Quantifying the effectiveness of early warning systems for natural hazards

M. Sättele[1] and M. Bründl[1] and D. Straub[2,]

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

[2] Technische Universität München, Engineering Risk Analysis Group, Munich, Germany

Correspondence to: M. Sättele (saettele@slf.ch)

Abstract

Early warning systems (EWS) are increasingly applied as preventive measures within an integrated risk management approach for natural hazards. At present, common standards and detailed guidelines for the evaluation of their effectiveness are lacking. To support decision-makers in the identification of optimal risk mitigation measures, a three-step framework approach for the evaluation of EWS is presented. The effectiveness is calculated in function of the technical and the inherent reliability of the EWS. The framework is applicable to automated and non-automated EWS and combinations thereof. To address the specifics and needs of a wide variety of EWS designs, a classification of EWS is provided, which focuses on the degree of automations encountered in varying EWS. The framework and its implementation are illustrated through a series of example applications of EWS in an alpine environment.

1 Introduction

A growing number of early warning systems (EWS) is developed and operated for reducing the risks imposed by a wide range of natural hazard processes. They can mitigate the consequences of hazardous events if information is issued timely. In recent years, EWS technologies have been improved significantly. In many fields, EWS are now cost-efficient alternatives to structural mitigation measures. They are applied for large scale hazard processes, such as severe weather, floods, tsunamis, volcanic eruptions or wildfires, where they complement structural measures and support the preparation and response to the hazard events (e.g. Sorensen, 2000; Zschau and Küppers, 2003; Grasso and Singh, 2009; Glade and Nadim, 2014). They are also popular as flexible and temporary mitigation measures on smaller scales. In mountain regions, they are successfully applied to mitigate risks from snow avalanches, debris flows, flash floods, rockfalls and landslides (e.g. Bell et al., 2010; Thiebes, 2012; Michoud et al., 2013; Stähli et al., 2015) .

Whether or not EWS are effective and efficient risk mitigation measures can be evaluated case-specifically through cost-benefit analyses, in which the life-cycle costs and the efficiency is compared to those of alternative mitigation measures (Penning-Rowsell E., 2005; SafeLand, 2012; Špačková and Straub, 2015). In cost-benefit analyses, the efficiency is defined as the risk reduction achieved with a mitigation measure and is expressed in monetary values. To avoid expressing the risk in monetary terms, cost-effectiveness analyses can be conducted instead (Bründl et al., 2009). The effectiveness E_w is quantifiable without expressing the risk in monetary terms. For EWS, one can define it as a function of the overall risk without the EWS, R, and the risk with the EWS, $R^{(w)}$ (Sättele et al., 2015a):

$$E_w = 1 - \frac{R^{(w)}}{R} \qquad (1)$$

The risks with and without the EWS are evaluated by summing or integrating over all n_{scen} possible scenarios j and all n_{obj} exposed objects i, which are persons or assets exposed to a hazardous scenario:

$$R = \sum_{j=1}^{n_{scen}} \sum_{i=1}^{n_{obj}} R_{ij} \qquad (2)$$

Both R_{ij} and $Rr_{ij}^{(w)}$ can be calculated from the probability of occurrence of a hazard scenario, p_j, the probability of exposure of object i in scenario j, pe_{ij}, the vulnerability of object i in scenario j, v_{ij}, and the value of object i, A_i (Fuchs, 2006; Bründl et al., 2009):

$$R_{ij} = p_j \times pe_{ij} \times v_{ij} \times A_i \tag{3}$$

When issuing timely information, EWS can reduce the exposure probability of persons and mobile objects (Dai et al., 2002; SafeLand, 2012; Thiebes, 2012) or their vulnerability (Einstein and Sousa, 2006). Detailed guidelines on how this risk reduction can be evaluated have been published for structural mitigation measures (e.g. Romang (2008)) but, to the best of our knowledge, not for EWS.

Even without detailed guidelines, the effectiveness of EWS has been investigated previously. Thereby, it is common practice to consider both the probability that an EWS detects hazardous events, as well as the probability that the EWS leads to a false alarm. If the EWS detects a hazard event, timely warnings can initiate preventive actions, such as an evacuation of endangered persons to prevent damage. On the other hand, frequent false alarms can lead to excessive intervention costs or reduce compliance with future warnings (Pate-Cornéll, 1986; Grasso et al., 2007; Schröter et al., 2008; Rogers and Tsirkunov, 2011; Ripberger et al., 2014). To account for the probability that events are correctly detected (hit) and the probability that false alarms are issued (Fig. 1), the effectiveness is typically evaluated based on concepts of signal detection theory, where a classifier (in the simplest case a predefined threshold) discriminates between alarm and no alarm (Swets, 1996).

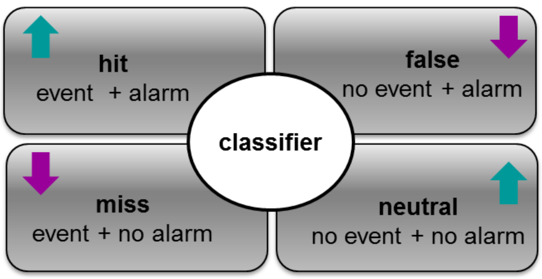

Figure 1: Following the principle of signal detection theory, a classifier (e.g. in form of a threshold) discriminates between correct and wrong outcomes of EWS: EWS correctly issues an alarm when an event occurs (hit) or no alarm when no event occurs (neutral), but can also wrongly issue false alarms or miss dangerous events.

An optimal EWS detects all hazardous events and never produces false alarms (Intrieri et al., 2013). In the operational application of EWS, false alarms cannot be avoided and an optimal trade-off between detected events and false alarm needs to be identified. To solve this optimization problem quantitatively, costs and utilities must be assigned to possible outcomes. Along these lines, Paté-Cornell (1986) suggests to optimize the effectiveness of fire warning systems operated in buildings in function of the probability that the event is detected (POD) and the probability that endangered persons comply with the warning (POC). The latter is modeled conditional on the probability of false alarms (PFA) by means of both descriptive (how do people react in real situations?) and normative (how should people optimally react?) approaches. In the normative model, the willingness of individuals to respond to an alarm is considered through a decision tree. Following that approach, decision trees have been used by others for the identification of decision rules that provide an optimal trade-off between POD and PFA (Einstein and Sousa, 2006; Rheinberger, 2013). In these two subsequent studies, the effect of false alarm on the compliance is not explicitly addressed, but the reliability is expressed in terms of POD and the PFA. This ability of the EWS to distinguish between hazard events and noise can be summarized graphically in receiver operator characteristic curves. This is the inherent reliability of an EWS and will be presented in Sec. 3.

As an alternative to decision trees, influence diagrams are applied to probabilistically model decision procedures associated with EWS (Einstein and Sousa, 2006; Martina et al., 2006). Influence diagrams are based on Bayesian networks (BN), which are graphical models that consist of nodes representing random variables and arcs describing the statistical dependencies among them (Jensen and Nielsen, 2007). They have been successfully applied in the field of environmental modeling and civil engineering due to their intuitive nature, their ability to deal with uncertainty and performing Bayesian analysis, and because of their strengths in representing dependence in large scale systems (Straub, 2005; Straub and Der Kiureghian, 2010). Causal relations between components are defined through conditional probability tables, describing the probability distributions of the variables conditional on their parent nodes. Influence diagrams extend BNs for decision analysis by including decision nodes and utilities (Shachter, 1986).

In Sturny and Bründl (2014), a BN has been constructed to model the technical reliability of a glacier lake EWS. In their study, it was possible to model the entire technical system with a BN, which was not possible with a fault tree in a previous study on the reliability of Swiss avalanche forecasting system (Bründl and Heil, 2011). The first BN, which models both the technical and the inherent reliability of a EWS, is described for a debris flow EWS in (Sättele et al., 2015a). In a subsequent case study, the reliability of a partly automated rockslide warning system is assessed (Sättele et al., 2015b). The automated part is again modelled in a BN and human decision-procedures of the non-automated part are assessed through a Monte Carlo analysis.

In the present contribution, a comprehensive framework approach for the evaluation of EWS is presented, with three main objectives. The first objective, addressed in section 2, is the development of a classification for EWS, which serves as a basis for a structured evaluation of EWS. The second objective is the development of evaluation methods for the technical and the inherent reliability of EWS. The third and final objective is the development of an overall framework for assessing the effectiveness of EWS. The individual steps of the framework approach are presented in section 3, illustrated by the insights gained in the case studies. The paper concludes with a discussion of the applicability of the framework, its limitations and future work (section 4).

2 Generic classification for EWS

EWS can be defined as "sets of capacities needed to generate and disseminate timely and meaningful warning information to enable individuals, communities and organizations threatened by a hazard to prepare and to act appropriately and in sufficient time to reduce the possibility of harm or loss" (UNISDR, 2007). EWS currently operated in practice have widely varying designs, because they are preliminary developed as prototypes to fit specific needs. They are ambiguously referred to as alarm, alert, detection, early warning, forecasting, monitoring and warning systems. To facilitate a structured evaluation of EWS, a recognized classification should be established.

A classification for landslide EWS is proposed by (Bell et al., 2010), in which monitoring systems, alarm and expert systems are distinguished. We adapt this proposal by classifying EWS in function of their degree of automation into: alarm, warning and forecasting systems (Sättele et al., 2012). In Fig. 2, each system class is depicted with the three main units for monitoring, data interpretation and dissemination. To indicate the degree of automation, components, which are operated automatically are highlighted in grey.

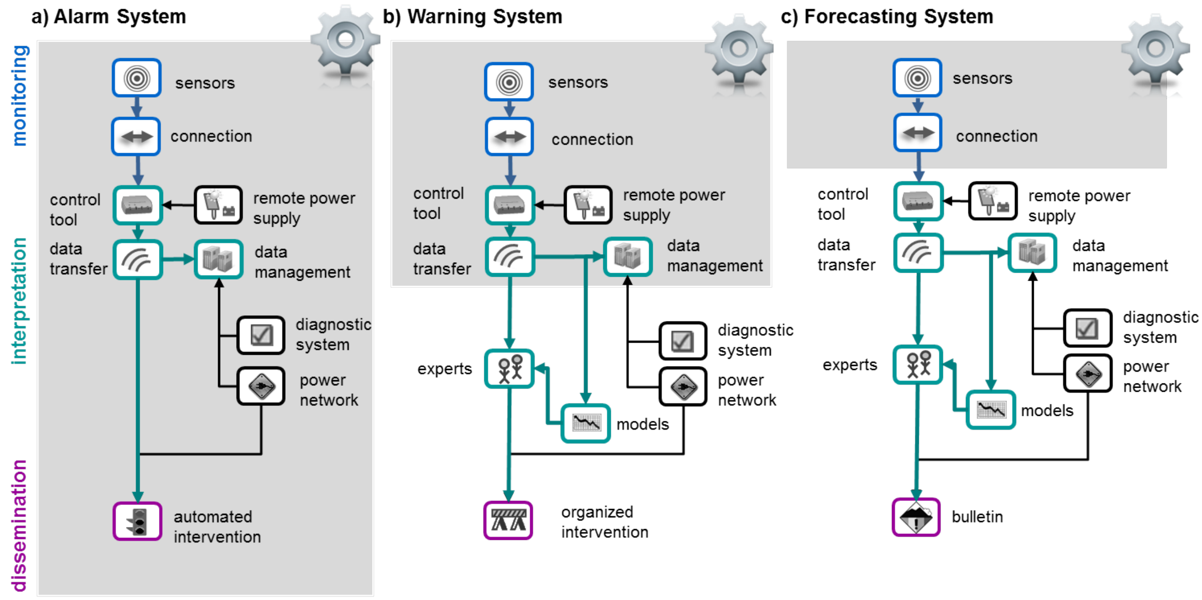

Figure 2: Classification for EWS: Each EWS class includes typical system components facilitating the monitoring, interpretation of data and dissemination of warnings. Automated system parts are highlighted in grey.

In this classification, monitoring systems are not considered as a stand-alone class, because they do not actively issue warning information (Schmidt, 2002; Glantz, 2003). They are a central unit of every EWS, in which the environment is observed and relevant data are collected to increase the processes understanding. As proposed by Bell (2010), alarm systems are understood as threshold-based fully automated EWS. The term “expert system” is omitted because it is already used in the field of artificial intelligence to signify computer systems that imitate the decision ability of humans (Jackson, 1990). Instead, the terms warning and forecasting system are used to distinguish to two types of partly automated EWS. All three classes are named according to how they disseminate information. While alarms are signals activated to inform endangered persons on on-going dangerous events, warnings provide information on imminent or probable events by including suggestions or orders on protective risk mitigation actions (Villagrán de León, 2013). Forecasts deliver more general information on the probability of hazard events in endangered or affected regions for certain time frames in the future (Hamilton, 1997).

The applicability of this novel classification was tested by assigning state-of-the-art EWS to the three classes (Sättele, 2015), including EWS installed worldwide for meteorological, flood, earthquake, tsunami, wildfire, volcanic eruptions and mountain hazards. The results are summarized in Fig. 3, where natural hazards are arranged according to the amount and expressiveness of available precursors and according to the lead time that typical EWS can provide.

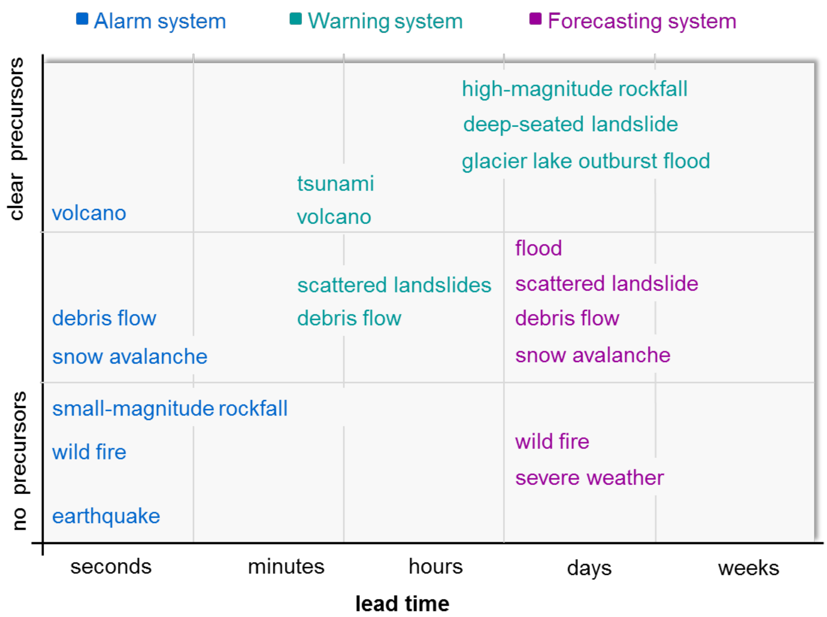

Figure 3: Assignment of natural hazard processes to the proposed classification for EWS: the system class depends on the availability and expressiveness of precursors and the available lead time.

In the following, general characteristics of each EWS class are introduced (see Table 1) and illustrated through a system example. These example systems have been investigated in detailed case studies previously (Sättele et al., 2015a; Sättele et al., 2015b) and key results of these case studies are used in Section 3 to demonstrate individual steps of the proposed framework approach.

Table 1: Characteristics associated with EWS classes.

Alarm system	Warning system	Forecasting system
Fully automated	Partly automated	Lowest degree of automation
Detect on-going process parameters	Monitor precursors	Monitor precursors
Short lead times	Extended lead times	Extended lead times
Thresholds serve as decision instance	First decision is based on threshold, the final one is made by experts	Experts conduct analysis in regular intervals and not based on thresholds
Automated intervention measures such as automated barriers on roads or interrupted power lines at railways	Organized intervention actions such as an evacuation	Forecast the danger level for predefined warning regions to enable preventive actions and preparation

2.1 Alarm system

Alarm systems are fully automated EWS (Table 1; Fig. 2a). In the monitoring unit, sensors are installed to detect process parameters of already ongoing hazard events. They are primarily installed for processes triggered rather spontaneously, such as earthquakes, wildfires, tornados, small rockfalls, debris flow or scattered landslides (Sättele, 2015). Thus, the remaining lead time is short and procedures include a minimal number of interfaces to ensure a reliable and fast information flow. Sensors are directly connected to a control tool, e.g. a data logger, in the interpretation unit. Here, data are analysed to issue and transfer automated warnings or to initiate mitigation actions when predefined thresholds are exceeded. Measured sensor data are transferred and stored in a central data management unit, which is commonly equipped with a diagnostics system. In the dissemination unit, automated intervention measures use optical signals or sirens to generate warnings. In some cases, power cut-offs are initiated to stop approaching trains. At the same time, risk-managers and system operators receive information.

Example: A fully automated alarm system is operated to protect persons from debris flows within the Illgraben catchment in Switzerland (Badoux et al., 2009). One single geophone in the upper catchment and two geophones and two radar devices some hundred meters below should detect ongoing events in real-time (Fig. 4). They measure the ground vibrations and the flow depth in the river bed. The upper geophone is controlled by one logger and another logger controls the remaining four sensors. An automated alarm is initiated if predefined thresholds are exceeded. The alarm information is transmitted via modem and communication devices to activate audible signals and red lights at three alarm stations. In parallel information is sent to system operators. The lead time of the alarm system is between 5 and 15 minutes.

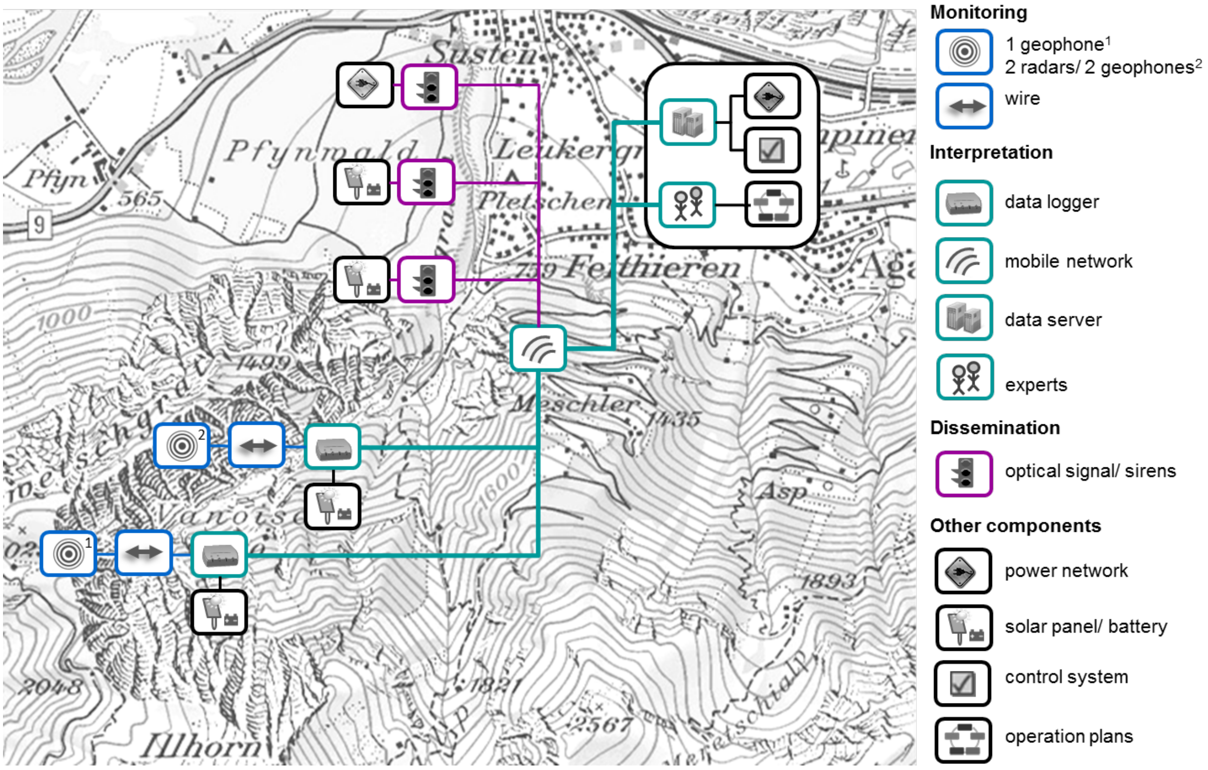

Figure 4: System sketch of the debris flow alarm system in the Illgraben catchment including automated procedures in the monitoring, interpretation and dissemination unit. [Figure based on pixmaps 2015 swisstopo (5704 000 000).]

2.2 Warning system

Warning systems are partly automated EWS (Table 1; Fig. 2b). In the monitoring unit, sensors or human observers monitor precursors of hazardous processes. Precursors are either events that trigger the hazard, such as intense rainfall, or relevant changes in the disposition that occur prior to the event. Therefore, warning systems are typically installed for natural hazard processes that evolve over time and provide precursors, such as tsunamis announced by earthquakes, volcanic eruption or large scale rockfalls (Sättele, 2015). Lead times are extended and enable a two-instance decision-making procedure in the interpretation unit. The first instance is automated: sensor data is transferred to a control tool that typically uses predefined thresholds to initiate automated warnings, similar to alarm systems. The warning is not directly issued to endangered persons but to experts, which are the second decision instance. Experts analyse measured sensor data, and to predict the final event they often apply models or consults additional information sources, such as remote sensing data or reports from local observers. In the dissemination unit, organized intervention actions, such as evacuations and/or closures of roads and railway sections, are set up to mitigate the risk.

Example: In Preonzo, Switzerland, a warning system was installed to predict a mid-magnitude rockslide (Willenberg et al., 2009; Loew et al., 2012), which eventually occurred on May 15, 2012, with about 300'000m^3 rock mass (Fig. 5). Five extensometers and a total station with 14 reflectors monitored increased displacement rates. In the automated part, warning information was sent when predefined thresholds were exceeded. In the non-automated part, displacement data was analysed by experts and the inverse velocity model was applied to predict the event timing, on the basis of which it was decided on further activities. Evacuations were ordered to protect the underlying factories and road. The available lead time is in the order of days.

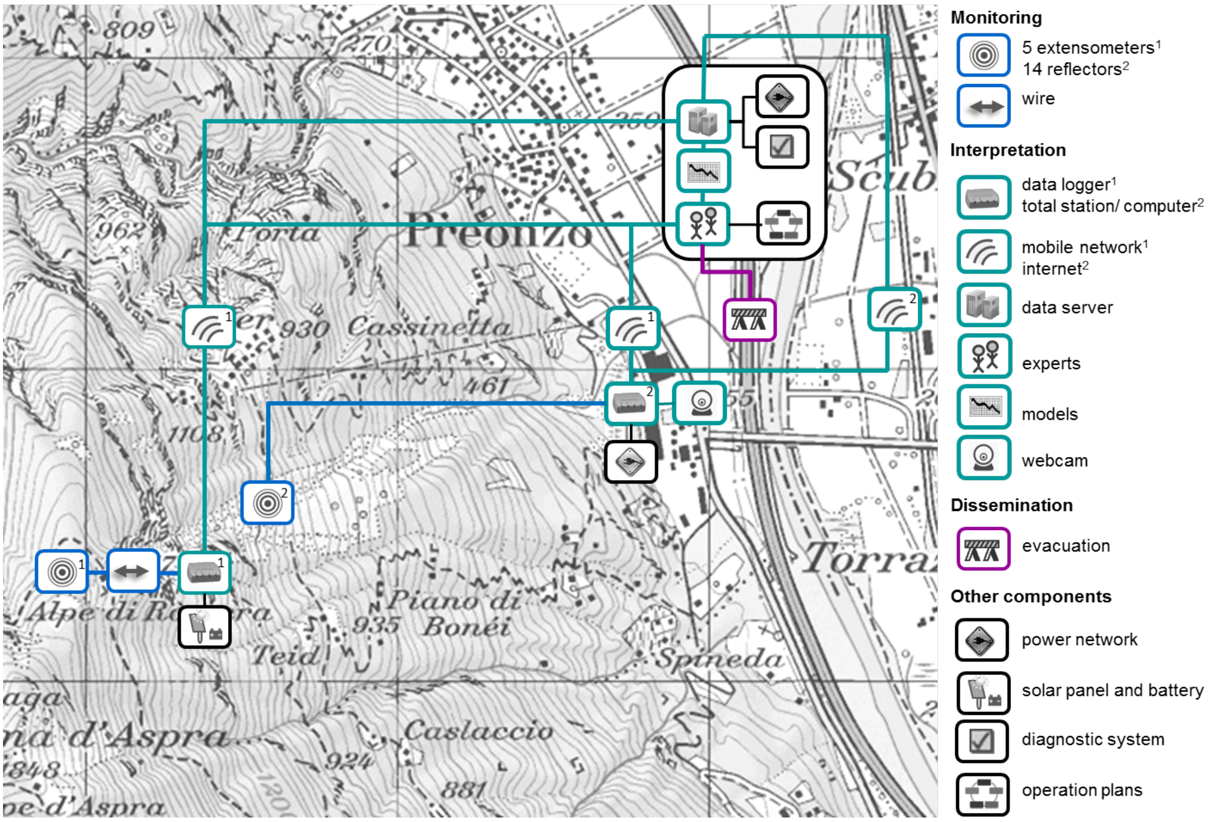

Figure 5: System sketch of the rockslide warning system in Preonzo including partly automated procedures in the monitoring, interpretation and dissemination unit. [Figure based on pixmaps 2015 swisstopo (5704 000 000).]

2.3 Forecasting system

Forecasting systems have the lowest degree of automation (Table 1; Fig. 2c). In the monitoring unit, sensors or human observers monitor precursors to indicate the likelihood of dangerous events. They are chiefly operated to extend the short lead time achieved with alarm systems for spontaneous processes, such as severe weather, wildfires or snow avalanches, but can also be found for processes that are more predicable such as rain induced flood events (Sättele, 2015). In contrast to warning systems, the data interpretation is not initiated when predefined thresholds are exceeded, but conducted at regular intervals. Measured sensor data are transferred to a central data management unit, where experts analyse data and apply models to forecast the danger level for predefined warning regions. If predefined danger levels are exceeded, information is disseminated to public and/or risk managers via media such as mobile phones, Internet, radio and TV. Based on this information and local assessments, risk managers typically initiate a chain of preventive measures by following operation and intervention plans.

Example: The Swiss avalanche system operated by the WSL Institute for Snow and Avalanche Research SLF is an example of a forecasting system (Fig. 6). A network of about 160 snow and weather stations monitors precursors, such as snow height, air and snow temperature and humidity, solar radiation, wind direction and wind speed at regular intervals. Observers transfer measurements and observations to the national centre (Techel and Darms, 2014). Data analysis is conducted by experts on a regular basis. They merge and analyse measured data and data collected by human observers; moreover they apply models and consult meteorological models to predict the danger level for the next day. The forecasts are disseminated in the form of a bulletin, in which warning regions are assigned to five danger levels defined in the uniform European Avalanche Hazard Scale (Meister, 1995). The bulletin is published via radio, TV and Internet, and if danger level four is exceeded, warnings are actively communicated to cantonal authorities and to the public by the National Emergency Operations Centre (Hess and Schmidt, 2012). Based on this information and local assessments, local avalanche safety officers take measures, such as road closures or controlled avalanche release.

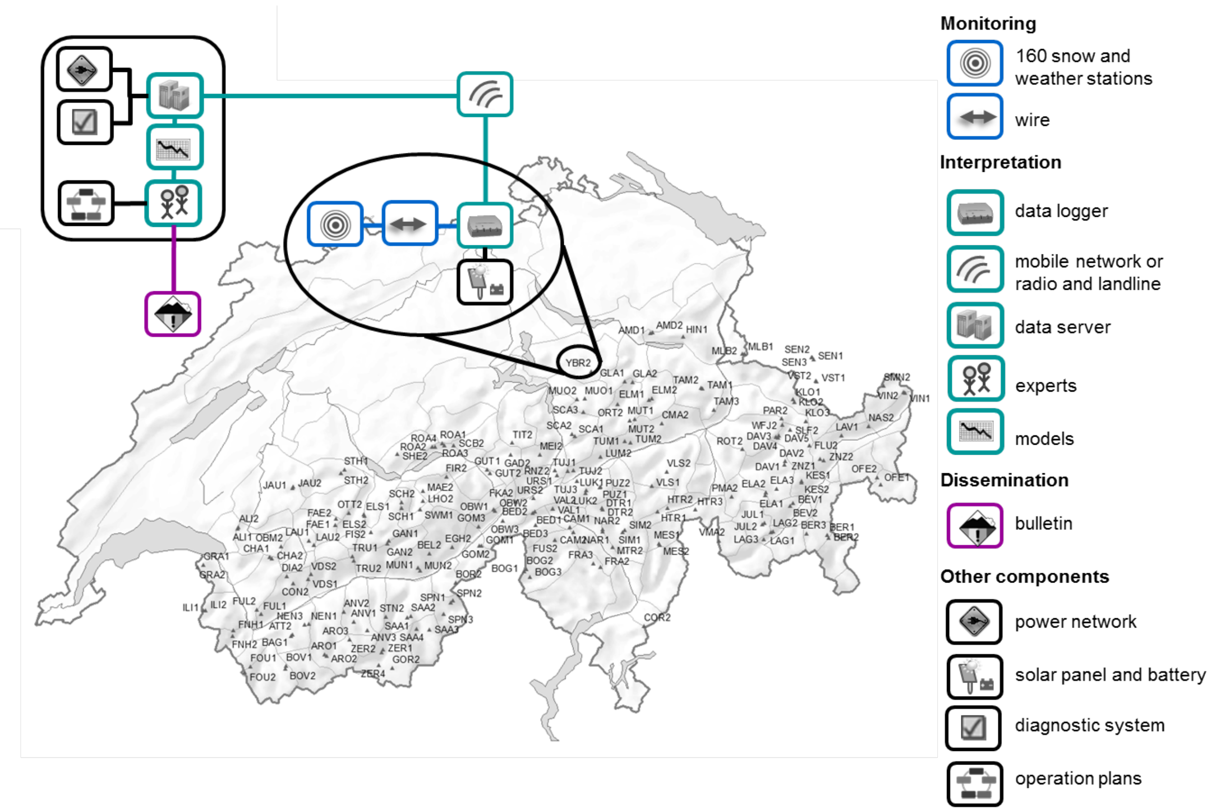

Figure 6: System sketch of the national avalanche forecasting system in Switzerland including mainly non-automated procedures in the monitoring, interpretation and dissemination unit. [Figure based on pixmaps 2015 swisstopo (5704 000 000).]

3 Framework for the evaluation of EWS

Based on the classification, we suggest a framework for a structured evaluation of EWS effectiveness, consisting of three parts as illustrated in Fig. 7. For fully automated alarm systems, parts I and III are sufficient, for partly automated warning and forecasting systems all three parts should be executed.

In parts I and II, reliability analyses are conducted, including the technical and the inherent reliability. The technical reliability analysis accounts for the availability of technical system components and their interdependencies in the system. The inherent reliability analysis differs for parts I and II. While the inherent reliability of automated EWS (part I) depends on automated decision instances such as signal thresholds, non-automated EWS (part II) rely primarily on human decision-making and the accuracy of models. In some cases, the model accuracy needs to be considered in part I as well, e.g. when earthquake alarm systems use models to detect events in real time. In both parts, the inherent reliability is expressed in terms of POD and PFA, as is the overall reliability.

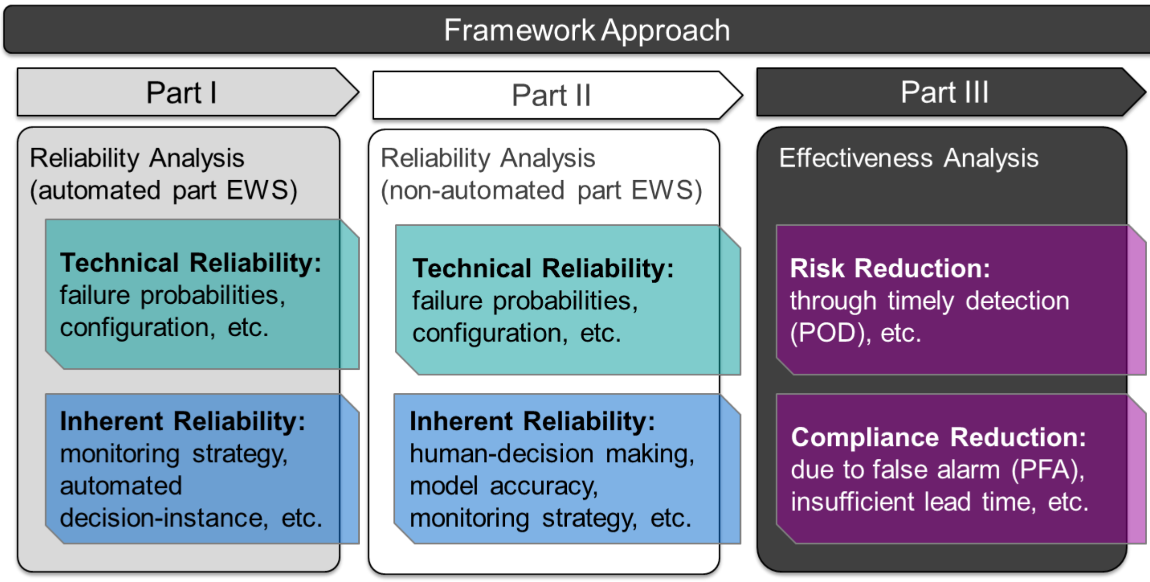

Figure 7: Framework approach comprises three major parts that can be selected dependent on the EWS class to quantify the effectiveness as a function of the reliability.

In part III, the EWS effectiveness is quantified as function of POD and PFA. The effectiveness is a direct function of POD, because timely detection leads to intervention measures that reduce consequences. A high number of false alarms may not only cause large costs for unnecessary interventions, but also decrease the probability that persons comply (POC). The POC is estimated from a basic compliance rate, combined with reduction factors to account for the effect of false alarms (PFA), insufficient lead time and possibly other effects related to the communication and perception of the alarm/warning.

In the following, the three parts of the framework are summarized and individual steps are demonstrated with results of the two case studies Illgraben and Preonzo (Sättele et al., 2015a; Sättele et al., 2015b).

3.1 Part I: Reliability analysis of automated EWS

In part I, the reliability achieved with fully automated alarm systems and the automated part of warning and forecasting systems is assessed in six steps (Fig. 8). Both the technical and inherent reliability are modelled together in a BN, which results in the POD and PFA of the automated system.

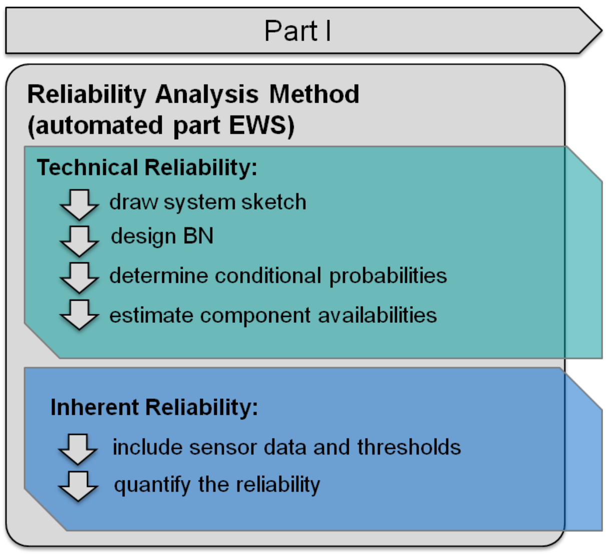

Figure 8: Part I includes six steps to model the technical and inherent reliability of automated EWS.

1^{st} draw system sketch: A system sketch is an essential basis to understand the EWS design and the dependencies among the components (see Fig. 4-6). It can be constructed according to the three main units of an EWS and contains all main system components. The information flow is indicated by arcs and components are represented in form of squares or nodes. Redundant system parts are depicted redundantly in the sketch.

2^{nd} design BN: The basic BN can be derived from the system sketch. It consists of nodes and arcs, which can be structured according to the same three units (see Fig. 9). Oval nodes represent system components, and they are arranged according to the causal chain from the *hazard event* to the *warning*. This includes the main functionalities such as *data measured*, *event indicated*, *warning issued*, *transmitted* and *released*. Redundant system components and functionalities are also depicted redundantly in the BN. The arcs in the BN are directed to follow the information flow between functionalities and components. Decision nodes (squared nodes) are added in the BN to specify decision criteria on varying levels (see 5^{th} step).

3^{rd} determine conditional probabilities: Interrelations between the components and functionalities in the causal chain can be specified in conditional probability tables of oval nodes. In many instances, AND or OR relations are sufficient to describe the dependencies of individual components and functionalities, but any other type of logical or probabilistic relation can also be specified. AND relations represent serial connections, in which all components must work to ensure the underlying functionality; OR-relations can be used to model redundant configurations.

4th estimate component availabilities: The availability of individual components is specified in the conditional probability tables of oval nodes representing components. If the component can assume exactly two states (functioning or fail), the random variable is binary. If additional states are possible, these are specified in the conditional probability tables. Availabilities can often be derived from failure rates specified by the supplier, to which one should add the rate of failures caused by external sources, such as extreme temperatures or disturbances due to human and animal activity.

5th include sensor data and decision instances: Decision instances, such as warning thresholds, are added as squared decision nodes on various levels, either for single sensors or to specify warning criteria to combine information from several sensors. Probabilities of measured sensor data to exceed these criteria are included in the conditional probability tables of the nodes representing sensor signals. These probabilities are estimated conditional on the occurrence of an event. This 5th step is not necessary for forecasting systems, which do not use automated decision instances.

6th quantify the reliability: The last node of the causal chain (*warning*) is used to assess the overall reliability of the EWS. POD and PFA are obtained by changing the status of the top node (*hazard event*) and evaluating the BN. If the top node is set to “event”, the probability of the last node being in state “alarm” is equal to the overall system POD. Similarly, the PFA is obtained by setting the top node to “no event”. The same BN facilitates that the technical and the inherent reliability are assessed together or separately. To model the technical reliability alone, the status of the node “event indicated” is set to “yes”; to assess the inherent reliability the status of all nodes representing technical system components is set to the state “functioning”.

3.1.1 Illustrative examples from the Illgraben and Preonzo case studies

The reliability of the fully automated Illgraben alarm system and the automated part of the Preonzo warning system is quantified following the six steps of part I (Fig. 8).

***1st draw system sketch**:* For the Illgraben and the Preonzo case study, system sketches are designed following the three main units for monitoring, data interpretation and information dissemination, as shown in Fig. 4 and 5. The sketch includes only main components to keep the following steps manageable. For example, the data logger is considered together with the included software.

2nd design BN: The BNs constructed for the Illgraben and Preonzo EWS vary strongly. For the fully automated Illgraben debris flow alarm system, a comprehensive reliability analysis for the entire warning chain from the *hazard event* to *warning* is conducted as illustrated in Fig. 9. The inherent and the technical reliability are evaluated together and are expressed in terms of POD and the PFA. Grey nodes represent the causal chain, white nodes the components and thresholds are defined through the black decision-nodes.

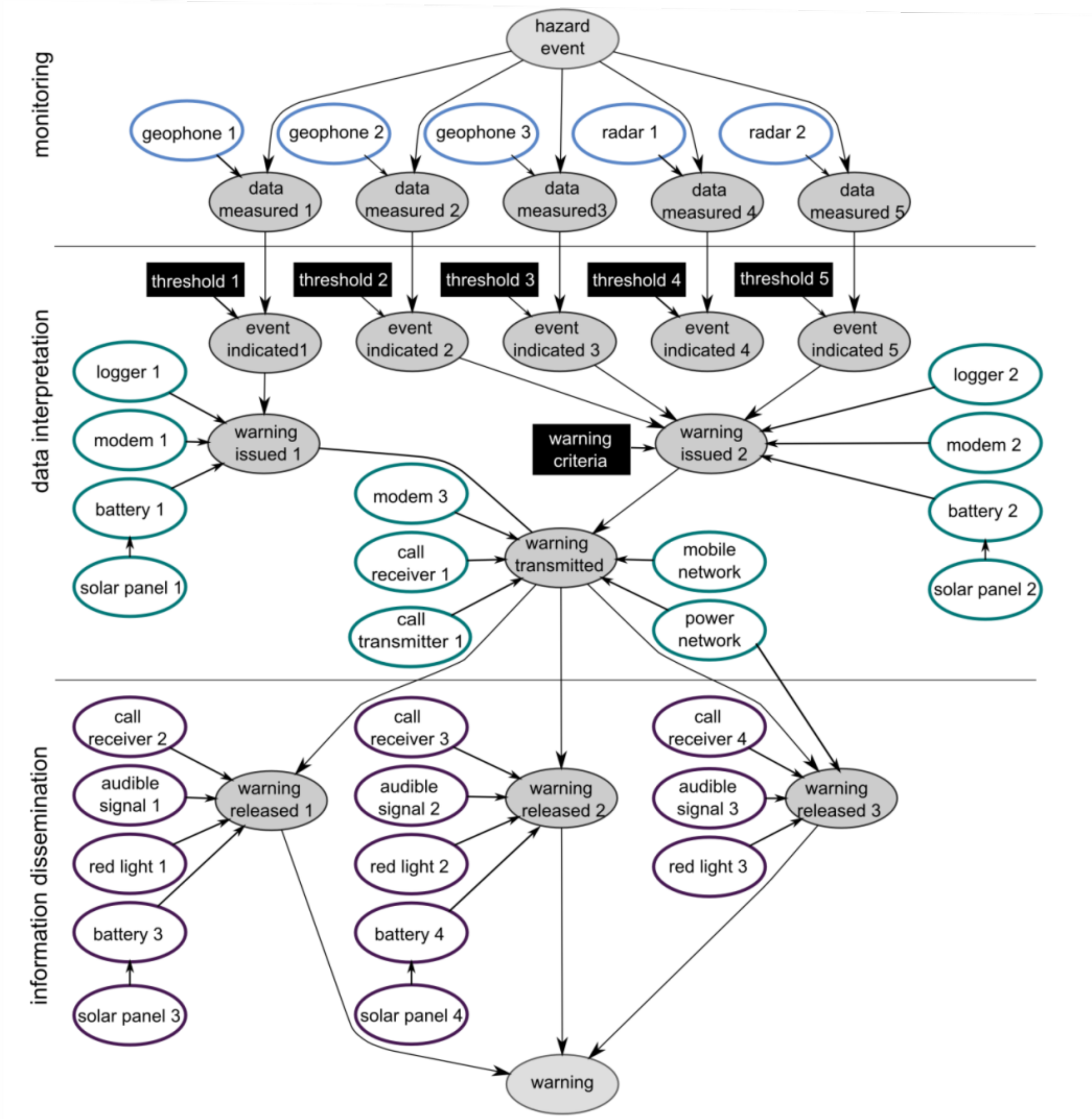

Figure 9: The BN to model the overall reliability of the Illgraben alarm system is structured according to three main units. Grey nodes represent main functionalities in the causal chain; white nodes represent components and squared black nodes the decision-instances on two levels, for details see (Sättele et al., 2015a).

For Preonzo a simplified BN is constructed to model the ability of the system to provide timely warning information to decision-makers (Fig. 10). Here, the technical reliability alone is modelled, and sensor data and decision nodes are not included, so that the PFA cannot be computed here. This simplification is possible because warnings are sent directly to experts whose compliance should not be reduced by frequent warning information.

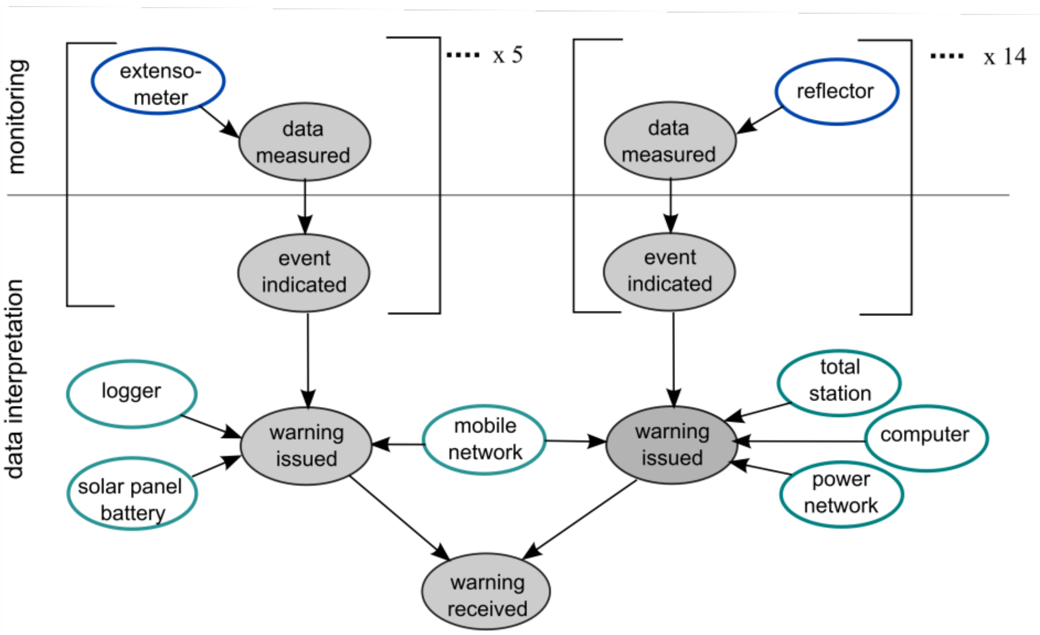

Figure 10: The BN to model the technical reliability achieved in the automated part of the Preonzo warning system. The redundant monitoring unit includes 5 extensometers and 14 reflectors. In the data interpretation unit, warning information is issued automatically to decision-makers. For details see (Sättele et al., 2015b).

3rd determine conditional probabilities: In both BNs, the interrelations among system elements are specified either deterministically or stochastically in the conditional probability tables of grey nodes. In the causal chain of the Illgraben BN, warning information is transmitted if either sensor unit 1 or 2 issues an event (Table 2a), but the warning in sensor unit 2 is only issued if a at least one of the geophones and one radar device indicates an event (Table 2b).

Table 2: The causal relations between functionalities and components are specified in the conditional probability tables of grey nodes. Here, two examples of deterministic nodes are shown. a) OR logic of the redundant sensor units; b) AND logic of sensors in monitoring unit 2.

a)

sensor unit 1 indicates event		yes		no	
sensor unit 2 indicates event		yes	no	yes	no
warning transmitted	yes	1	1	1	0
	no	0	0	0	1

b)

event indicated 1 (geophone 1)		yes								no							
event indicated 2 (geophone 2)		yes				no				yes				no			
event indicated 3 (radar 1)		yes		no		yes		no		yes		no		yes		no	
event indicated 4 (radar 2)		yes	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
warning issued 2	yes	1	1	1	0	1	1	1	0	1	1	1	0	0	0	0	0
	no	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	1

4^th^ estimate component availabilities: In both case studies, availabilities of components are specified in the CPTs of white nodes. All components can assume exactly two states; functioning and failed. For the Illgraben case study, availabilities A of system components are calculated following Eq. 4 and are in the order of 0.9995 for most components (Sättele et al., 2015a).

$$A \approx 1 - (\lambda_{IF} + \lambda_{EF}) \times \mathrm{E}[T_r] \qquad (4)$$

λ_{IF} are internal failure rates and λ_{EF} are external failure rates; $\mathrm{E}[T_r]$ is the expected time it takes to detect and repair a failure. Internal failures rates λ_{IF} are derived from the specified mean time to failure (MTTF) and the mean time between failure (MTBF) values and external failure rates λ_{EF} are estimated by experts.

5^th^ include sensor data and decision instances: In the Illgraben case study, past event data from 44 events are used to determine probabilities of thresholds being exceeded on both event and non-event days (see Table 1 in (Sättele et al., 2015a)). The BN constructed for the warning system in Preonzo is developed to facilitate the assessment of the technical reliability alone and does not include thresholds or measured sensor signals (details see 2^nd^ step).

***6^th^ quantify the reliability**:* In the Illgraben case study, the inherent reliability for varying thresholds is modelled for each sensor separately (see Fig. 11). Besides the threshold, the positioning of the sensors has a major influence on the EWS reliability, whereas technical failures of individual components have a comparatively low impact due to high redundancies (Sättele et al., 2015a).

For Preonzo we find that the technical reliability, i.e. the POD of the automated part, is high (0.988) due to multiple redundancies in the sensor unit and a diagnostic system that immediately detects and reports component failures to minimize downtimes of the system. The inherent reliability is close to one, but is not assessed quantitatively with the BN. This is not necessary because the warning threshold were set low to ensure that the EWS sends timely information to the expert team responsible for the final decision on an evacuation. The system is furthermore designed as fail-safe, i.e. in case of a technical failure, the experts are alerted.

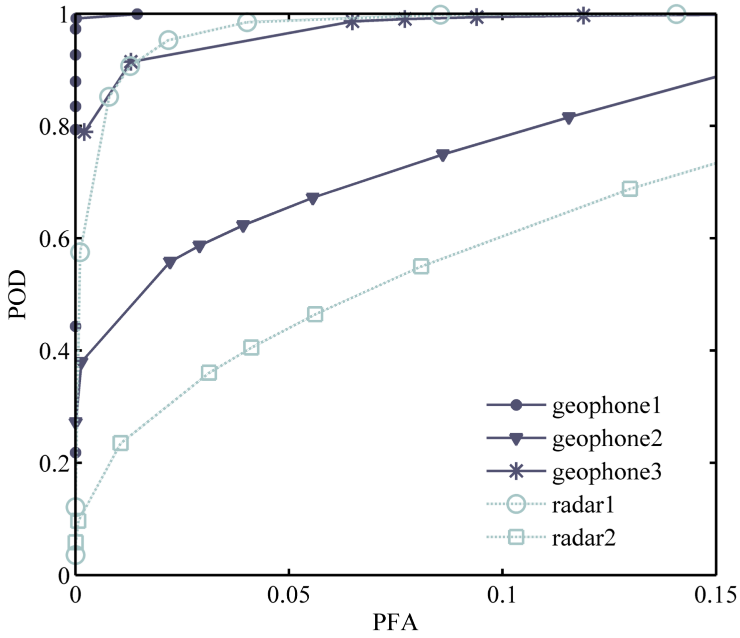

Figure 11: Reliabilities of individual sensors in the Illgraben alarm system vary strongly and can be graphically summarized as receiver operator characteristic curves, in which the dependence between POD and PFA is shown (Sättele et al., 2015a).

3.2 Reliability analysis II: non-automated EWS

In part II, reliability analyses of non-automated parts of warning and forecasting systems are conducted. Here, the ability of the decision-makers to correctly predict or forecast events is evaluated. This ability depends on (potentially complex) human and model-based decision procedures, which are difficult to quantify in practical applications. If the reliability cannot be expressed quantitatively in terms of POD and PFA, a qualitative or semi-quantitative analysis should be conducted instead. This evaluation should address both the technical and the inherent reliability and can be conducted in five steps (Fig. 12).

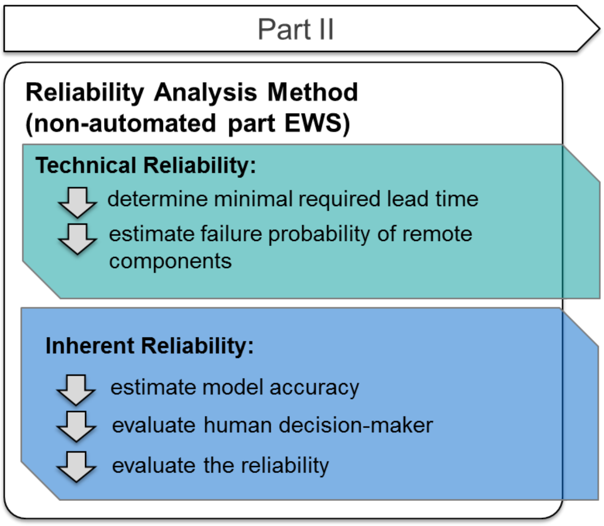

Figure 12: Part II includes five steps to model the reliability of non-automated EWS.

1st determine minimal required lead time: Lead times associated with the non-automated part of warning and forecasting systems are typically larger than those of alarm systems, often in the range of one to several days (see Section 2.2). During this time period, additional data and information is collected and predictions become increasingly accurate (see e.g. Grasso et al., 2007; Schröter et al., 2008). The reliability analysis in part II is therefore conducted as a function of the lead time. The reliability can either be evaluated for a fixed lead time or for a set of lead times. For a given lead time, one should consider the reliability associated with that lead time, as well as the related intervention costs, e.g. those caused by an early evacuation.

2nd estimate failure probabilities of remote components: Non-automated EWS measure precursors and thus provide extended lead times. Nevertheless, their reliability increases with shorter lead times. For some EWS, destructive pre-events can lead to an increased failure probability of system components, e.g. sensors, as the event approaches. A typical example is provided by the Preonzo case study, summarized in Sec. 3.2.1. The technical failure probability associated with the minimum required lead time is the input for determining the remaining number of sensors, which will in turn affect the forecast accuracy that is evaluated in the next step.

3rd estimate model accuracy: Experts often apply models to predict the event magnitude, time and spatial dimensions. Flood forecast are for example based on coupled hydro-meteorological models, which become probabilistically when Hydrological Ensemble Prediction Systems are used (Wetterhall et al., 2013). The accuracy of models depends on their capabilities, their case-specific applicability and on the quality of the available input data. The quality of the data is determined by the number, the type and the positioning of sensors. The model accuracy is evaluated for the selected minimal lead time and expressed qualitatively or semi-quantitatively (see 5th step). The estimated model accuracy directly influences the ability of decision-makers to set up intervention measures correctly. If no models are applied, this step can be skipped.

4th evaluate human decision-makers: In the non-automated part of EWS, the final decision is made by humans. The involved decision procedures are typically complex and can only in some cases be assessed quantitatively (see 3.2.1). In most cases, a qualitative or semi-quantitative analysis is more suitable, in which possible outcomes, the degree of risk aversion and the expertise of individuals and effects associated with group dynamics are addressed. Decision-makers are evaluated according to their ability to correctly detect dangerous events (POD) and avoid false alarms (PFA). Both terms can be rated in predefined evaluation scales e.g. as low, medium or high.

5th evaluate the reliability: The reliability achieved in the non-automated part of the EWS is evaluated as a function of the lead time. It depends on the procedures to initiate and carry out intervention measures following a warning, such as evacuation. The decision on a warning is influenced by the accuracy of the applied forecasting models and the quality of available information from different sources, such as measured sensor data, data from other sources and reports from human observers. The quality of the input information directly influences the forecast ability of models and the success of human decision-making. Whether damage is successfully prevented depends also on the quality and the feasibility of predefined intervention plans. In a comprehensive reliability analysis, all these factors and their dependencies are considered. In most cases, this analysis will be qualitative. However, the final reliability should be expressed (semi-)quantitatively in terms of POD and PFA. To this end, values for POD and PFA may be assigned to qualitative rating scales (e.g. low (POD=0.90 and PFA=0.1), medium (medium POD=0.95 and PFA=0.05) and high (POD=0.99 and PFA=0.01).

3.2.1 Illustrative example from the Preonzo case study

In a detailed case study, the reliability of the non-automated part of the Preonzo warning system is assessed. To enable a quantitative reliability evaluation, a post-event analysis of a large event (about 300’000m^3) that occurred on May 15, 2012 is conducted, following the five steps of part II.

1st determine minimal required lead time: If decision-makers release the information one day in advance, the evacuation can be carried out successfully and sufficient time for intervention teams to set up protective measures is available. The quality of the prediction is also maximum for short lead times, and the intervention costs, which occur due to business interruptions in the underlying factory buildings, can be kept relatively low. Hence, one day is selected as the lead time.

2nd estimate failure probabilities of remote components: Sensors fail before the event in May 2012, and shortly before the instable mass collapses, a majority of sensors are destroyed. To account for the increasing failure rate, a function is fitted to the number of observed failures (Fig. 13). The estimated failure probability of sensors at the minimal required lead time (t = 1 day) necessary to set up an evacuation successfully is 0.4.

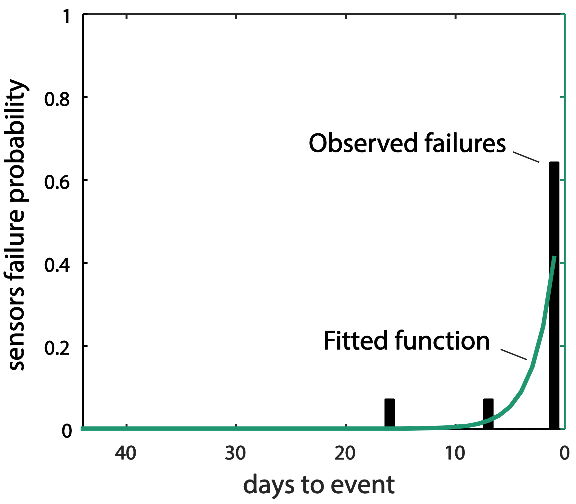

Figure 13: Shortly before the event in May 2012 a large number of sensors is destroyed: the green function is fitted to the observed percentage of destroyed sensors (Sättele et al., 2015b).

3rd estimate model accuracy: To predict the event time, the inverse velocity model is applied on sensor data measured in Preonzo before May 15. In Fig. 14, the predicted event dates modelled between April 1 and May 14 by sensors installed close to the release area are summarized. As the event approaches, the prediction made by individual sensors becomes more uniform. One day before the event occurred, at the minimal lead time, ten out of twelve available sensors predict the event to occur on the next day. However, on May 6, most sensors predict the event for the next day and an unnecessary evacuation is set up on May 7 and annulled a day later when accelerations slow down again.

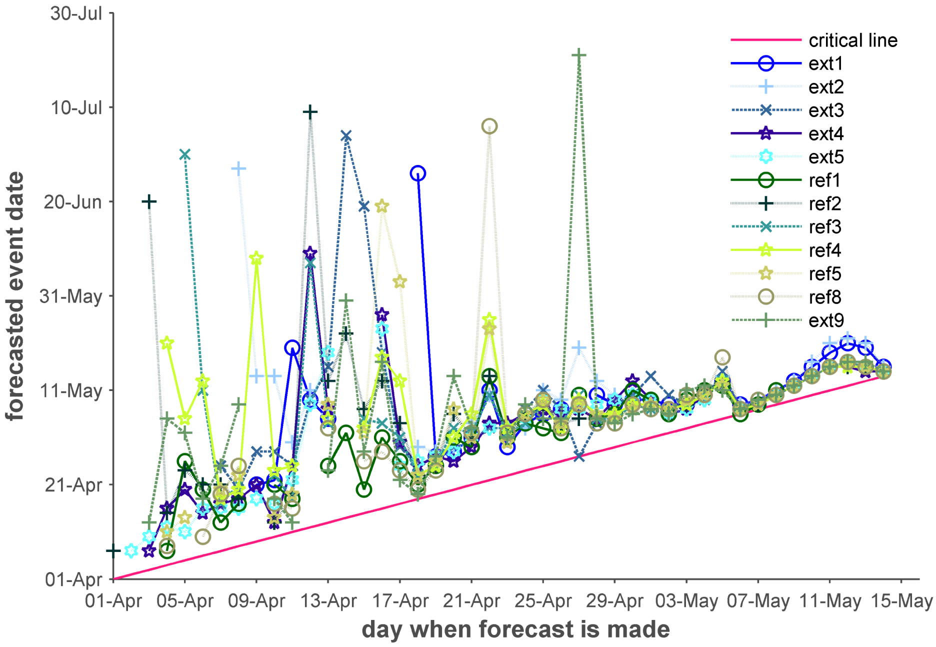

Figure 14: In Preonzo, the model accuracy increases with decreasing lead time. In April, sensor forecasts made with the inverse velocity model vary strongly among different sensors. On May 14, ten out of twelve sensors predict the event correctly for the next day (Sättele et al., 2015b).

4th quantify human decision-makers: In Preonzo, the final decision on setting up intervention measures is made by an expert team. As a first attempt to quantify the decision-making procedure, the experts are characterized by simple decision rules. According to these rules, an evacuation is set up if less than a certain amount of initial sensors remain intact (technical criterion) or if a certain percentage of initial sensors predict the event for the following day (inherent criterion), as summarized in Table 3. The amount of initial sensors is varied in the Preonzo study from 5 to 50.

Table 3: To quantify the human decision-maker, two risk types are specified with different evacuation criteria (Sättele et al., 2015b).

risk type	**technical evacuation criterion, evacuate when:**	**inherent evacuation criterion, evacuate when:**
less risk tolerant	less than 6 sensors are functioning	20% of sensors forecast the event for the next day
more risk tolerant	less than 3 sensors are functioning	50% of the sensors forecast the event for the next day

5th quantify the reliability: The overall reliability achieved in the non-automated part of the Preonzo warning system is assessed probabilistically through a Monte Carlo simulation. The model accuracy and the sensor failures are randomized, to quantify the probability that evacuation measures are set up on the day of the event (POD) (Fig. 15a). In addition, the costs for intervention are calculated, which are decreasing with increasing number of sensors, and which are smaller for the risk-tolerant decision-maker (Fig. 15b). Analyses are conducted for a varying number of initial sensors and two risk types (see Table 3) and confirm that the risk tolerance of human-decision makers have a significant influence on the reliability of non-automated parts of EWS. Figure 15a shows that even with a high number of sensors, the probability of the risk tolerant decision maker to detect the event is never exceeding 0.85.

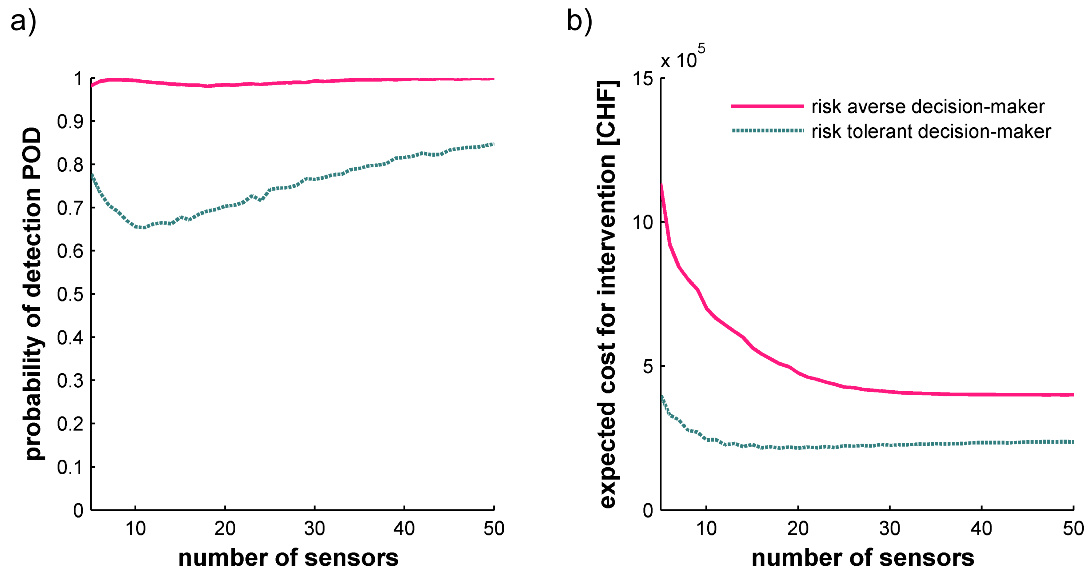

Figure 15: The reliability (POD) and costs for intervention are modeled for two decision makers and varying number of initial sensors: a) the less risk tolerant decision-maker reaches high values of POD independent of the number of sensors; the risk tolerant decision-maker only reaches a POD up to 0.85; b) the more risk tolerant decision-maker creates lower expected costs, which reach a minimum of 215,000 CHF with around 20 sensors or more; for details see (Sättele et al., 2015b).

3.3 Part III: Effectiveness Analysis

The effectiveness of an EWS, E_w, is defined as the relative risk reduction achieved with the EWS and can be quantified following Eq. 1 as a function of the risk without the EWS R and the risk with the EWS $R^{(w)}$. EWS reduce the risk when timely information leads to intervention measures that decrease either the exposure probability pe_{ij} or in some cases the vulnerability in Eq. 3. By combining Eqs. 1-3, the effectiveness of an EWS can be calculated as:

$$E_w = 1 - \frac{\sum_{j=1}^{n_{scen}} \sum_{i=1}^{n_{obj}} p_j \times pe_{ij}^{(w)} \times v_{ij}^{(w)} \times A_i}{\sum_{j=1}^{n_{scen}} \sum_{i=1}^{n_{obj}} p_j \times pe_{ij} \times v_{ij} \times A_i} \qquad (5)$$

To determine $pe_{ij}^{(w)}$ and $v_{ij}^{(w)}$, the POD and PFA estimated in the reliability analyses of part I and II, are used.

The exposure probability $pe_{ij}^{(w)}$ is reduced when persons are successfully evacuated or when intervention measures avoid that persons enter endangered areas. Organized evacuations are often initiated by warning and forecasting systems installed for tsunami, flood, volcanic, large scale slope failures and wild fires. Automated measures for keeping people from the endangered area are activated by alarm systems installed for debris flows, avalanches and small magnitude rockfalls.

The vulnerability $v_{ij}^{(w)}$ is reduced if the EWS sends timely information that leads to temporary measures, which decrease the susceptibility of objects to damage. If storm events are announced timely, movable objects can be fixed; if flood warnings are issued, protective temporary measures such as sandbags or wooden barriers can be installed. Modern earthquake alarm systems can slow down trains or shut down critical processes in factories when strong shaking is detected in time.

The reduction of the exposure probability and the vulnerability is equal to the probability that the event is detected and intervention measures are initiated (POD) and that endangered persons comply with the warning (POC). The latter is not relevant for fully automated intervention measures such as power cut-offs. If EWS issue warnings to persons, a high POC is crucial. It can be quantified as a function of the general compliance rate POC_0 and reduction factors RF, e.g. due to false alarms $RF(PFA)$ or insufficient lead time $RF(ILT)$:

$$POC = POC_0 \times RF(PFA) \times RF(ILT) \qquad (6)$$

The basic compliance rate and the reduction factors must be determined case-specifically. The basic compliance rate depends on type of intervention measures and human decision-making. If, for example, barriers are closed on a road, car drivers have to comply, while red lights can be ignored. Moreover, it can be assumed that regular trainings and education leading to a higher awareness of potential consequences can improve the basic compliance rate.

The reduction factor due to false alarms $RF(PFA)$ accounts for the cry-wolf effect, namely that people have an increased tendency to ignore warnings after experiencing (multiple) false alarms. This effect depends, among other factors, on past experiences, expected consequences and the degree of risk aversion of the recipients.

The reduction factor due to insufficient lead time $RF(ILT)$ express the ability to comply. In certain cases, EWS have to be constructed in a way that the available lead time may not be sufficient and not everybody willing to comply can successfully evacuate. In the case of earthquake alarm systems, lead times are in the range of just a few seconds; or for avalanche alarm systems constructed above railways, the lead time is limited by the distance from the railway to the release point.

3.3.1 Illustrative example from the Illgraben case study

In the Illgraben case study, the effectiveness E_w is calculated as a function of POD and PFA. The alarm system reduces the exposure probability of persons in the Illgraben catchment. Therefore, the effectiveness is equal to the reduced exposure probability with the EWS. To simplify the analysis, different debris flow types are not distinguished, and only one scenario j is considered. The exposure probability is the same for all persons i, $pe_{ij} = pe_j$, and it follows:

$$E_w = 1 - \frac{p_j \times pe_j^{(w)} \times \sum_{i=1}^{n_{pers}} v_{ij} \times A_i}{p_j \times pe_j \times \sum_{i=1}^{n_{pers}} v_{ij} \times A_i} = 1 - \frac{pe_j^{(w)}}{pe_j} \quad (7)$$

The reduced exposure probability is evaluated as a function of the POD and the POC:

$$pe_j^{(w)} = pe_j(1 - POD \times POC) \quad (8)$$

Inserting in Eq. (7), the effectiveness becomes

$$E_w = POD \times POC \quad (9)$$

POD values result from the reliability analysis and POC is calculated as a function of PFA. To this end, we adapt the basic compliance rate $POC_0 = 0.95$ from published traffic analyses (Rosenbloom, 2009; Johnson et al., 2011) and the $RF(PFA)$ is adapted from a existing case study in which the compliance frequency of students as a function of false alarms is assessed (Bliss et al., 1995). As illustrated in Figure 16, the compliance frequency strongly decreases with an increasing ratio of false alarms.

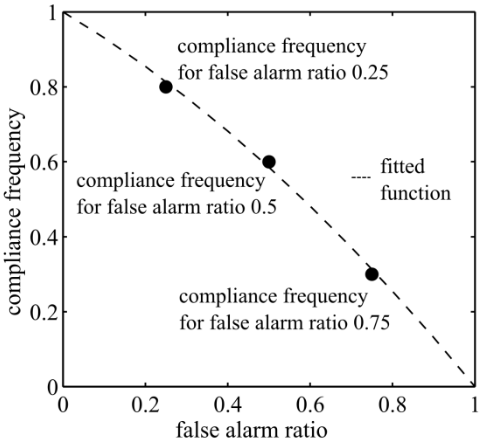

Figure 16: Compliance frequency in function of the false alarm ratio (Sättele et al., 2015a).

In the Illgraben case study we extend the BN to a decision graph and identify the threshold combination that leads to a maximal effectiveness following Eq. (9). In Fig.17, the resulting effectiveness is shown as a function of POD and PFA, together with the POD and PFA values associated with the best system configurations. For this highly reliable EWS, the effectiveness decreases faster with increasing PFA than with increasing POD.

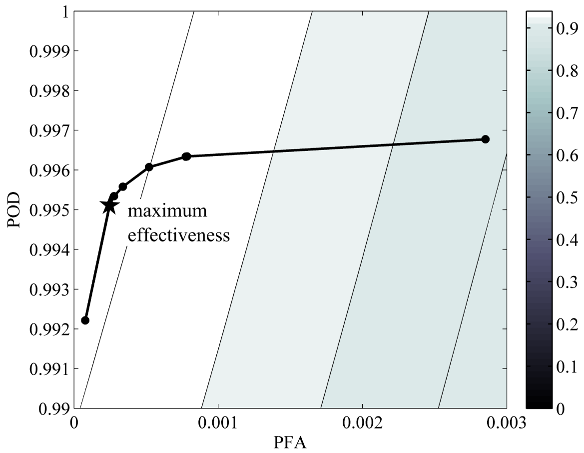

Figure 17: The effectiveness of the Illgraben alarm system could be quantified as a function of POD and PFA; i.e. the reliability (Sättele et al., 2015a).

4 Discussion

The proposed classification for EWS distinguishes alarm, warning and forecasting systems according to their degree of automation, their lead time, and the expressiveness of the available precursors (Figs. 2 and 3). The selection of an EWS class depends strongly on the underlying natural hazard process. Different process types allow for different monitoring strategies, which are associated with different lead times and degrees of automation. Earthquakes, for example, occur without clear precursors and damage can only be reduced by fully automated alarm systems with very short lead times. In contrast, large river floods provide clear precursors and damage can be reduced when warnings or forecasts are made early enough to set up temporary intervention measures.

A differentiation of EWS according to their degree of automation has proven to be a valuable basis for evaluating EWS. The system requirements differ strongly between automated and non-automated EWS and these should be addressed separately. Typical procedures conducted within automated EWS parts are less complex than human- and model-based decision procedures that are part of non-automated EWS. Part I of the proposed framework consists of a six step method for a quantitative reliability assessment of automated EWS; and part II contains five steps for a qualitative or semi-quantitative evaluation of non-automated parts.

Through the two case studies, we demonstrate that this framework approach is applicable to assess alarm and warning systems installed for gravitational processes in mountain regions. With the Preonzo case study, we moreover show that under some conditions the reliability of non-automated EWS can be quantified as well. Here, a post event analysis is conducted, in which human-decision makers are specified through simple decision rules. When specifying less risk tolerant decision rules (Table 3), the analysis leads to similar recommendations than the ones that were actually made by the experts. However, to refine the framework approach for the application on EWS operated for earthquakes, floods, meteorological hazards, tsunamis, volcanic eruptions and wildfires, the following steps of the procedure should be further enhanced.

In part I, the technical and the inherent reliability of automated EWS are quantified in a BN. For the construction of the BN, a system sketch forms the basis for understanding key system components and their interrelations. To keep the complexity of the BN and the proceeding steps low, only essential components should be considered. In step 4, availabilities of individual system components are estimated. Internal failure rates can be derived from specifications of manufacturers, but external failure sources such as extreme temperatures and lightning, which are more difficult to estimate, must be considered as well. However, for many EWS such as the Illgraben case study, the influence of technical reliability is low compared to the inherent reliability, i.e. the ability to interpret data correctly. The assessment of the inherent reliability is challenging in the design phase of EWS or for EWS installed for rare events such as large-magnitude rockfalls. In these cases, sensor data are not yet available to estimate probability distributions of EWS signals. Other EWS, such as earthquake alarm systems, use real-time models to estimate the magnitude on a spatial dimension whenever unexpected ground shakings are detected. Here, measured signals are vector-values and vary in space and time; they need to be further processed in models before a classifier can be applied to distinguish critical events from non-occurrences. In these instances BN must be enhanced; e.g. to model the reliability dependent on the lead time.

In part II, a qualitative or semi-quantitative evaluation is proposed, to assess time dependent human and model based decision procedures. Although a concrete evaluation method, such as the BN of part I, is not provided, the overall procedure for the evaluation of non-automated EWS is presented. The reliability is estimated as a function of the lead time. In step 2, the increase in sensor failure probability before the event must be addressed, as demonstrated in the Preonzo case study. Another example is provided by the 2011 Tohoku earthquake in Japan 2011, where a majority of the offshore sensors failed before the tsunami hit the mainland (Wei et al., 2013). It may be possible that no sensor data are available for an event prediction in the critical phase. The accuracy of predictive models (step 3) depends on the capacity of the model, its applicability and the availability of sensor data. For natural hazards EWS, it is common practice to express the accuracy of models in terms of POD and PFA (see Simmons and Sutter, 2009). As we demonstrate, the framework enables to include the possibility of technical system component failures into POD and PFA, to obtain a single measure of EWS reliability. In some cases, e.g. for flood models, the ability to spatially and temporarily predict the event should be addressed in the reliability analysis (Wheater et al., 2005). In these cases, the reliability is ideally described by the prediction errors of the timely forecasted discharge and not (only) in terms of POD and PFA. In non-automated EWS, the final decision is made by humans, often together with models applied on available sensor data. In most cases, human-decisions are not rule-driven and cannot be quantified easily, but depend on factors such as experience, risk tolerance and the environment in which the decision is made. To account for those factors, a qualitative evaluation is proposed, in which the performance of human decision makers is rated in predefined scales (e.g. low, medium, high) as it is common for the evaluation of structural mitigation measures (Margreth and Romang, 2010). The final reliability should then be evaluated in a semi-quantitative procedure where values for POD and PFA are assigned to different rating scales, e.g. high POD (0.95-1.0), limited POD (0.8-0.95) and low POD (0-0.8).

In part III, the effectiveness is quantified as a function of POD and PFA. The reduction of the exposure probability and vulnerability is a direct function of POD. In some instances, the EWS effectiveness is directly proportional to POD, as demonstrated in the Illgraben case study. The PFA determines the probability that persons comply with the warning (POC). It is also used to estimate the costs caused by unnecessary evacuations. The costs and the effectiveness are main criteria for the identification of optimal risk mitigation measures for natural hazards.

The overall user-friendliness of the novel framework can be improved if a convenient software tool is provided. Such a software tool can be developed following the three steps defined by the framework approach. The reliability evaluation for automated system parts can be done by running a BN in the background. The user interface should be designed user-friendly, including simple input fields in which e.g. system components, their technical failure probabilities and dependencies can be specified in order to optimize a system. Finally, it could be embedded in a software environment in which risk reduction of an EWS can be compared to alternative measures to support decision makers in the identification of optimal mitigation measures.

5 Conclusion

With the proposed framework approach, the effectiveness of EWS is evaluated as a function of the reliability through three main parts. To enable a structured evaluation of EWS, a generic classification is provided, differentiating EWS into alarm, warning and forecasting systems according to their degree of automation, lead time and the availability of clear precursors. In function of the EWS class, different parts of the framework are selected. Each part is structured along predefined steps, which are here illustrated with the result of two case studies. The reliability assessment of the automated part of EWS is performed quantitatively through a Bayesian network. To evaluate non-automated EWS parts, which involve the decision making of experts, a qualitative or semi-quantitative approach is generally preferable. However, as exemplified in the Preonzo case study, a quantitative assessment can be possible and provide insights.

The framework should be tested and further developed through additional case studies. Findings of these studies can be implemented in the existing approach, which is flexible enough to cover various needs.

6 Acknowledgements

This framework is a main result of the project ReWarn, which was financially supported by the Swiss Federal Office for Civil Protection (FOCP). We want to thank Christoph Graf (Swiss Federal Institute for Forest, Snow and Landscape Research WSL), Giorgio Valenti and Marco Franzi (Canton Ticino) as well as Lorenz Meier (Geopraevent AG) for their support in the case studies. We also thank Unni Eidsvig (NGI) and an anonymous reviewer for their valuable comments, which helped to improve an earlier version of the paper.

7 References

Badoux, A., Graf, C., Rhyner, J., Kuntner, R., and McArdell, B. W.: A debris-flow alarm system for the Alpine Illgraben catchment: design and performance, Nat Hazards, 49, 517-539, 2009.

Bell, R., Mayer, J., Pohl, J., Greiving, S., and T, G.: Integrative Frühwarnsysteme für gravitative Massenbewegungen (ILEWS): Monitoring, Modellierung, Implementierung, Klartext Verlag, Essen, p. 270 pp., 2010.

Bliss, J. P., Gilson, R. D., and Deaton, J. E.: Human probability matching behaviour in response to alarms of varying reliability, Ergonomics, 38, 2300-2312, 1995.

Bründl, M., Romang, H., Bischof, N., and Rheinberger, C.: The risk concept and its application in natural hazard risk management in Switzerland, Natural Hazards and Earth System Sciences, 9, 801-813, 2009.

Bründl, M., and Heil, B.: Reliability analysis of the Swiss avalanche warning system, in: 11TH International Conference on Applications of Statistics and Probability in Civil Engineering, edited by: Faber, M., Köhler, J., and Nishijima, K., CRC Press an imprint of the Taylor & Francis Group, Zürich, 881-887, 2011.

Dai, F., Lee, C., and Ngai, Y. Y.: Landslide risk assessment and management: an overview, Engineering Geology, 64, 65-87, 2002.

Einstein, H. H., and Sousa, R.: Warning systems for natural threats, Proceedings Geohazards, Lillehammer, Norway, 2006, 2006.

Fuchs, S.: Cost-Benefit Analysis of Natural Hazard Mitigation, in: Encyclopedia of Natural Hazards, Springer, 121-125, 2006.

Glade, T., and Nadim, F.: Early warning systems for natural hazards and risks, Nat Hazards, 70, 1669–1671, 2014.

Glantz, M. H.: Usable science 8: early warning systems: do’s and don’ts, Report of workshop, Shanghai, China, 2003.

Grasso, V. F., Beck, J. L., and Manfredi, G.: Automated decision procedure for earthquake early warning, Engineering Structures, 29, 3455-3463, 2007.

Grasso, V. F., and Singh, A.: Early Warning Systems: State-of-Art Analysis and Future Directions, Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi, 2009.

Hess, J., and Schmidt, F.: Towards optimised early warning developments in Switzerland, 12th conference INTERPRAEVENT 2012, Grenoble, France, 2012, 2012.

Intrieri, E., Gigli, G., Casagli, N., and Nadim, F.: Brief communication "Landslide Early Warning System: toolbox and general concepts", Nat. Hazards Earth Syst. Sci., vol. 13, pp. 85-90, 2013.

Jackson, J.: Introduction to expert systems, Addison-Wesley Longman Publising Co. Inc., United States, 1990.

Jensen, F. V., and Nielsen, T. D.: Bayesian networks and decision graphs, 2 ed., Information Science and Statistics, edited by: Jordan, M., Kleinberg, J., and Schölkopf, B., Springer Science + Business Media, New York, 2007.

Johnson, M., Newstead, S., Charlton, J., and Oxley, J.: Riding through red lights: The rate, characteristics and risk factors of non-compliant urban commuter cyclists, Accident Analysis & Prevention, 43, 323-328, 2011.

Loew, S., Gischig, V., Moore, J., and Keller-Signer, A.: Monitoring of potentially catastrophic rockslides, Proc. of 11th Int. and 2nd North Am. Symp. on Landslides and Engineered Slopes, Banff, Canada, 2012.

Margreth, S., and Romang, H.: Effectiveness of mitigation measures against natural hazards, Cold Regions Science and Technology, 64, 199-207, 2010.

Martina, M. L. V., Todini, E., and Libralon, A.: A Bayesian decision approach to rainfall thresholds based flood warning, Hydrology and earth system sciences, 10, 413-426, 2006.

Meister, R.: Country-wide avalanche warning in Switzerland, Proceedings International Snow Science Workshop, Utah, USA, 1995.

Michoud, C., Bazin, S., Blikra, L., Derron, M.-H., and Jaboyedoff, M.: Experiences from site-specific landslide early warning systems, Natural Hazards & Earth System Sciences, 13, 2013.

Pate-Cornéll, M. E.: Warning Systems in Risk Management, Risk Analysis, 6, 223-234, 1986.

Penning-Rowsell E., J. C., Tunstall S., Tapsell S., Morris J., Chatterton J., Green C.: The Benefits of Flood and Coastal Risk Management: A Handbook of Assessment Techniques, Middlesex University Press, London, 2005.

Rheinberger, C. M.: Learning from the past: statistical performance measures for avalanche warning services, Nat Hazards, 65, 1519-1533, 2013.

Ripberger, J. T., Silva, C. L., Jenkins-Smith, H. C., Carlson, D. E., James, M., and Herron, K. G.: False alarms and missed events: the impact and origins of perceived inaccuracy in tornado warning systems, Risk Analysis, 35, 44–56, 2014.

Rogers, D., and Tsirkunov, V.: Implementing Hazard Early Warning Systems, Global Facility for Disaster Reduction and Recovery, 2011.

Romang, H.: Wirkung von Schutzmassnahmen, Nationale Plattform für Naturgefahren PLANAT, Bern, 2008.

Rosenbloom, T.: Crossing at a red light: Behaviour of individuals and groups, Transportation Research Part F: Traffic Psychology and Behaviour, 12, 389-394, 2009.

SafeLand: Quantitative risk-cost-benefit analysis of selected mitigation options for two case studies. Deliverable 5.3, SafeLand Project - Living with landslide risk in Europe: Seventh Framework Programme for research and technological development (FP7) of the European Comission, 2012.

Sättele, M., Bründl, M., and Straub, D.: A classification of warning system for natural hazards, Probabilistic Workshop, Stuttgart, 2012.

Sättele, M.: Quantifying the Reliability and Effectiveness of Early Warning Systems for Natural Hazards, PhD Ing., Technische Universität München TUM, Munich, 2015.

Sättele, M., Bründl, M., and Straub, D.: Reliability and effectiveness of warning systems for natural hazards: concept and application to debris flow warning, Reliability Engineering and System Safety, 142, 192–202, 2015a.

Sättele, M., Krautblatter, M., Bründl, M., and Straub, D.: Forecasting rock slope failure: How reliable and effective are warning systems?, Landslides, 605, 2015b.

Schmidt, R.: Warnsysteme in Wildbacheinzugsgebieten, Institut für Alpine Naturgefahren und Forstliches Ingenieurwesen, Universität für Bodenkultur, Wien, 2002.

Shachter, R. D.: Evaluating influence diagrams, Operations Research, 34, 871-882, 1986.

Simmons, K. M., and Sutter, D.: False Alarms, Tornado Warnings, and Tornado Casualties, Weather, Climate, and Society, 1, 38-53, 2009.

Sorensen, J.: Hazard Warning Systems: Review of 20 Years of Progress, Natural Hazards Review, 1, 119-125, 2000.

Špačková, O., and Straub, D.: Cost-benefit analysis for optimization of risk protection under budget constraints, Risk Analysis, 35, 941-959, 2015.

Stähli, M., Sättele, M., Huggel, C., McArdell, B. W., Lehmann, P., Van Herwijnen, A., Berne, A., Schleiss, M., Ferrari, A., Kos, A., Or, D., and Springman, S. M.: Monitoring and prediction in Early Warning Systems (EWS) for rapid mass movements, Natural Hazards and Earth System Sciences, 15, 905-917, 2015.

Straub, D.: Natural hazards risk assessment using Bayesian networks, 9th International Conference on Structural Safety and Reliability, ICOSSAR, Rome, Italy, 2005.

Straub, D., and Der Kiureghian, A.: Bayesian network enhanced with structural reliability methods: Methodology, Journal of engineering mechanics, 136, 1248-1258, 2010.

Sturny, R. A., and Bründl, M.: Bayesian networks for Assessing the reliability of a Glacier Lake warning System in Switzerland, Interpraevent 2014 in the Pacific Rim - Natural Disasters Mitigation to Establish Society with the Resilience, Nara, Japan, 2014,

Swets, J. A.: Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers, Lawrence Erlbaum Associates, Inc, New York, 1996.

Techel, F., and Darms, G.: Schnee und Lawinen in den Schweizer Alpen - Hydrologisches Jahr 2012/13, WSL-Institut für Schnee- und Lawinenforschung SLF, 2014.

Thiebes, B.: Landslide Analysis and Early Warning Systems: Local and Regional Case Study in the Swabian Alb, Germany, Springer, 2012.

Wei, Y., Chamberlin, C., Titov, V. V., Tang, L., and Bernard, E. N.: Modeling of the 2011 Japan tsunami: Lessons for near-field forecast, Pure and Applied Geophysics, 170, 1309-1331, 2013.

Wetterhall, F., Pappenberger, F., Cloke, H. L., Thielen-del Pozo, J., Balabanova, S., Daňhelka, J., Vogelbacher, A., Salamon, P., Carrasco, I., Cabrera-Tordera, A. J., and others: Forecasters priorities for improving probabilistic flood forecasts, Hydrology and Earth System Sciences Discussions, 10, 2215-2242, 2013.

Wheater, H., Chandler, R., Onof, C., Isham, V., Bellone, E., Yang, C., Lekkas, D., Lourmas, G., and Segond, M.-L.: Spatial-temporal rainfall modelling for flood risk estimation, Stochastic Environmental Research and Risk Assessment, 19, 403-416, 2005.

Willenberg, H., Eberhardt, E., Loew, S., McDougall, S., and Hungr, O.: Hazard assessment and runout analysis for an unstable rock slope above an industrial site in the Riviera valley, Switzerland, Landslides, 6, 111-119, 2009.

Zschau, J., and Küppers, A. N.: Early Warning Systems for Natural Disaster Reduction: With 79 Tables; EWC'98;[this Volume is the Result of the International IDNDR-Conference on Early Warning Systems for the Reduction of Natural Disasters, Held at the Geo-Forschungszentrum in Potsdam, Germany from 7-11 September 1998., Springer, 2003.