Training decision-makers in hazard spatial prediction and risk assessment: ideas, tools, strategies and challenges

A. G. Fabbri¹ & C. J. Chung²

¹DISAT, University of Milano-Bicocca, Italy ²Department of Earth Sciences, University of Ottawa, Canada

Abstract

Hazard prediction and risk assessment over regions exposed to natural and technological processes are complex tasks that require exposure to quantization of its uncertainty related to the prediction of future events through statistical methods, spatial data analysis, case studies and process evolution interpretation in conditions of uncertainty. All too often decision makers, DMs, similarly to judges in environmental legal practice, do not have technical training to enable them to communicate/understand the associated uncertainty from technical specialists. In particular communication is a challenge with those who can provide prediction maps and associated statistics to support decisions on disaster prevention, avoidance or mitigation. An interactive short course was prepared to overcome such obstacles to responsible land use planning and proactive measure taking, for example, by asking a set of questions. A first phase in the training follows steps that are to facilitate the comprehension of a spatial database on landslide hazard, of its data processing, and of the interpretation of the analysis results. Integral parts of a second phase are the theory of predictive methods, the strategy in prediction map generation and visualization, including validation via blind tests and the representation of the associated spatial and prediction uncertainties. A successive third phase of the training brings in environmental and socioeconomic spatial indicators to assign vulnerabilities and values to exposed elements in the spatial database. Scenarios for hazard development in the future are then provided. They allow to estimate the uncertainty associated with the probabilities of hazardous occurrences and to resolve the risk equation for different settings. The DM training course includes interactive and iterative



techniques for general purpose "target mapping" and helps to derive spatial uncertainties to comprehend hazard and risk maps.

Keywords: hazard prediction, cross-validation, blind tests, mathematical models, analytical strategy, training tools, training strategy, decision-makers.

1 Introduction

The idea of dedicating a course on hazard and risk analysis to decision-makers may not be new. However, given the disquieting aspects of some recent disasters (e.g. May 1998 mudflow in Sarno, Italy, [1]; the much debated 2003 Hurricane Katrina in Louisiana, USA, [2]; or the recent April 2009 earthquake in L'Aquila, Italy, [3]) there seem to be a need for applied research to effectively communicate with decision-makers. As advocated by Alexander [4] the three classes of stakeholders in geologic hazard, researchers, practitioners and consumers, are to avoid a "dialogue of the deaf." Consumers of research are the general public, industry and commercial business, and decision-makers are generally not scientists so that administrators, politicians and members of the public are not readers of technical writings. This leads to a major communication gap.

Living with risk, nevertheless, is a human condition that has affected all living beings since they came about. In his "story of risk", Bernstein [5] wrote that (pp. 1-2) "the revolutionary idea that defines the boundary between modern times and the past is the mastery of risk, the notion that the future is more than the whim of the gods and that man and women are not passive before nature ... the ability to define what may happen in the future and to choose among alternatives lies at the heart of contemporary societies. Risk management guides us over a vast range of decision-making." One would expect to see that developed societies had established criteria and measures to cope with risk in a coordinate manner. For instance, Monmonnier [6] noted that methods of anticipation have followed traditional descriptive practices and provided examples of "cartographies of danger." He concluded that hazard zone mapping is a recent phenomenon that seems to focus on forecasting and monitoring and this means that risk maps involve models for estimating the likelihood of rare events that are inherently uncertain.

It turns out that the estimation aspect is in itself a technical obstacle to risk management. A difficulty in comprehending risk maps and risk estimates is the lack of exposure to them of decision-makers. Posner [7] in his book "Catastrophe: Risks and Response" expressed doubts that risk or major disasters can be managed, understood or predicted: risk analysis being still a scientific endeavour and often judges like himself are not in the position to evaluate how to decide on risk matters or controversies, mostly because of their lack of exposure to science, so that they are not prepared to understand technical information.

This problem is compounded by the scarcity of legal instruments. Bonachea [8] reviewed environmental legislation on natural hazard and observed that in Europe there is no law or directive to manage the territory in relation with natural



risks, except for the October 16, 1989 Council Resolution for the prevention of natural and technological risks (Official Journal of the EU, 1989), which calls for the preparation of a statement on natural and technological risks. Within Europe the situation is different from country to country. France has, for instance, three classes of risk maps concerning floods and landslides, since the 1990s. Legislation in other European countries is more recent and not yet satisfactorily applied so that unified norms can be used and a clear mandate for whom to make them is established.

Alexander [9] remarks that after the loss of 153 lives in a mudflow disaster in Sarno, Italy, in 1998, a national law was passed for landslide risk estimation and analysis, obliging the country's 20 regions to assess landslide hazard in relation to vulnerability of human settlements, communications and activities. A web search on geologic hazard and risks and on the activity of many local and national administrations, however, shows that much has still to be done to put that into practice and construct risk maps. This appears to be so also worldwide even if risk maps appear of increasing concern [10, 11].

An example that could represent a new trend is a set of initiatives by the State of Oregon in the US. In a website [12] mention is made of regional landslide hazard maps and reports to be released that use laser based terrain mapping. Each report contains three different maps: (1) a LIDAR-based landslide inventory map; (2) a shallow-seated landslide susceptibility map, and (3) a deep-seated landslide susceptibility map. The inventory of past landslides is used to create the susceptibility maps that display areas at risk for future landslides. The maps are to be used in the planning and decision process. They are also available for sale on CD-ROM and as printed sheets (DOGAMI, 2009: Open File Report O-08-09). This initiative is a consequence of previous measures taken by the City of Portland Oregon Bureau of Development Services since 2003, to reduce the risk of private and public losses as a result of landslides, after the February 1996 storms.

Such initiatives are not widespread as yet. However, some progress is being made locally by various administrations worldwide, so that the need to improve and maintain communications between applied research and decision-making is becoming increasingly felt. In particular it is critical for decision-makers to be exposed to how hazard is estimated, to what assumptions are implied in prediction modelling, and to what scenarios can be constructed for risk assessment.

Hazard prediction and risk assessment over regions exposed to natural and technological processes are complex tasks that require exposure to quantization of its uncertainty related to the prediction of future events through statistical methods, spatial data analysis, case studies and process-evolution interpretation in conditions of uncertainty.

This contribution introduces a short course for decision-makers, DMs, on hazard prediction modelling and risk assessment. Phases in training and analysis are linked with critical questions pointing at the importance of understanding how to model using a quantitative logical framework. A database on landslide occurrences in northern Spain is used to explore software techniques for spatial



modelling. Expected results and further challenges for DMs and for trainers are indicated.

2 A short course

An interactive short course tailored to decision-makers was prepared to overcome such obstacles to responsible land use planning and proactive measure taking, for example, by asking a set of questions to expose the motivations for modelling. Its duration is about two days consisting of six lecture hours and over six hours of interactive exercises with special-purpose software on hazard spatial prediction and risk assessment. Scope of the course is to familiarize decision-makers with the methods and techniques of spatial analysis via real-life case studies on natural hazard. The context of spatial analysis is a representation of the landscape and the environmental conditions in it through the spatial database. To construct, use and comprehend the significance of such a database a number of fundamental questions need to be asked and answered. They are to place the decision-makers well within the role of a modeller.

Some of the questions are as follows. The first three questions are preliminary to applying prediction models.

- 1. Is natural hazard in the area of interest a priority in itself and in the context of other hazards such as highway traffic's or epidemics', for instance?
- 2. Natural hazard is the consequence of natural processes encroaching human activities. Should decision-makers have some degree of knowledge of such natural processes or should all such knowledge remain fully within the research and the user community?
- 3. Is it necessary, feasible or convenient to predict or mitigate hazard and risk? If the answer is a yes, who is to undertake the role and the responsibility of such a task? If the answer is a no, what is the consequent damage to society?

At this point prediction modelling is justified.

- 4. Given a study area in which natural hazard is known to have occurred and is expected to occur in the future, what sorts of land-use decisions or measures are desirable or feasible? Where? For how long in the future?
- 5. Quantification of spatial data for hazard representation and prediction implies a number of unavoidable assumptions. Who is to have knowledge of them? So is for the prediction models, their theoretical background and applications to spatial databases.
- 6. Given some estimate of hazard levels and distribution over the area of concern, what are the benefits and the costs of taking alternative actions including the do-nothing alternative?
- 7. What degree of transparency and communication to the general public would you consider acceptable of the representations of hazard and risk, e.g. maps, over the area of concern? This when considering also documented case studies at hand worldwide.



- 8. The concept of hazard, either natural or technological, is based on the probability of occurrence of damaging events. What exposure should decision-makers have on how to obtain such probabilities?
- 9. An assessment of the vulnerability of all elements exposed to the hazard is a prerequisite to risk analysis. What degree of insight on them is proper for decision-makers?
- 10. What reasonable scenarios for risk assessment should be used to take decisions on risk reduction?
- 11. Many consultants or agencies claim to be able to provide hazard and risk maps upon demand. What should a decision-maker know in order to evaluate them? Is the good name of such consultants or agencies sufficient ground to take decisions? Where lays the ultimate decisional responsibility?

These and many more similar questions need to be considered in order to focus on what should or would decision-makers want to know from the short course. The course stems from the results of a case study developed in EC Research Network Projects on the assessment of landslide risks and mitigation in mountain areas [13–15].

3 Training strategy phases

The training strategy consists of three phases each comprising several steps (within brackets the associated questions listed in Section 2):

(i) Spatial prediction of hazard

- comprehension of spatial database (1, 2, 3)
- interpretation of analyses of spatial relationships of hazardous events (4, 5)
- exposure to data processing (5)

(ii) Interpretation of hazard maps

- theoretical background to a spatial prediction model: the likelihood ratio function (5, 6)
- visualization of predicted hazard (7)
- validation of prediction maps via blind tests (7)
- representation of prediction-map uncertainties (7)
- strategy of prediction-map generation (8)

(iii) Spatial risk-assessment

- comprehension of spatial socioeconomic environmental indicators (9)
- assignment of vulnerability and value to exposed elements (9)
- simulation and verification of scenarios of future hazard development (10)
- application of risk equation (11)
- assessment of impact of hazard uncertainties on risk assessment (11)



Phase (i). Spatial prediction of hazard requires a number of assumptions concerning the database, its resolution and relevance in terms of characteristics of predictable settings of hazardous occurrences, about the quantification of the spatial data, the processing steps to obtain spatial relationships, and for the mathematical models to be used to obtain prediction maps. Out of several prediction models applicable to spatially distributed data, such as those based on Bayesian probability, linear or logistic regressions, fuzzy sets, Dempster-Shafer's evidential theory, and the likelihood ratio function, the latter is discussed and applied in the course given its more immediate interpretability.

Phase (ii). Interpretation of hazard maps requires exploring briefly the theoretical background of prediction modelling and how the models can be applied to the database to obtain estimates of the likelihood of hazardous occurrences. Furthermore, it is critical in it to be able to visualize and interpret the results of the spatial predictions, either in terms of coloured maps with prediction classes, or as distribution of future events in those classes. Given that the prediction classes obtainable are only interpretable in relative terms, it becomes essential to plan a number of validation experiments in which other prediction maps are generated using only the distribution of part of the known hazardous occurrences. Such prediction maps are then compared with the distribution of the part of the occurrences not used in the prediction. The new distributions provide empirical assessment of the quality, "goodness", of the prediction maps. Generally the above partition of occurrences are used to generate sets of prediction maps, whose predicted values are then overlaid and compared to obtain the associated uncertainty map. Acceptable levels of uncertainty are then selected to identify the distribution of the most reliable predicted classes of hazard and proceed in the risk assessment to follow.

Phase (iii). Spatial risk-assessment requires to construct and discuss scenarios concerning the temporal and spatial significance of the predicted hazard and to estimate the probability of occurrence as maps and functions. Time partition of the hazardous occurrences enables to obtain predictions in time intervals. Such intervals have to be compatible with those for which there is a social concern, e.g. 20, 30, 50 or 100 years. Additionally, several risk maps will be generated for people, dwellings, infrastructures, or land uses, according to the spatial information available on their distribution, value and vulnerability to the hazard. Also an aggregate risk map is generated.

Scope of the three phases is to gradually develop awareness of the significance of data, assumptions, predictions, estimates and spatiality in order to support the decision-making process.

4 A database for hazard prediction and risk assessment

The database used for the short course deals with landslide hazard. However, the hazard prediction methodology is of complete generality so that it can be directed to many more types of hazard, from flooding, subsidence, erosion, seismic activity, volcanism, etc., in cases in which spatial information is available to characterize the condition of occurrence of all hazardous events.



The study area, is part of a spatial database for the lower Deba valley in the Basque province of Guipúzcoa in northern Spain. The database has been the subject of a number of hazard studies [13, 16]). It has been later expanded with addition of socioeconomic information to enable risk assessment [8, 14, 15]. It covers approximately 140 km², with elevations ranging between 0 m and 700 m a.s.l., and the main annual rainfall is 1500 mm. Geologically, the terrain consists of moderately folded and faulted limestones, marls, claystones, sandstones, flysch facies and volcanic rocks of the Cretaceous and Paleogene of the Basco-Cantabrico Pyrenees. Slopes are generally steep (average slope gradient is 22°) and there are surficial deposits of different composition and thickness that determine the occurrence of many hundreds of landslides triggered by intense rainfall episodes. Extensive field surveys and photo-interpretations of different sets of flight coverage allowed mapping landslides of different dynamic types in temporal groups [16]. For this study 1206 shallow translational landslides and associated flows were considered: 300 that occurred post-1997 (1998-2001), and 906 pre-1997 (1991-1997). Because some were reactivated after 1977, a subtotal of 1123 landslides was eventually used, with only 217 considered as post-1977.

The digital database consisted of rasterized images of 1886 columns x 1555 rows of 10 m pixel resolution. Within that image space the area of concern occupied 1,393,541 pixels. The average size of the landslide main failure is about 400 m², and the location of each was assigned to a single pixel of 10 m resolution. The digital images that contain the location of the 1123 landslides (906 pre-1977; 217 post-1977), represent the evidence of landslide occurrence and are considered as *direct supporting data layers*. The six "causal factor" data layers that were used are three continuous ones (elevation, slope and aspect) and three categorical ones (geology with 28 map units, vegetation-land use with 7 units, and thickness of surficial deposits with 3 units, between 0.5 and 3.0 m). They are believed to sufficiently represent the typical settings of the landslide occurrences and are considered as *indirect supporting data layers*.

The spatial database has been used in different experiments in which distinction is made between the *training area*, i.e. a sub-area within which the spatial relationships are established between the direct and indirect supporting data layers, and the *study area*, a sub-area across which those relationships are extended to obtain prediction classes of hazard. Such classes represent the *target image*, resulting from the prediction modelling, i.e. the likely location of the future occurrences that we would like to obtain. The term support relates to *the proposition* that a pixel will be affected by a landslide given the combined presence of the mapping units at its location. Such a proposition is supported with different levels in the prediction classes. For prediction modelling the basic information extracted from the database are the frequencies of landslide occurrences in the different mapping units and the frequencies of the mapping units in correspondence with the pixels representing the occurrences.

For some of the experiments the database was spatially subdivided in two: one for the NW region of 533,904 pixels, and one for the SE region of 857,159 pixels. The corresponding landslide occurrences were distributed as follows. The NW region contains 403 pre-1997 and 101 post-1997 occurrences. The SE



region contains 503 pre-1997 and 116 post-1997 occurrences. The spatiotemporal subdivision was used for distinct predictions and validations.

5 Tools and techniques

To add further notions that are to benefit communication between the research community, the user communities, and the decision-makers, some insight on computer tools and techniques is necessary. This is obtained by means of interactive exercises and experiments using Spatial Target Mapping software and Spatial Risk Assessment software [17], in conjunction with a general-purpose spreadsheet and optionally a geographic information system. Fabbri *et al* [18] and Chung *et al* [19] discussed earlier versions of the software. However, to avoid having to become familiar with the peculiarities of a specific system, procedural steps are documented in guided exercises that help in constructing the predictions and in developing the different strategies. This is done exploiting the particular application database described in the previous section.

The main steps in the Spatial Target Mapping analyses are described in Figure 1, as a flowchart with five groups of routines for which inputs and outputs are specified. The SPM routines are run first to obtain a prediction image for the *training area* and the related statistics. Then, the CV routines are run to perform a cross-validation to interpret the prediction image generated by SPM. If the *study area* is the same as the *training area*, the TMP routines are run next to obtain a file of all the prediction images requested by the statistics of the CV routines. If the *study area* is not equal to the *training area*, the EXT routines are run after the SPM routines to use the statistics from the *training area*, to generate a prediction image in the *study area*. Outputs of TARGET MAPPING are the *target image*, the associated *uncertainty image* and the *target-uncertainty combination image*. In addition, tables are generated with the distribution of the future (validation) occurrences in the prediction classes. A visualization of this distribution, the prediction-rate curve, describes the "goodness" of the prediction.

The prediction-rate curve is then modelled, into a probability of occurrence curve, to obtain an estimate of the probability of occurrence at each pixel so that it can be used in the risk equation, $R = H^*V^*E$, where at each pixel, R is the risk value assessed, H is the probability of occurrence of the hazardous event, V the vulnerability of the elements exposed E. The modelling of the prediction-rate curve into the probability of occurrence curve for each pixel requires a scenario in which the expected number of pixels to be affected by future hazardous occurrences is anticipated. This is done next using the Spatial Risk Assessment software.

The steps in the Spatial Risk Assessment procedure are described in Table 1. They are satisfied by modules: TL, Pr-Ta, and RISK. Module TL is run first for all available socio-economic data layers, and the output file is edited into a value and vulnerability table for each data layer. Then, depending on the characteristics of the prediction-rate table and curve obtained either from the CV or the SPM routines of STM, module Pr-Ta is run to obtain the probability of



occurrence table. Finally, the RISK module uses the socioeconomic data layers and the associated vulnerability and value tables, the *target image*, and the probability of occurrence table to generate the *risk image* and the associated statistics. If more than one socioeconomic data layer is in the input of RISK, the output *risk image* will represent the aggregated risk.



Figure 1: Procedural flow of target mapping.

Module	Description
Socio-economic	Input:
data	(1) socio-economic data layers.
preparation,TL	Output:
	(1) statistics file of all input data layers to be manually
	edited into text files, one for each layer by adding pixel
	monetary value and vulnerability for each thematic class.
Generation of	Input:
empirical	(1) the prediction-rate tables manually extracted of the
probability of	statistics from either a CV or an SPM module output of
occurrence, Pr-	STM (see fig. 1);
Та	(2) the # of pixels in study area;
	(3) the # of pixels expected/assumed to be affected by
	future hazardous events.
	<u>Output:</u>
	$\overline{(1)}$ table of empirical probability of occurrence per pixel
	for risk analysis.
Generation of	Input:
risk images,	(1) socio-economic data layers and associated value-
RISK	vulnerability tables;
	(2) target image from STM;
	(3) probability table from Pr-Ta.
	Output:
	(1) Risk image;
	(2) statistics from the RISK module.

 Table 1:
 Spatial risk assessment procedure.

6 Expected results and challenges

In developing the short course for DMs, motivation and feasibility of hazard prediction are fundamental starting points. Several questions need to be considered to justify and understand the spatial prediction modelling of natural hazards and consequent risks. A training strategy linked to those questions is applied to the analysis of a case study on landslide hazard. Computer tools and techniques are also provided.

The consequences of the strategy discussed are targeting the general problem of constructing hazard and risk maps for safer land-policy decisional approaches where transparent concern and objective quantitative support guide to alternatives of land development or of land sterilization. This contribution stems from the observation of poor or lacking communication between those who model the evolution of natural hazards and those who are to act to protect people and assets from unacceptable risks.

The challenges appear to reside not only in the unbalance between objective and subjective considerations but also between physical and human factors in decision-making. The lack of transparency is often brought up in decisions



taken on occasions of disaster situations, either before or after the occurrence of damaging events. For that reason, measures pre-disaster are easily criticized as excessive and post-disaster measures are seen as insufficient. Such common circumstances point to the need of bridging the gap between research and decision-making, therefore, the target of the training course is to contribute to the DM's comprehension, sensibility and respect of the physical relevance of hazardous processes that must be part of decisions in which socioeconomic and unavoidable political priorities are competing or even conflicting.

Acknowledgement

This contribution was partly supported by the European Commission's Project Mountain Risks: from prediction to management and governance (MRTN-CT-2006-035978, 2007-2010).

References

- [1] www.colorado.edu/hazards/research/qr/qr131/qr131.html
- [2] www.globalsecurity.org/military/facility/new-orleans-imagery.htm
- [3] google-latlong.blogspot.com/2009/04/laquila-italy-earthquakeimagery.html
- [4] Alexander D. E., Making research on geological hazard relevant to stakeholders' needs. *Quaternary International* **171-172**, pp. 186-192, 2007.
- [5] Bernstein P.L., *Against the Gods. The remarkable story of risk.* John Wiley and Sons. Ltd.: New York, 383 p., 1996.
- [6] Monmonnier M. S.,: *Cartographies of Danger: Mapping Hazards in America*. University of Chicago Press: Chicago, 363 p., 1997.
- [7] Posner R. A., *Catastrophe, Risk and Response*. Oxford University Press Inc.: New York, 334 p., 2004.
- [8] Bonachea, J., Desarrollo, aplicación y validación de procedimientos y modelos para la evaluación de amenazas, vulnerabilidad y riesgo debidos a procesos geomorfológicos. Ph.D. thesis, Universidad de Cantabria, Santander, Spain, 356 p., July 2006 (http://www.tesisenred.net/TDR-1124106-134112/index_cs.html).
- [9] Alexander D. E., A brief survey of GIS in mass-movement studies, with reflections on theory and methods. *Geomorphology*, **94**(3-4), pp. 261-267, 2008.
- [10] Fabbri A.G., Chung C.F. & Jang D.H., 2005, Spatial visualization of risk: a challenge for its management. *Proc. of IAMG '05 on GIS and Spatial Analysis*, eds., Q. Cheng & G. Bonham-Carter, Toronto, Canada, Aug. 21-26, 2005, vol. 1, pp. 346-351, 2005.
- [11] Zlatanova S. & Fabbri A. G., Geo-ICT for risk and disaster management. Geospatial Technology and the Role of Location in Science, H.J. Scholten et al, eds., GeoJournal Library, DOI 10.1007/978-90-481-2620-0_13, Springer Science+Business Media B.V.: Berlin, 2009, in press.



- [12] www.oregongeology.org/sub/news&events/archives/press-release-2009-02-18.pdf
- [13] Remondo, J., González-Díez, A., Díaz de Terán J.R. & Cendrero, A., Landslide susceptibility models utilising spatial data analysis techniques. A case study from the lower Deba valley, Guipúzcoa (Spain). *Natural Hazards*, **30**, pp. 267-279, 2003a.
- [14] Remondo, J., Bonachea, J. & Cendrero, A., A statistical approach to landslide risk modelling at basin scale: from landslide susceptibility to quantitative risk assessment. *Landslides*, **2**(4), pp. 321-328, 2005.
- [15] Fabbri A.G., Remondo J., Ballabio C., Poli S., Chung C.F. & Scholten H.J., Occurrence neighborhoods and risk assessment from landslide hazard in northern Spain. *Risk Analysis VI. Simulation and Hazard Mitigation*, C.A. Brebbia & E. Beriatos, eds., WIT Press: Southampton, Boston, pp. 29-41, 2008. Also available as WIT Transactions on Information and Communication, v. 39, www.witpress.com, ISSN 1743-3517 (on-line), doi:10.2495/RISK080041.
- [16] Remondo, J., González-Díez, A., Díaz de Terán J.R., Cendrero, A., Fabbri, A. & Chung, C.F., Validation of landslide susceptibility maps; examples and applications from a case study in northern Spain. *Natural Hazards*, **30**, pp. 437-449, 2003b.
- [17] www.spatialmodels.com
- [18] Fabbri, A.G., Chung, C.F. & Jang, D. H., A software approach to spatial predictions of natural hazards and consequent risks. *Risk Analysis IV*, C.A. Brebbia, ed., Boston, WIT Press: Southampton, pp. 289-305, 2004.
- [19] Chung, C.F., Fabbri, A. G., Jang, D. H. & Scholten, H. J., Risk assessment using spatial prediction model for natural disaster preparedness. *Geoinformation for Disaster Management*, eds. P. van Oosterom, S. Zlatanova & E.M. Fendel, Springer: Berlin, pp. 619-640, 2005. Procs. of Gi4DM, The First Symposium on Geo-information for Disaster Management, Delft, Netherlands, March 21-23, 2005.

