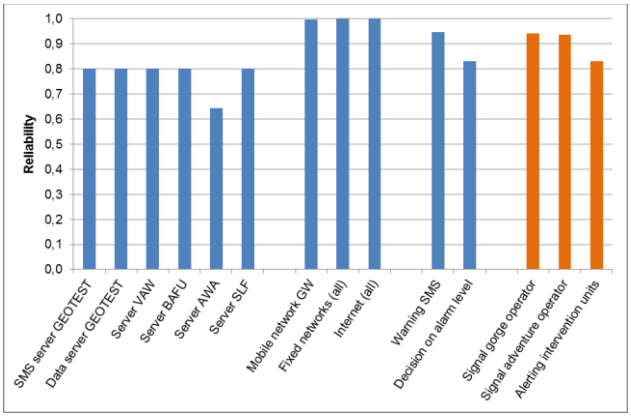


Fig. 4 Reliability values calculated by the BN - part 2. The elements on the x-axis correspond to the data management and warning and alarm components, cf. **Fig. 1** and **Fig. 5**.

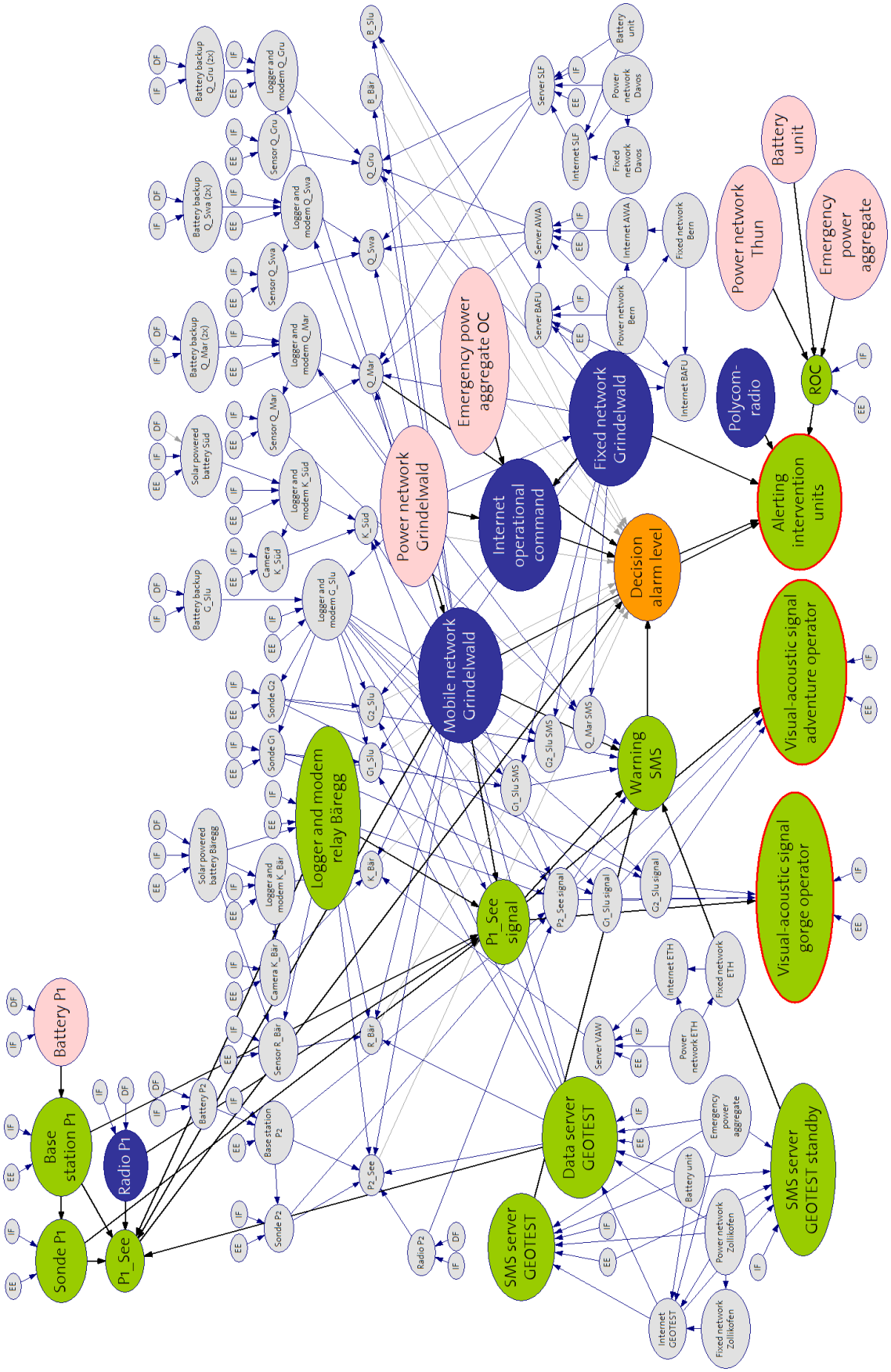


Fig. 5 Bayesian Network for the glacier lake warning system. The coloured nodes show the important components from the level meter "P1_See" to the alerting units. A graph with higher resolution can be provided by the corresponding authors upon request.

The two alerting signals have almost the same reliability of 0.94 whereas the alerting of the intervention units reaches a reliability of only 0.83 (orange bars in **Fig. 4**). A critical situation arises when the failure coincides with a flood event. The probability of an event to occur within the considered time period is 0.0055 per day. It was calculated by dividing the number of events that occurred in the summer of 2009 by the number of days the warning system was activated (1/183d). Given that the probability of occurrence is independent of a system failure, we multiply the two values:

$$P(signal) = 0.0055 * (1 - 0.94) = 0.00033$$
$$P(alerting) = 0.0055 * (1 - 0.83) = 0.00094$$

For a summer season, the probability that an event occurs at the same time as the corresponding warning units fail – which is the worst case – is 3.3 x 10^{-4} and 9.4 x 10^{-4}, respectively.

4.4 Scenarios

By setting one or more values to a specific state, we can generate scenarios. If we assign the state "faulty" to the node "Base station P1", we can model the consequences for other nodes of the system. This approach applies for parent and for child nodes. The development of scenarios allows the investigation of the influence of individual components on relevant nodes. We are interested in the consequences on the two nodes "signal adventure operator" and "alerting of intervention units" (**Fig. 6**). The node "signal gorge operator" is not further considered, because the two visual-acoustic signals have very similar reliability values (**Fig. 4**).

In general, the node "signal adventure operator" (SAO) is more resilient than the "alerting of intervention units" (AIU; **Fig. 6**): the failure of the mobile network Grindelwald, the power network of Grindelwald, and the logger including the modem of the relay Bäregg lead to a reliability reduction of 0.11 (0.94 - 0.83). The failures of the four sensors in the lake and in the gorge cause a sure failure (reliability = 0) of the SAO whereas the reliability of AIU is still 0.39. For all other examined scenarios, the AIU has lower reliability values: *e.g.* the non-availability of the fixed network leads to a reduction of 0.52 (0.83 - 0.31). The failure of the operational command's internet and the breakdown of the regional operations centre (ROC) that manages emergency calls causes a sure failure of the AIU.

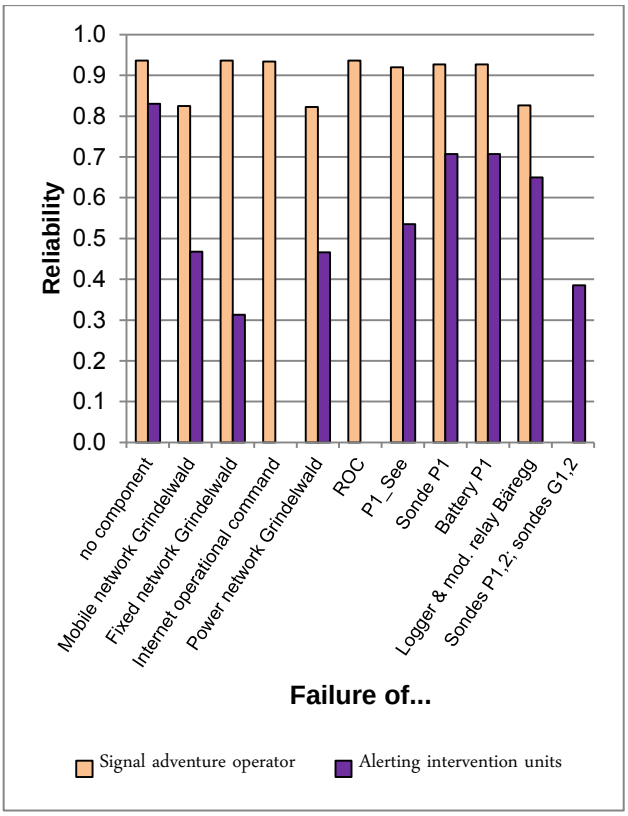


Fig. 6 Scenarios for the two main alerting units: the reliability values are depicted for the failure of different components of the warning system.

All the components which are able to cause a sure failure of one of the alerting units alone have a high reliability close to one.

4.5 Sensitivity analysis

In order to assess the influence of certain nodes on the alerting nodes, the CPT of the following nodes containing the most uncertain assumptions have been modified: the server software, all diagnostic system failures and the external event of the signal adventure operator. The results are shown in **Table 3**.

Table 3 Results of the sensitivity analysis conducted for the least certain input values.

Compo-nent	Initial fail.prob.	Scenario fail.prob.	Consequence on	
			SAO	AIU
Initial calculation	-	-	9.4 x 10^{-1}	8.3 x 10^{-1}
Server software	1 x 10^{-1}	5 x 10^{-2}	9.4 x 10^{-1}	8.6 x 10^{-1}
	1 x 10^{-1}	2 x 10^{-1}	9.4 x 10^{-1}	7.7 x 10^{-1}
All DF	4 x 10^{-3}	2 x 10^{-3}	9.4 x 10^{-1}	8.3 x 10^{-1}
	4 x 10^{-3}	8 x 10^{-3}	9.4 x 10^{-1}	8.3 x 10^{-1}
EE of SAO	1 x 10^{-2}	5 x 10^{-3}	9.4 x 10^{-1}	8.3 x 10^{-1}
	1 x 10^{-2}	2 x 10^{-2}	9.3 x 10^{-1}	8.3 x 10^{-1}

Variations within the reliability values of the alerting nodes are low. The most important influence arises from the server software. Doubling its original failure probability (0.1→0.2) leads to a reduction of the AIU's reliability by 0.06 (0.83 - 0.77).

5. DISCUSSION

In the case study we illustrated, how reliability data for the different components can be derived from various sources (**Table 1**). Nevertheless, for some components, assumptions had to be made. These assumptions were tested regarding their influence on the final results and most of them have proven to exert a minor influence. Only the reliability of the server software has a considerable influence on the reliability of the alerting units. Therefore, the call for fostered research in this domain by Kashi and Nachiappan [2010] is confirmed.

The Mean-Time-Between-Failure (MTBF) rate is a statistical measure for hardware components. Generally, these values are calculated by the manufacturer by extrapolating failure data from the past or stress tests – it cannot be excluded that some values are optimised [*Elerath and Shah*, 2004]. Nevertheless, comparing several similar components produced by different manufacturers, *e.g.* loggers, revealed that the various values correspond well to the order of magnitude which we consider as precise enough for the goal of this study. The second reason for the use of MTBF values is the integration of external events in the BN. They are based on real events and for nearly all components they are higher than the MTBF values – except for the servers. Their placing in protected server rooms entails that an external event, *e.g.* flooding, fire etc., is less probable than an internal failure.

The possibility to weigh nodes has proven to be an extraordinary strength of BN: it enabled us to define the CPT of four nodes by incorporating the different importance of their parent nodes. This was done for the nodes "visual-acoustic signal" (2x), "warning SMS" and "decision on alarm level". All other CPT of nodes contain either 0/1 or IF/EE/DF probabilities (**Fig. 2**).

The node "alerting of intervention units" resulted in a lower reliability value (P = 0.83) than the "visual-acoustic signals" (P = 0.94) mainly because of the data detour via the servers before the intervention units can be alerted. The value of 0.83 is *acceptable* because of the fact that a system failure has to coincide with the occurrence of a major event. This probability that the system works was calculated as

$$P = 1 - 9.4x10^{-4} = 9.9906x10^{-1}$$

for one summer season. Future studies are needed to be able to discuss the applicability on similar settings.

The scenario analyses revealed that a reliable component has a higher influence on the overall system reliability than a less reliable component. Two reasons are possible: Either the failure of a low reliability component leads to a smaller reduction in reliability of the system, or, the system was developed with more reliable components because of its important function. To these important components we have assigned higher weights, which explain the higher influence on the system as a whole.

No component can have a reliability of 1. The high reliability of the presented warning system is confirmed by the fact that all components which might cause a sure failure of one of the alerting units are close to zero.

Receiver Operating Characteristic Curves (ROC-Curves) are a method to depict the relation of probability of detection (POD) and probability of false alarms (PFA), which can be used as an indicator for the reliability of a warning systems [*Krzysztofowicz et al.*, 1998]. Since no false alarms were issued to the public in the past years, it was not possible to apply this method.

In the present study, human factors have been widely neglected. Thus, the following assumptions apply for all statements made: (1) the sensor positions and types measure the correct parameters; (2) the experience of decision makers, personal risk aversion and emotions affect the decision on the alarm level in a positive way; (3) the threshold values have been set accurately; and (4), evaluations of prevailing conditions are carried out correctly. Despite these assumptions, we judge the negligence of human factors in this study as justified because the operational command in Grindelwald consists of several experienced experts being highly familiar both with the warning system and the hazard processes in the catchment area. Nevertheless, future studies should include a detailed analysis of human factors.

6. CONCLUSION

The present study shows how Bayesian Networks (BN) can be used to conduct a quantitative reliability analysis for a warning system protecting several communities from glacier lake

outburst floods. Our study revealed that two important strengths of BN can be pointed out: firstly, once the structure of the BN is defined, knowledge on probabilities must only be available for the nodes not having any parent nodes, at least for our example. All other probabilities are calculated automatically after the weighing of the parent nodes have been assigned in the corresponding child node CPT. Secondly, the development of failure scenarios can be done in a straightforward way, *i.e.* we reveal the influence of single components – or of component combinations – on final nodes, such as alerting units.

The investigated warning system contains two main warning units: visual-acoustic signals ("signal adventure operator", SAO) and the alerting of intervention units (AIU). Our analyses resulted in a reliability of 0.94 for the SAO and of 0.83 for the AIU, for one summer season. A critical situation would only arise if a coinciding alerting unit fails at the same time an event occurs ($P = 5.5 \times 10^{-3}$); the probability is 3.3×10^{-4} and 9.4×10^{-4}, respectively.

The analysis of failure scenarios has shown that the failure of single components has a limited influence on the reliability of the alerting units. The overall high reliability is achieved through numerous redundancies in the monitor, data management and power supply units. The results imply that the data transmission of monitor units to several servers considerably increase the system reliability. Furthermore, it becomes evident that the use of several identical components increases the reliability and that the probabilities of external events are higher than internal failures.

For further studies it will be of high interest to reduce the number of assumptions, in particular for server software, to evaluate longer periods of data measurements, to study the role of sensor position and to integrate human factors quantitatively in order to fill the CPT yet more accurately. We therefore expect more studies engaging in the application of BN on other warning systems.

ACKNOWLEDGMENT: This paper summarises the results of a Master Thesis at ETH Zurich. We thank Prof. Dr. Hansruedi Heinimann, ETH Zurich, for fruitful and constructive discussions and Martina Sättele, WSL-SLF, for her contributions in an earlier phase of the project; finally, we thank three anonymous reviewers for their valuable comments.

REFERENCES

Bobbio, A., Portinale, L., Minichino, M. and Ciancamerla, E. (2001): Improving the analysis of dependable systems by mapping fault trees into Bayesian networks, Reliability Engineering & System Safety, Vol. 71, No. 3, pp. 249-260.

Campbell Scientific (2012): CR1000 data logger. www.campbellsci.co.uk/index.cfm?id=1156. Access: 04/08/2012.

Elerath, J.G. and Shah, S., (2004): Server class disk drives: how reliable are they?, Reliability and Maintainability, Annual Symposium - RAMS, pp. 151-156.

Franke, J. (2011): Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen, Mitteilungen am Institut für Wasserbau der Universität Stuttgart, Vol. 202, pp. 1-316.

Gregoriades, A., Sutcliffe, A., Papageorgiou, G. and Louvieris, P. (2010): Human-Centered Safety Analysis of Prospective Road Designs, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, Vol. 40, No. 2, pp. 236-250.

Grêt-Regamey, A. and Straub, D. (2006): Spatially explicit avalanche risk assessment linking Bayesian networks to a GIS, Nat. Hazards Earth Syst. Sci., 6(6), pp. 911-926.

Hählen, N. (2012): pers. communication.

Intrieri, E., Gigli, G., Mugnai, F., Fanti, R. and Casagli, N., (2012): Design and implementation of a landslide early warning system, Engineering Geology, 147–148(0), pp. 124-136.

Jensen, F. V. and Nielsen, T. D. (2007): Bayesian Networks and Decision Graphs. New York: Springer.

Krzysztofowicz, R., Kelly, K. S., and Long, D. (1994): Reliability of Flood Warning Systems. Journal of Water Resources Planning and Management, 120(6), pp. 906-926.

Meier, L. (2012): pers. communication.

openPR (2008): press release by Degetel Datenfunk GmbH. www.openpr.de/news/202086.html. Access: 04/08/2012.

Pearl, J. (1988): Probabilistic Reasoning in Intelligence Systems: Networks of Plausible Inference. Morgan Kaufmann, San Francisco, 552pp.

Romang H., Dufour F., Gerber M., Rhyner J., Zappa M., Hilker N. and Hegg C. (2009): Flood Warning in Small Catchments. In: P. Samuels; S. Huntington; W. Allsop and J. Harrop (editors), Flood Risk Management: Research and Practice. Taylor & Francis Group, London, pp. 1201-1207.

Sättele, M., Bründl, M. and Straub, D. (2012): Classification of warning systems for natural hazards. In: C. Moormann; M. Huber and D. Proske (editors), Proceedings of the 10th International Probabilistic Workshop, 15/16 November 2012. Mitteilung des Instituts für Geotechnik, Institut für Geotechnik der Universität Stuttgart, Stuttgart, pp. 257-270.

Sättele, M., Bründl, M.; Straub, D. (2013): Bayesian networks to quantify the reliability of a debris flow alarm system. In Deadatis, G; Ellingwood, B.R. and Frangopol, D.M. (editors), Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures, Proceedings of the 11th International Conference on Structural Safety & Reliability ICOSSAR. Taylor & Francis, New York, 3661-3668.

Straub, D. (2005): Natural hazards risk assessment using Bayesian networks. In: Augusti, G., Schuëller, G. I. and Ciampoli, M. (editors), 9th International Conference on

Structural Safety and Reliability ICOSSAR'05. Rome, Millpress.

Stürm, P. (2012): pers. communication.

University of Pittsburgh (2012). GeNIe 2.0. http://genie.sis.pitt.edu.Access : 17/04/2014.

Vishwanath, K.V. and Nagappan, N. (2010): Characterizing cloud computing hardware reliability, Proceedings of the 1st ACM symposium on Cloud computing (SoCC '10), New York, USA, pp. 193-204.