Stochastic Design of Early Warning Systems

Z. Medina-Cetina, F. Nadim
International Centre for Geohazards / Norwegian Geotechnical Institute, Oslo Norway

ABSTRACT: Early Warning Systems (EWS) are designed to avoid, or at least to minimize the impact imposed by a threat. Identifying information patterns generated by the different information sources is fundamental for predicting a catastrophic event, and thus for defining realistic warning levels. Since EWS are time sensitive or stochastic, it becomes necessary to develop a methodology that integrates the different monitoring information sources and that takes into account all possible performing scenarios of the system. This paper discusses the development of a Stochastic Design of an Early Warning System (SDEWS) introducing a risk measure as the reference variable which enables the integration of the different effects captured by the monitoring instruments. In this way, the risk measure serves as a rational index for the definition of warning thresholds, but also introduces EWS within a decision-making framework. For this purpose, a Bayesian approach is proposed as a suitable tool for integrating and updating the joint states of information, for updating the warning level of the system, and for facilitating the decision-making required for the issuing of the warnings. Some of the methods proposed for implementing the SDEWS are also discussed (Bayesian smoothing, Bayesian filtering and Bayesian networks).

1 INTRODUCTION

Early Warning Systems (EWS) are monitoring devices designed to avoid, or at least to minimize the impact imposed by a threat on humans, damage to property, the environment, or/and to more basic elements like livelihoods. The recent catastrophic natural disasters, like the Indian Ocean tsunami of 26 December 2004, the global concern of the impact of climate change, and evidence of growing social disruption as a consequence of natural hazards have triggered the interest on advancing the state of knowledge on EWS. This is corroborated in the Report on Global Survey of Early Warning Systems released by the United Nations (2006). The report recognizes that although existing technology covers almost all types of natural-type hazards, the optimization of technological resources, the maximization of inferences from available information resources, and the use of moderate skills for their implementation could have a strong impact in communities with deep economical needs, or in communities where the investment on an EWS could guaranty sustainable development.

EWS, when efficiently designed, can effectively help reduce the hazard, vulnerability or consequences associated to a threat. From a review on current applications of EWS spanning the natural, social and economical hazards, it was found that inferences gained from information generated by the monitoring devices can be improved methodologically (see section 2). In general, statistical methods considered for determining the possibility of issuing (or not) an early warning concentrate only on identifying trend or rate changes observed on the monitoring data. The capacity to include the influence that uncertainty plays into the decision-making involved in the operation of an EWS is in most cases disregarded, not to mention the influence across decision variables, or the capacity to forecast possible events conditioned on historic data. Actually, in many cases, it is common to find contrasting well-developed EWS at the technological level, but underdeveloped for data assimilation and consequently for making inferences.

Identification of patterns of information generated by a limited number of monitoring instruments, and the required simplification for interpretation, independently of the sophistication of the methods considered, becomes fundamental for predicting a catastrophic event, and for defining and issuing the corresponding warning levels. The incorporation of available knowledge, empirical, theoretical and
even subjective, for the integration of the joint state of knowledge about the potential threat, would help improve the predictive capability of the EWS, something that is missing in most of the existing systems. Furthermore, since EWS are time sensitive or stochastic, it becomes necessary to account for a methodology that integrates the different monitoring information sources in a time-framework (and eventually to space), that accounts for multi-variable and possibly multi-level information sources, and that considers all possible performing scenarios of the system (system reliability). Such methodology must also be capable of incorporating the uncertainty from participating information sources, so that more realistic scenarios can be predicted with the aim of facilitating the decision-making for the issuing (or not) of the early warning.

This work discusses the development of a Stochastic Design of an Early Warning System SDEWS, introducing a risk measure as the reference variable that integrates the effects of each of the information sources, serves as a rational index for the definition of warning thresholds, and naturally incorporates EWS within a decision-making framework. It is worth mentioning that although the concept of risk has proved to be a consistent and robust parameter in many engineering fields (NRC, 1994); there is a growing interest in many social disciplines to extend its definition with the aim to make it more people oriented (ISDR, 2005). A discussion on current and emerging risk definitions are presented later in this paper (see section 3).

Under the previous premises, a Bayesian approach is proposed as a suitable avenue for integrating and updating the joint states of information included into the EWS. Consequently, a Bayesian formulation would allow for a systematic updating of the current warning level. This is possible by applying quantitative methods capable of managing multi-variable and multi-level data for Bayesian smoothing (learning from past data) and Bayesian filtering (learning from predictions based on past data). A summary of some of these methods is discussed later in this work (see section 4). Also, a work in progress of a Bayesian Network (BN) is presented to illustrate the benefits of this technique to facilitate the decision-making process for the issuing of an early warning in a geohazard context (see section 5).

2 EARLY WARNING SYSTEMS

In light of recent cases of significant social disruption caused by natural hazards (Sousa and Einstein, 2006), advancing the state of knowledge about EWS has become an urgent need because in many situations the risk associated with natural threats can only be mitigated by EWS. According to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change IPCC Working Group I (2007a), the global concern on the increase of magnitude and intensity of hydro-meteorological events is unfortunately not a hypothesis anymore but a fact. In this regard, following the findings of Working Group I, Working Groups II and III (2007b and 2007c) make a clear statement on focusing international efforts on adapting to imminent consequences of climate change, and in the need of implementing global mitigation measures. These findings align with previous results presented at the Hyogo Framework for Action 2005-2015 (2005), where “Risk identification, assessment, monitoring and early warning” was outlined as one of the five disaster reduction priorities.

Natural hazards are thus one of the primary fields of application for EWS (UN, 2006). Some of the fields where EWS have developed within this area are introduced in Table 1. From a literature review on current applications of EWS, it was found that inferences gained from information generated by the monitoring devices can be significantly improved by introducing advanced information methodologies. In general, statistical methods considered for determining early warnings concentrate only on identifying trend or rate changes on the monitoring data. This is the case of transforming the parameter space into a reduced space where controlling variables can be identified using Principal Component Analysis (PCA), or/and special statistical techniques to identify atypical observations (Jiji et al., 2003). In this regard, further development of data assimilation can be developed since information sources sometimes require not only data identification but also data classification (Duda et al., 2001), and to include space and time parameters, particularly when defining threshold levels using multiple information sources.

Additionally, the capacity to include the influence that uncertainty plays into the decision-making involving EWS is in most cases disregarded, not to mention the influence it has across decision variables, or the potential it has to forecast its influence conditioned on historic data. This means that
the capacity to include model-based predictions about the threat, to continuously update the model parameters, and to even include expert beliefs can be achieved by implementing advanced inference tools. Some cases found in the literature where uncertainty was incorporated to some extent as part of EWS, were applied in the fields of coastal earthquakes (Cervone, etc, 2006), erosion in highlands (Hj and Kia Hui, 2006), dam engineering (Smith, 2006), bank financial operations (Sarkar and Sriram, 2001), shipboard fire-detection (Kuo and Chang, 2003; Rose-Pehrsson et al., 2003), hydropower stations (Liu et al, 2004), design of underground facilities (Olsson and Stille, 2002), earthquake engineering (Iwata et al, 2005), and weather forecasting (Libbert et al., 2006), among others.

<table>
<thead>
<tr>
<th>Table 1 Common Early Warning Systems applied to natural hazards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydro-meteorological</td>
</tr>
<tr>
<td>• Floods</td>
</tr>
<tr>
<td>• Tropical cyclones</td>
</tr>
<tr>
<td>• Severe storms</td>
</tr>
<tr>
<td>• Drought</td>
</tr>
<tr>
<td>• Extreme temperatures</td>
</tr>
<tr>
<td>• Air pollution</td>
</tr>
<tr>
<td>• Dust and Sandstorms</td>
</tr>
<tr>
<td>• Snow Avalanches</td>
</tr>
<tr>
<td>• Famine</td>
</tr>
<tr>
<td>Geological</td>
</tr>
<tr>
<td>• Earthquakes</td>
</tr>
<tr>
<td>• Tsunamis</td>
</tr>
<tr>
<td>• Volcanoes</td>
</tr>
<tr>
<td>• Landslides</td>
</tr>
<tr>
<td>• Near-earth objects</td>
</tr>
<tr>
<td>Biological</td>
</tr>
<tr>
<td>• Epidemics</td>
</tr>
<tr>
<td>• Locust swarms</td>
</tr>
<tr>
<td>Environmental Degradation</td>
</tr>
<tr>
<td>• Desertification</td>
</tr>
<tr>
<td>• Wildfire</td>
</tr>
</tbody>
</table>

3 RISK ASSESSMENT

This section discusses the development of a stochastic design of an early warning system. It introduces a risk measure as the reference variable that integrates the effects of each of the warning information sources, serves as a rational index for the definition of warning thresholds, and naturally incorporates EWS within a decision-making framework. Following the work by Sousa and Einstein on ‘Warning Systems for Natural Threats’ (2006), the engineering definition of Risk $R$ for a specific state of information is:

$$R = P[T] \cdot u(C) \quad (1)$$

where $P[T]$ is the hazard or probability of occurrence of possible threats $T$, and $u(C)$ is the loss or utility of a set of consequences $C$ that are all certain to happen to some extent. In the case where the loss or utility is weighted by a link term that conditions the occurrence of $C$ to certain level of intensity of $T$, the risk concept is redefined as:

$$R = P[T] \cdot P[C|T] \cdot u(C) \quad (2)$$

where the link factor $P[C|T]$ is called vulnerability.

EWS are considered risk mitigation agents within a decision-making framework. In this case, the risk mitigation mechanism is illustrated in Fig.1, which shows that the impact of early warning can be due to Active Countermeasures $AC$ acting on the hazard reduction (i.e. influencing directly to reduce or avoid the threat), and/or due to Passive Countermeasures $PC$ acting on the vulnerability reduction (i.e. influencing directly to reduce or avoid the vulnerability or the consequences). The implementation of either one of them generates the costs $u(AC)$ and $u(PC)$ respectively. Therefore, the trade-off between the savings induced by the hazard or vulnerability reduction and the costs associated to their implementation is what defines the risk measure, and consequently the warning levels.
Due to the stochastic nature of EWS, the risk measure $R$ can be defined as a stochastic process, either referenced only to time:

$$R(t) = P[T(t)] \cdot P[C(t)|T(t)] \cdot u(C(t))$$  \hspace{1cm} (3)$$

or to time and space:

$$R(X, t) = P[T(X, t)] \cdot P[C(X, t)|T(X, t)] \cdot u(C(X, t))$$  \hspace{1cm} (4)$$

where $X$ represents the spatial parameters and $t$ the time parameter. Such definitions implicitly assume that the threat and the consequences are stochastic processes. For instance, the impact of a meteorological condition such as a hurricane varies in time and location. Similarly, consequences such as traffic due to commuting are also a function of location and time. In this way, the complexity of warning systems can be in some cases resolved properly if adequate references of space and time are defined (e.g. stochastic risk maps), and if data assimilation tools capable of dealing with spatio-temporal analyses are available (i.e. those capable of managing extreme stochastic conditions such as non-Gaussianity and non-stationarity). Additionally, the integration of a system with temporal or spatio-temporal characteristics demands an efficient representation and of the dynamics of the decision-making under the presence of EWS. Fig.2 shows a simple scheme that illustrates the sequence for estimating the different information sources that integrate the risk measure, including the updating effect introduced by the EWS.
It is worth mentioning that although the concept of risk has proved to be a consistent and robust parameter in many engineering fields, there is a growing interest in many social, environmental and economic disciplines to extend its definition with the aim to make it more people-oriented. A specific shift on the risk definition is observed for including some of non-physical factors, particularly related to risk and vulnerability. The current structure for estimating a risk measure is considered robust enough as to incorporate risk components such as exposure and coping capacity, and to manage different information levels (i.e. individual, local, regional, national and international). A summary of recent advances on the development on new risk and vulnerability definitions based on this approach is presented by Villagran de Leon (2006).

4 BAYESIAN METHODS

As mentioned before, the Bayesian paradigm fits well the needs for integrating multi-variable, multi-level, objective and subjective information sources, and even spatio-temporal referenced risk measures. Most importantly, it solves the smoothing and filtering conditioned on data available, resulting in a more rational assessment of the uncertainty, and consequently on the impact it has when making decision about issuing early warnings. Elemental concepts related to the Bayesian theory are introduced below, as a way to familiarize the readers with the fundamental concepts in which the design proposed herein was based.

The Bayesian paradigm is defined as:

\[
\pi(\theta | d_{\text{obs}}) = \frac{f(d_{\text{obs}} | \theta) \pi(\theta)}{\int f(d_{\text{obs}} | \theta) \pi(\theta) d\theta} = \frac{f(d_{\text{obs}} | \theta, g(\theta)) \pi(\theta)}{\int f(d_{\text{obs}} | \theta, g(\theta)) \pi(\theta) d\theta} 
\]

(5)

where the prior \( \pi(\theta) \) represents the a-priori state of information associated to the a set of parameters \( \theta \), the likelihood \( f(d_{\text{obs}} | \theta) \) represents the a-priori state of information associated to the potential of the parameters \( \theta \) to match the observations \( d_{\text{obs}} \), or to help to match them if they are embedded into a predictive model \( g(\theta) \). The posterior \( \pi(\theta | d_{\text{obs}}) \) is thus the joint probability function between the a-priori states of information associated to both the prior and the likelihood.

For estimating first and second moments of the posterior, or for estimating marginal statistics for each parameter included in \( \theta \), it is required to integrate the posterior with respect to the parameters domain. This integral can become cumbersome when the number of parameters spans in a high dimensional space. Fortunately the integral can be solved numerically. This solution yields a full description of the uncertainty associated to the model parameters \( \theta \) conditioned on the available data.

An efficient way to solve the posterior integral is using Markov Chain Monte Carlo (MCMC) (Robert and Casella, 2004). A useful property of the MCMC is that it converges to the target joint density as the sample integration grows. A common decision rule that determines which samples are ‘accepted’ or ‘rejected’ is the Metropolis-Hastings (MH) criteria. Therefore, the posterior integration at the MH ‘state’ of the chain \( s+1 \) iteration is obtained by sampling a candidate point \( Y \) from a proposal distribution \( q(\hat{\theta}_s) \), where the candidate point \( Y \) is accepted or rejected as the next step of the chain with probability given by:

\[
\alpha(\hat{\theta}_s, d_{\text{obs}}) = \min \left\{ 1, \frac{\pi(Y | d_{\text{obs}}) q(\hat{\theta}_s | Y)}{\pi(\hat{\theta}_s | d_{\text{obs}}) q(Y | \hat{\theta}_s)} \right\} 
\]

(6)

For the MCMC sampling the distribution of interest \( f(d_{\text{obs}}) \) appears as a ratio, so that the
constant of proportionality cancels out. Additionally, the evaluation of the posterior requires discarding the first iterations called the burn-in points, before it reaches the stationary condition from which the statistical inferences are generated. In the case where the risk measure is defined as a stochastic process, Equations 5 and 6 can be reformulated introducing spatio-temporal referred parameters \( \theta(X,t) \).

The previous Bayesian formulation helps to solve the smoothing, which allows for identifying and classifying information based on past observations (Denison et al., 2002; Duda, 2001). This means finding the probabilistic definitions of \( \theta \) that will best approximate the available data through the model \( g(\theta) \). Filtering in the other hand, aims at estimating the probability that certain conditions will occur in the future also conditioned in past observations. Within the Bayesian framework, filtering is defined as the posterior (Barker et al., 1995; Van der Merwe et al, 2000; Fox et al., 2003):

\[
\pi(d_q | d_{obs}) = \frac{f(d_{obs} | d_q) \pi(d_q)}{\int f(d_{obs} | d_q) \pi(d_q) dd_q} = \frac{f(d_{obs} | d_q, h(d_q)) \pi(d_q)}{\int f(d_{obs} | d_q, h(d_q)) \pi(d_q) dd_q}
\]  

(7)

where \( d_q \) represents a vector of predictions at time \( q \) of all information sources taking part on the estimate of the risk measure, and \( h \) represents the forecasting model.

The integration of smoothing and forecasting results for the assessment of the risk measure is placed at the centre of many research efforts since it represents the conjunction of different states of probabilities. Tarantola (2005) proposes a solution for the conjunction of probabilities as:

\[
f_z \cap f_2(d) = \frac{1}{\nu} \frac{f_z(d) \cap f_2(d)}{\mu(d)}
\]

(8)

where \( f_z \cap f_2(d) \) represents the joint of two states of probability associated to the state vector \( d \), \( \nu \) is a normalization constant given by the integral of the second factor in Equation 8, and \( \mu(d) \) represents the homogeneous distribution, which reflects the minimum state of knowledge about \( d \). In the case where the probability space for the state vector \( d \) is referenced to space or time, the stochastic definition of the joint states of information would be the extension of Equation 8, but now including the spatio-temporal parameters \( X \) and \( t \).

5 SCHEMATIC STOCHASTIC DESIGN OF AN EWS USING BAYESIAN NETWORKS

In the geohazards community in Norway, there is an interest in the risk assessment for dynamic systems, particularly the ‘landslide-tsunami’ interaction as a possible threat for the coastline of a fjord or lake. For this purpose, a conceptual and methodological review has been developed by the authors for the SDEWS. The current work in progress spans to the implementation of Bayesian tools of smoothing and forecasting into a decision-making framework. In this sense, the authors find that Bayesian Networks BN (Heckerman et al, 2000; Korb and Nicholson, 2004) is the appropriate avenue for taking into account the dependencies between information sources and for using it as a robust probabilistic template capable of updating the risk measure, as real time information flows through the EWS.

Fig.3 illustrates a simple schematic design for the assessment of the risk measure under a decision-making framework that includes the SDEWS. In this case, information sources about rain precipitation, snow precipitation, pore pressure, slope inclination, slope material composition, fjord bathymetry, and even the motion of the landslide itself are monitored by an arrange of sensors that report to a central station referred as EWS. A direct dependency is considered between the landslide and the tsunami, and from the tsunami to the vulnerability. The risk measure is thus calculated using the updated version of Equation 2 (Fig.1), after assessing the hazard imposed by the tsunami, the vulnerability, and the corresponding consequences.

Each of the components of the BN are assumed as possible events that pass the contribution of the
information in form of probability measure throughout the network until the risk measure. The potential benefits of such approach is that it allows for the incorporation of sudden information changes at each event (e.g. a sensor stops working), and for the corresponding updating of the risk measure which includes the collected uncertainty from the network. The process for tuning of the Bayesian smoothing and forecasting, called probabilistic calibration (Medina-Cetina, 2006), becomes the process for defining and later issuing the most appropriate warning levels.

Future work on the development of this particular system, include data exploration for the definition of the probability states of each of the events considered in the network, as well as the development of computational resources for the smoothing, forecasting, and the updating of the risk measure given by the BN. Although validation of the EWS cannot be performed unless there is data available triggered by actual threats, simulations of the system under extreme conditions can be performed as a way to test the inference system and the decision making.

6 SUMMARY

This work has introduced the concept of Stochastic Design of Early Warning Systems. Conceptual and methodological elements required for its implementation were discussed. A unified approach based on risk measure is proposed as a reference index for the definition and issuing of early warnings. Bayesian principles were proposed for continuously integrating different information sources and consequently for updating the risk measure. This is possible under multi-variable, multilevel, and objective and subjective conditions. A work in progress was introduced related to the SDEWS associated to the ‘landslide-tsunami’ threat, where Bayesian Networks were considered as an appropriate venue for updating of the risk measure based on decisions based under the presence of an EWS.

Fig.3 Bayesian Network for the risk assessment of the ‘landslide-tsunami’ threat

ACKNOWLEDGMENTS.

The work presented in this paper was supported by the Research Council of Norway through the Centre of Excellence “International Centre for Geohazards” (ICG). The authors gratefully acknowledge this support.
REFERENCES


