Grant Agreement No.: 226479

SafeLand

Living with landslide risk in Europe: Assessment, effects of global change, and risk management strategies

7th Framework Programme

Cooperation Theme 6 Environment (including climate change)

Sub-Activity 6.1.3 Natural Hazards

Deliverable D2.8

Recommended Procedures for Validating Landslide Hazard and Risk Models and Maps

Work Package 2.3 - Development of procedures for QRA at regional scale and European scale

Deliverable/Work Package Leader: UPC Revision: [3] –**Final**

July, 2011

Rev.	Deliverable Responsible	Controlled by	Date
0	UPC		
1			
2			

Note about contributors

The following organisations contributed to the work described in this deliverable:

Lead partner responsible for the deliverable:

Technical University of Catalonia (UPC)

Deliverable prepared by:
Jordi Corominas
Olga Mavrouli

Partner responsible for quality control:

Bureau des Recherches Géologiques et Minières (BRGM)
Audrey Baills

Contributors:

Technical University of Catalonia (UPC)
Cristina Baeza
Nieves Lantada
José Moya
Jordi Corominas
Olga Mavrouli

Università degli studi di Salerno (Unisa)
Leonardo Cascini
Settimio Ferlisi

Analisi e Monitoraggio del Rischio Ambientale (AMRA Scarl)
Luciano Picarelli
Antonio Santo
Giulio Zuccaro
Guido Rianna
Giuseppe Di Crescenzo

.

Eidgenössische Technische Hochsschule Zürich (ETHZ)
Harikrishna Narasimhan
Michael H. Faber

Università degli studi di Firenze (Unifi)
Veronica Tofani
Nicola Casagli

Joint Research Centre (JRC)
Miet Van Den Eeckhaut
Javier Hervás

Universitá degli Studi di Milano- Bicocca (UNIMIB)
Paolo Frattini
Federico Agliardi

Faculty of Geo-information Sciences and Earth Observation (ITC), University of Twente
Pankaj Jaiswal
Cees van Westen

Centre National de la Recherche Scientifique (CNRS)
Jean-Philippe Malet

Geological Institute of Romanian (GIR)
Raluca Maftei

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

(UPC)

1.1 NEED OF VALIDATION

An important issue in the development of a model is assessment of its "quality". Any procedure, model or map must have documented evidence that is has been appropriately verified and validated. Quantifying the confidence and predictive accuracy of model calculations provides the decision-maker with the information necessary for taking the appropriate decisions Some authors claim a model should be validated (Tsang, 1991; Power, 1993) before its application; others argue that models cannot be validated per se by any procedure but rather evaluated (Konikow and Bredehoeft, 1992; Oreskes et al., 1994; Oreskes, 1998). Aside from these conflicting conceptual and terminological views, evaluation a landslide model is always a difficult task.

According to AIAA (1998), *Verification* is the process of determining that the implementation of the model accurately represents the developer's conceptual description of the model and its solution while *Validation* is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

Physically based models for landslides can be verified because an analytical solution usually exists for most of landslide mechanisms and they can also be validated provided that experimental or field data are available. By means of verification, the errors may be identified and removed from the model. Landslide susceptibility and hazard maps instead, particularly when prepared based on heuristic and statistical methodologies, are difficult to verify as no analytical or accurate benchmark solution exist. Because of this, it is fundamental that these documents could be at least validated. This deliverable will mainly focus on the existing procedures for validation of landslide models and maps (susceptibility, hazard and risk maps).

The accuracy of each susceptibility, hazard or risk model or map must be determined. That is to say, the agreement is between the predicted and the observed results. For a landslide susceptibility map, for instance, accuracy is used to indicate whether the observed landslides have occurred in the most suceptible cells and if a correct distinction between potentially landslide free and landslide prone areas is made. As Frattini et al (2009) have stated, when a landslide susceptibility model is applied in practice, the classification of land according to susceptibility results in economical consequences. For instance, terrain that is classified as stable can be used without restrictions, increasing its economical value, whereas unstable terrain is restricted in use, and is consequently reduced in value. The misclassification of terrain in a model may come from different sources such as the quality and resolution of the input data; the size of the study area; or the adequacy of the approaches or models in which the susceptibility analyses are based on. Therefore, model evaluation should consider several criteria (David and Goodrich, 1990): (a) data reliability, resolution and completeness; (b) adequacy in describing the system; (c) robustness (response to small changes in the input data); and (d) accuracy of the landslide models and maps.

This deliverable is organized following the above mentioned criteria.

1.2 GLOSSARY

Accuracy: The difference between the true value and an estimated value.

Adequacy: The decision that the model fidelity is sufficient for the intended use

Calibration: Process of adjusting numerical or physical modeling parameters in the computational model for the purpose of improving agreement with experimental or field data.

Conceptual model: Collection of assumptions, algorithms, relationships, and data that describe the reality of interest from which the mathematical model and validation experiment can be constructed

Confidence: Probability that a numerical estimate will lie within a specified range

Model: A representation of a physical system or process intended to enhance our ability to understand, predict, or control its behavior

Prediction: Use of a model to foretell the state of a physical system under conditions for which the model has not been validated

Reliability: The rate of success of an algorithm or system

Robustness: The degree of stability of an algorithm under variations in more of its inputs parameters. A robust model or procedure is that consistently producing high accuracy or reliability in their results.

Uncertainty: A potential deficiency in any phase or activity of the modeling or experimentation process that is due to inherent variability (irreducible uncertainty) or lack of knowledge (reducible uncertainty).

Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

Verification: The process of determining that the implementation of the model accurately represents the developer's conceptual description of the model and its solution.

2 DATA RELIABILITY, RESOLUTION AND COMPLETENESS

2.1 SOURCES OF UNCERTAINTY IN LANDSLIDE SUSCEPTIBILITY, HAZARD AND RISK ASSESSMENT

(ETHZ)

Engineering risk assessment is generally based on a Bayesian interpretation of probabilities (Faber and Stewart, 2003). Within this framework, it is useful to distinguish two fundamentally different types of uncertainties, namely epistemic and aleatory uncertainties. This distinction has been considered for risk assessment of technical systems, e.g., Apostolakis (1990) or Helton and Burmaster (1996), and increasingly for natural hazards, e.g., Hall (2003), Apel et al. (2004) or Straub and Der Kiureghian (2007), but has been discussed also for general geological applications by Mann (1993). Aleatory uncertainties are interpreted as random uncertainties, which, for a given model, are naturally inherent to the considered process; epistemic uncertainties are related to our incomplete knowledge of the process, often because of limited data and can be characterised in the form of model uncertainties and statistical uncertainties.

The absolute and relative magnitudes of aleatory and epistemic uncertainty are markedly case-specific. This differentiation in uncertainties is introduced for the purpose of setting focus on how uncertainty may be reduced, rather than calling for a differentiated treatment in the risk assessment and decision analysis process. The distinction is relevant because aleatory uncertainty cannot be reduced for a given model. In contrast, epistemic uncertainty can be reduced, for instance, by collecting additional information. For this reason, a clear identification of the epistemic uncertainties in the analysis is crucial, as these may be reduced at a later time. Furthermore, neglecting epistemic uncertainty can lead to strong underestimation of the risk, see Coles et al. (2003) for an example.

As with most natural hazards, the uncertainty related to the occurrence of the hazard is generally large for landslide hazards. In the literature, this uncertainty is generally represented by an exceedance frequency and various methods are proposed for identifying the exceedance frequency at a specific site. These include:

i) the analysis of historical datasets, e.g., Hungr et al. (1999) or Dussauge-Peisser (2002),
ii) empirical models which describe hazard as a function of different indicators (observable parameters) such as topography and geology, e.g., Budetta (2004) or Baillifard et al. (2004),
iii) phenomenological (mechanical) models, e.g., Jimenez-Rodriguez et al. (2006) or Duzgun et al. (2003), and
iv) expert opinion, e.g., Schubert et al. (2005).

All these methods are useful in a particular context. While methods i) and ii) are generally more appropriate for the analysis of larger areas with less accuracy, iii) and iv) are more suited for the detailed analysis of a specific site.

An example involving the modelling of rockfall hazard is considered. Rockfall is generally considered an inherently uncertain process, i.e., it is not possible to deterministically predict the time and the extent of the next event. However, it is possible to describe rockfall using a

probabilistic model, describing the frequency $H_V(v)$ with which a rock of a certain volume V or larger is detached. Because the assessment of rockfall is based on limited data and simplified models, the probabilistic model is itself subject to uncertainty itself; this can be represented by modelling the parameters of $H_V(v)$ as random variables. In this case, we write $H_V(v|\boldsymbol{\theta})$ to indicate that the model is defined conditional on the values of its parameters $\boldsymbol{\theta}$. This epistemic uncertainty on $\boldsymbol{\theta}$ can be depicted by credible intervals (which can be considered as the Bayesian equivalent of confidence intervals) on the exceedance frequency curve as demonstrated in Figure 2.1.1.

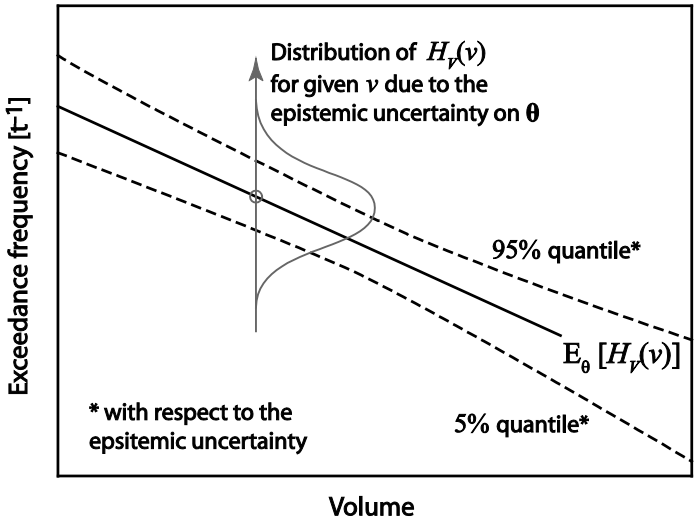


Figure 2.1.1 Exceedance frequency – illustrating the difference between epistemic and aleatory uncertainty.

It is mathematically convenient to express the exceedance frequency in a parametric format. Traditionally, a power law has been applied to describe the relation between rock volume V and the exceedance frequency:

$$H_V\left(v|\boldsymbol{\theta}\right) = av^{-b} \quad (2.1.1)$$

The statistical parameters of the model characterising the shape of the exceedance frequency curve are $\boldsymbol{\theta} = [a,b]^{\mathrm{T}}$. The epistemic uncertainty is included in the analysis by modelling $\boldsymbol{\theta}$ as a random vector. Using the probability density function of $\boldsymbol{\theta}$, $f_{\boldsymbol{\Theta}}\left(\boldsymbol{\theta}\right)$, the unconditional exceedance frequency is computed as:

$$H_V\left(v\right) = \int_{\boldsymbol{\Theta}} f_{\boldsymbol{\Theta}}\left(\boldsymbol{\theta}\right) H_V\left(v|\boldsymbol{\theta}\right) d\boldsymbol{\theta} \quad (2.1.2)$$

There are various sources for epistemic uncertainties in large scale models, preventing an exact prediction of the exceedance frequency for a particular site. A brief description of these is provided below.

Statistical uncertainty

The parameters of the large scale models are derived empirically from data sets. Because of the limited size of these data sets, the estimated parameters are subject to statistical uncertainty.

Measurement uncertainty

Measurements and recordings of the geological properties are typically subject to uncertainty and observations of historical events are often incomplete and biased and must rely on local experts. As an example, rocks on a road will generally be reported and documented, but those that missed the road may often not be. Measurement uncertainty also results from derives from equipment, operator/procedural and random measurement effects.

Model uncertainty

Extrapolation of the statistical models to areas other than those for which observations are available leads to additional uncertainty as the geological and topographical characteristics will be different for these areas. GIS-based models will take into account some of these parameters, but the omitted parameters will lead to an uncertainty in the model predictions. Uncertainty also occurs due to the approximations and simplifications inherent in empirical, semi-empirical, experimental or theoretical models used to relate measured quantities to non-measurable numerical parameters used in estimation. Finally, although the power-law is, for example, commonly assumed to express exceedance frequency in the case of rockfall hazard, it has not been justified by phenomenological considerations. Thus, it is not ensured that the parametrical model accurately represents the actual behaviour.

Spatial variability

The frequency of hazard events varies in space. The observations represent an average over an area and the resulting parameter values, therefore, do not reflect the variations from the average.

Temporal variability

The frequency of hazard events varies in time. When working with annual frequencies, the seasonal changes do not affect the analysis, but the frequency may change over the years or may be dependent on extreme events (e.g., earthquakes). However, in certain instances, e.g., when temporal closure of the road is considered as a risk reduction measure, seasonal variations must be explicitly addressed by the analysis.

How can these uncertainties be quantified? Statistical uncertainty can be quantified by using standard statistical methods such as Bayesian analysis, see, e.g., Coles (2001). Measurement uncertainty can generally be estimated when the data collection method is known. Unfortunately, no simple analytical method is available for estimating model uncertainties. A solution is to rely on expert opinion, i.e., to ask experts about their confidence in the models. It is also possible to compare the model with observations which have not been used in the calibration of the model (model validation) or to compare different models. Furthermore, it is possible to include additional parameters in the formulation of the exceedance frequency, expressed in Equation 2.1. The model uncertainties are then reduced while the statistical uncertainties increase, but the latter can then be estimated analytically. Coles et al. (2003) demonstrate this for the analysis of rainfall data. The spatial and temporal variability can be analysed quantitatively, if data is available in sufficiently small scale; a data-set showing the spatial distribution of rockfall events is presented in Dussauge-Peisser et al. (2002). Spatial variability can be described by the spatial correlation of the relevant characteristics. In most practical cases, however, a simplified approach is favourable, whereby smaller areas are determined within which the spatial variability can be neglected. Temporal (typically seasonal) variability can be described by time-dependent parameters in the exceedance frequency model, corresponding to the assumption of the hazard event (e.g. rockfall) following an inhomogeneous Poisson process.

2.2 QUALITY AND RELIABILITY OF THE COLLECTED DATA

(UNIFI and JRC)

Following the definition proposed within Safeland project (Deliverable D8.1, SafeLand project handbook), hazard can be defined as a condition with the potential for causing an undesirable consequence. The description of landslide hazard should include the location, volume (or area), classification and velocity of the potential landslides and any resultant detached material, and the probability of their occurrence within a given period of time. This, essentially, means that landslide hazard assessment procedures must take into account both space and time prediction (Catani et al., 2005).

Predictions based solely on spatial probability of occurrence are, however, very common, due to the fact that they are relatively easier to carry out. In such cases the term “landslide susceptibility” should be considered more appropriate (Dai et al. 2002).

According to Dai et al. (2002) the factors which determine the probability of landsliding for a particular slope can be grouped into two categories: (1) preparatory variables which make the slope susceptible to failure without triggering it, such as geology, slope gradient and aspect, elevation, soil geotechnical properties, vegetation cover and long term drainage patterns and weathering; and (2) triggering variables such as intense rainfall, rapid snowmelt, water-level change, volcanic eruptions, earthquake shaking and human activities (Wieczorek, 1996).

If triggering variables are not considered in the analysis the term susceptibility must be used to define the likelihood of occurrence of a landslide.

As stated by Metternicht et al. (2005) remote sensing has been mostly used for the detection and identification of diagnostic features of preparatory variables, and to a lesser extent, for the detection

of potentially triggering factors as shown by studies of Kniveton et al. (2000), Buchroithner (2002), Huggel et al. (2002), Kaab et al. (2003).

Susceptibility analysis can be based on a number of techniques and data. According to Soeters and van Westen (1996), van Westen et al. (1997) and Aleotti and Chowdhury (1999), they can be divided into inventory, heuristic, deterministic and statistical approaches.

The most straightforward method to any study of landslide hazard is the compilation of a landslide inventory (Dai et al., 2002). The output of a landslide inventory provides the spatial distribution of mass movements, represented as polygons or points (Wieczorek, 1983) and for this reason it can be used as elementary form of hazard map, though they fail to identify areas that may be susceptible to landsliding unless landslides have already occurred (Dai et al., 2002).

In the heuristic approach expert opinions are used to estimate landslide potential from data on preparatory factors. These models are based on the assumption that the relationships between the landslide susceptibility and preparatory factors are known and specified in the models (Dai et al., 2002).

Deterministic methods, in which the relative probability of spatial occurrence of a mass movement is usually derived from the computation of the factor of safety, are probably the most objective means of hazard assessment (e.g. Montgomery and Dietrich, 1994; Wu and Sidle, 1995). Unfortunately, due to the large spatial variability of the mechanical, hydrological and geometrical parameters involved in the equations (e.g. Burton et al. 1998), their application is very difficult and therefore at the basin scale while numerous studies are available at the slope scale (Anderson and Lloyd, 1991; Crosta, 1998; Iverson, 2000; Crosta and Dal Negro, 2001, Tofani et al., 2006).

The uncertainties in the definition of strength parameters can be at least partially overcome by the application of probabilistic methods (for a more in-depth description of such techniques refer to Baecher and Christian, 2003).

The methods mentioned so far do not result in real hazard maps as defined by Varnes (1984). Assessing the probability of occurrence at a certain location within a certain time period is only possible when the temporal prediction is defined.

Canuti and Casagli (1996) propose three different approaches for landslide frequency analysis, i) Analysis of the time series of the landslide events, ii) Analysis of the time series of triggering factors, iii) Monitoring.

The first approach is one of the most efficient tool for the definition of the recurrence of landslides. It is based on the historical analysis of the past reactivations of the phenomena through documents, reports and population depositions. Some examples of the of this type of approach are reported in Skempton et al. (1989), Lee et al. (1991), and Del Prete et al. (1992).

The second approach investigating the relationship between the occurrence of landslides and the frequency of triggering factors has been widely exploited. Two different approaches have been presented in the literature to explain the relationship between rainfall and slope failures: statistical and deterministic approach.

The statistical approach consists on the creation of rainfall thresholds on empirical basis. Some examples are reported in Caine (1980), Crozier (1986), Wieczorek (1987), Crosta and Frattini (2001), Guzzetti et al. (2008). The deterministic approach consists on the combination of

hydrological models with slope stability analyses (Montgomery and Dietrich, 1994; Wilson and Wieczorek, 1995; Crosta, 1998; Terlien, 1998). Some application s of the distributed stability models at the basin scale are reported in Borga et al. (1998), Aleotti et al. (2003) and Crosta et al. (2003).

Temporal prediction carried out by means of monitoring systems gives the more detailed and reliable information; anyway it is the most expensive and complex method especially if carried out by means of ground-based instrumentation. Remote sensing, and in particular A- DInSAR techniques, overcome these problems giving the availability of long and complete time series of deformations.

2.2.1 Preparatory factors

Landslide spatial prediction, or susceptibility assessment, can benefit of both optical and radar remote sensing for the identification of preparatory factors as well as landslide inventory maps (Metternicht et al., 2005).

To determine location and extent of landslide (especially after of large triggering events) aerial photographs are being progressively replaced with high and very-high resolution satellite images. The increasingly higher spatial and temporal resolution of optical satellite observations enables (i) more detailed and reliable identification of affected areas, (ii) an immediate response minimizing the risk of omission (due to landslide traces fading away with time), and (iii) repeated observations potentially leading to multi-temporal inventories, which can be easier related to specific events (e.g. Saba et al., 2010). Multi-temporal inventories are still available for relatively few thoroughly studied areas (Guzzetti et al., 2009; Jaiswal and van Westen, 2009; Witt et al., 2010), whereas at the same time a lot of archived and continuously acquired imagery still remains unexploited to create multi-temporal inventories for other areas. Even if time-series from field-surveys are readily available, very high resolution (VHR) optical images may reveal a considerable additional number of landslides for a given area (Fiorucci et al., in press) and should at least be considered for visual interpretation.

There is a large number of studies which proposed, applied and compared automated (both pixel and object-based) techniques for landslide mapping with optical data (Hervás et al., 2003; Cheng et al., 2004; Nichol and Wong, 2005; Whitworth et al., 2005; Barlow et al., 2006; Borghuis et al., 2007; Danneels et al., 2007; Rau et al., 2007; Park and Chi, 2008; Joyce et al., 2009; Marcelino et al., 2009; Yang and Chen, 2010; Martha et al., 2010; Di et al., 2010; Gao and Maro, 2010; Lu et al., 2011;). Martha et al. (2010) used optical data in combination with topographic information derived from a 10 m resolution DEM for the creation of a landslide inventory in an object-based approach. In contrast to the relatively high number of studies on (semi-)automatic landslide identification, on little work has been dedicated to the use of automatically mapped inventories as input for hazard or susceptibility assessment. Exceptions are for example the studies presented by Park and Chi (2008) and Barlow et al. (2009), but in the first study still (presumably considerable) manual editing was involved to refine the automatically mapped landslides. Not all studies on automated mapping included proper accuracy assessments, but those who did showed mapping accuracies between approximately 30% (e.g. Marcelino et al., 2009) and 90% (e.g. Yang and Chen, 2010) depending

largely on the complexity of the ground conditions but also on the employed techniques and image types.

Airborne LiDAR techniques show particular strength for the mapping of old landslides under forest (e.g. Van Den Eeckhaut et al., 2007) but can also be used to support the mapping of newly triggered shallow landslides (Ardizzone et al., 2007; Lu et al., 2011). There seems to be a general agreement that LiDAR based mapping yields more accurate and complete inventories than field surveys alone (Ardizzone et al., 2007; Schulz, 2007; Van Den Eeckhaut et al., 2007). Resulting inventories have been employed for efficient susceptibility models whereby the LiDAR derived terrain model provides also valuable input to extract influential topographic variables (Van Den Eeckhaut et al., 2009; Van Den Eeckhaut et al., 2006). The acquisition of airborne LiDAR datasets is still rather costly and for a frequent updating of landslide inventories other techniques may remain more feasible. With an eye towards analysis of such datasets over larger areas some automated mapping techniques (McKean and Roering, 2004; Glenn et al., 2006; Booth et al., 2009) with partially promising accuracies (Booth et al., 2009) have been proposed. It appears desirable to compare susceptibility models based on automated techniques with those based on manual mappings to assess the sensitivity of the spatial susceptibility before further inferences on the hazard are drawn.

The application of the interferometric techniques to radar images is a powerful tool for landslide detection and mapping at large scale. In particular the A-DinSAR technique can contribute to the creation of landslide inventory maps which can be used for susceptibility mapping over large scale. There are a few of examples in literature of the use of landslide inventory maps derived from radar images for landslide hazard assessment (Singhroy et al., 1998; Catani et al. 2005, Farina et al., 2006). It's worth noticing that the realization of landslide inventory maps with InSAR technology benefits from the integration with optical imagery, geological and topographic information (Singhroy, 2008).

Radar interferometry, especially through the PS technique has provided a tool for landslide back-monitoring. The availability of deformations time series provided for long period by the radar satellites as ERS and ENVISAT can be used for the definition of the recurrence of landslides. At the same time the comparison of the time series with triggering factors data can define the causes and the deformation thresholds.

Optical imagery can give a great a contribution in the landslide susceptibility assessment especially in the definition of the preparatory factors. In particular the use of optical images is quite common when the analysis is carried out over large areas since these data can be easily collected.

As reported in Metternicht et al. (2005) optical imagery, as well as aerial photographs, have provided the main contribution for the mapping of landslide related factors.

Optical imagery can be used for defining the land cover, tectonics lineaments and for the set up of a Digital Elevation Model (DEM) which can be the inputs data for the heuristic and statistical susceptibility approach.

DEMs are very important in landslide susceptibility analysis since many predisposing factors can be derived, such as slope gradient, slope aspects and plan profile curvatures. SPOT and ASTER scenes can produce DEMs with spatial resolution of 15 m and vertical accuracy of 20 m (Reinartz et al. 2004; Fujisada et al., 2005; Toutin, 2006) and are particularly affordable while the use of high resolution images such as Quickbird and IKONOS in landslide studies is hampered by the high

acquisition costs (van Westen et al., 2008). From the DEMs many derivate maps can be derived such as slope gradient, slope aspects and plan profile curvature.

Recently also radar images have been employed for the definition DEMs. An example is the DEM derived from NASA Shuttle Radar Topographic Mission (SRTM): it obtained a new class of elevation data on a near-global scale to generate the most complete high-resolution digital topographic database of the Earth (Rabus et al., 2003; Rodriguez et al., 2003). SRTM consisted of a specially modified radar system that flew onboard the Space Shuttle Endeavour during an 11-day mission in February of 2000. . It is currently distributed free of charge by USGS and is available for download from the National Map Seamless Data Distribution System, or the USGS ftp site. The SRTM data is available at 3 arc second, which means approximately 90m of ground resolution. A 1 arc second data product, 30m resolution, was also produced, but is not available for all countries. 30m SRTM data have been used to get topographic factors (slope, aspect etc) as reported in Hong et al. (2008). The DEM derived is used for the definition of the global landslide susceptibility by means of a statistical analysis.

Land use and land cover and their associated changes can have important impact on landslide activity. The effect of vegetation on slope stability can be classified either hydrogical and mechanical (van Westen et al. 2008). Land cover maps are made on a routine basis from medium resolution satellite imagery such as LANDSAT, SPOT, ASTER, IRS1-D (van Westen et al., 2008); the application of spectral indices to optical images to retrieve land cover parameters is widely used at regional scale. NDVI (Normalized Difference Vegetation Index) is derived from the reflectance of red image and near infrared image and is widely applied using Landsat and SPOT satellites: as landslides, especially shallow and rapid slope movements, resulted in non-vegetated areas, vegetation indices could give a contribution to a computer-assisted approach to interpret landslides and making a landslide map (Liu et al. 2002).

Apart from topography and land cover / land use also geology is often a key factor for the occurrence of landslides (Pachauri and Pant, 1992; Gerrard, 1994; Sarkar and Kanungo, 2004), and combined with other information layers (in particular lineaments, faults, land use, soil, ...) can be used to map the landslide susceptibility on a regional scale. The required scale for such usage is about 1:25000. We can assume that for a more local analysis (mapping monitoring an indiviual landslide), a more detailed map (1:5000) will be required. The spectral analysis of the solar radiation backscattered by the ground surface (reflectance spectrometry) could be an important source of remote sensing information about the chemical and mineralogical characteristics of the materials of the ground surface. In fact, the reflectance spectra (visible and near infrared) of the rocks have specific absorption bands that allow the identification of the constituent minerals, by analyzing the position, the shape and intensity of this bands. The analysis of hyper-spectral and multi-spectral imagery is therefore a powerful tool for identifying mineralogy from space.

Apart from lithological information also structural information is very important for landslide hazard assessment. For the landslide susceptibility mapping, lineaments associated with fractures, discontinuities and shear zones are useful information that could be combined to other data. In most cases such information will be used to refine the lithological information and fault positions. Sarkar and Kanungo (2004) for example propose a more direct use of this information by including the density of lineaments as indicator in the susceptibility computation. An extensive review of the use of lineament maps in landslide susceptibility and hazard mapping was recently undertaken by Ramli et al. (2010). They concluded that lineament maps are a vital part of landslide hazard assessment but

also demonstrated that none of the reviewed studies used automated mapping techniques. The authors pointed out that manual mapping yield generally high uncertainties and only few studies apply instruments to evaluate such discrepancies of different lineament maps.

Soil information required for landslide susceptibility assessment includes soil type, including geotechnical and hydrological properties, and soil depth. These data are desired for statistical modelling, but essential for deterministic modelling. Pedologic soil maps classify the soils based on the upper soil horizons (van Westen et al., 2008). Geotechnical characteristics of the soils can be derived from in situ and laboratory measurements. A detailed characterization of the soil often hampers the application of determinstic models over wide areas. Several authors have modelled the variation of soil thickness over wide areas, such as Terlien et al., (1995), Dietrich et al. (1995), Segoni et al, (2009) and Catani et al. (2009).

2.2.2 Triggering factors

The temporal prediction of landslides is related to analysis of the time series of the landslide events or the analysis of the time series of triggering factors or the monitoring of landslides .

One of the most intriguing applications currently being investigated regarding the use of remote sensing is the temporal prediction of shallow landslides. Drawing on recent advances of satellite remote sensing technology, experimental landslide prediction models are developed to identify the timing for landslides induced by heavy rainfall (Hong et al., 2006; Adler et al., 2000).

Satellites have provided global estimates of precipitation over various temporal and spatial scales since the 1970s (Kidd, 2001). A long history of development in the estimation of precipitation from space has culminated in sophisticated satellite instruments and techniques to combine information from multiple satellites to produce long-term products useful for climate monitoring (Adler et al., 2003). In November 1997, the Tropical Rainfall Measuring Mission (TRMM) was launched with the primary objective of making accurate measurements of rainfall and latent heating from space (Kummerow et al., 2000). In particular the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA, Huffman et al., 2006) has been used for landslide prediction at global scale (Hong et al., 2006, Hong et al., 2007).

SAR interferometry is used to map and monitor active faults deformation since the 90's for co-seismic displacements (Massonnet et al., 1993). However, on the one hand, fault displacement during high magnitude earthquakes (i.e.most of the earthquakes mapped by conventional differential InSAR) is generally too high. This causes loss of signal coherence and enables mapping of the displacement(De Michele et al., 2010; Raucoules et al., 2010). Sometimes, the lack of signal coherence in the InSAR signal of the displacement field is an indirect evidence of the seismogenic fault and can be used to detect un-mapped active faults or blind seismogenic structures (Talebian et al., 2004). On the other hand, slow (cm/yr) aseismic movements on faults are difficult to observe on single interferograms as the displacement values are lower than the displacement detection capabilities of the technique, taking into account the presence of different sources of noise in the interferometric phase.

Recently the prediction of shallow landslides has benefited also by the use of Numerical Weather Prediction Models (NWP). The numerical models used for the weather forecasts are a simplified

schematic representation of physical reality, described through a set of equations that simulate the behaviour of nature. Two different approaches exist to the NWP problems:

- Deterministic approach: it postulates that, at least over a certain time period, the laws of physics, as applied to the atmosphere, can be solved (integrated forward in time) to find the forecast fields given initial data describing the current conditions.
- Probabilistic approach: it is based on the idea of starting a set of forecast integrations from slightly different initial conditions, reflecting the range of uncertainty in the estimated initial state. This ensemble approach allows a probability to be assigned to the likelihood of rainfall .

With regard to the time period covered, different types of numerical weather forecasts exist:

- Short-range forecasts are made for a time period up to 48 hours. Due to the chaotic nature of the atmosphere in this range the forecasts are generally more accurate than the other types of forecasts.
- Medium-range forecasts are for a period extending from about three days to seven days in advance.
- Long-range forecasts are for a period greater than seven days in advance but there are no absolute limits to the period.

In the framework of Safeland project the regional model COSMO-LM has been used for the prediction of shallow landslides at the regional and local scale (Mercogliano et al, 2010).

The COSMO-LM model provides precipitation forecasts taking also into account orographic effects, an important factor in triggering mass movements. It contains parameterization providing reliable estimates of convective cell dynamics, which are the cause of the most frequent extreme rainfall events. The COSMO-LM model has been used in its 2 configurations: the one with 7km of horizontal resolution (providing 72 hour forecasts) and the other one with the higher horizontal resolution of 2.8 km (providing 24 hour forecasts).

3 ADEQUACY OF THE APPROACHES USED TO PREDICT LANDSLIDE SUSCEPTIBILITY AND HAZARD

3.1 Understanding landslide mechanisms

(UNISA)

The knowledge of the processes leading to slope instability (triggering stage) until the end of the landslide propagation stage represents a fundamental step in the analysis and zoning of landslide susceptibility and hazard (Fell et al. 2008a; Corominas & Mavrouli, 2011). Restrictions associated to this knowledge derive from the method adopted for a proper characterization of the whole landslide mechanism. In this regard, the selection of the most appropriate method depends on several factors such as: availability, quality and accuracy of data; resolution of zoning; required outcomes; scale of zoning, etc. The Deliverable D2.4 (Corominas & Mavrouli, 2011) thoroughly discusses this topic and distinguishes the methods among heuristic, statistical and deterministic (Soeters and vanWesten, 1996).

In heuristic methods the expert opinion of the person carrying out the zoning is used to assess the susceptibility and hazard. These methods combine the mapping of the landslides and their geomorphologic setting as the main input factors for the analysis purposes.

The adoption of the heuristic methods seems to lead to results conditioned by the apparent subjectivity of the person doing the study. Anyway, in several cases, a proper use of the heuristic methods can lead to conservative estimations. This is the case of the travel distance prediction carried out on the basis of the location of ancient landslide deposits (or talus cones), in turn obtained via geomorphological criteria and – possibly – corroborated by information achieved through historical documents. In effect, the derived travel distance corresponds to most extreme observed events that are attained by only a minority of the landslide cases (Hungr et al., 2005). The conservative (but not unrealistic) nature of this approach (Figure 3.1.1) may be very appropriate in preliminary assessment of landslide susceptibility and hazard (Cascini, 2005). On the other hand, it can be observed that the heuristic methods, which are often empirical in nature, could have a significant uncertainty which should be allowed for in developing the susceptibility maps for landslides which will travel beyond the source landslide. An example is provided by the rockfalls whose reach angle calculation needs the collection of incident data, so reducing the errors related to the identification of both the source location and end points of the travel paths.

As far as statistical methods are concerned, they are based on the observed relationships between each factor and the past distribution of landslides. As Fell et al. (2008b) claim, this approach usually involves the mapping of the existing landslides, the mapping of a set of factors that are supposed be directly or indirectly linked to the stability of the slopes, and the establishment of the statistical relationships between these causal factors and the instability process. Hence susceptibility or hazard zoning is conducted in a largely objective manner whereby factors and their inter-relationships are evaluated on a statistical basis (Brabb et al., 1972; Carrara, 1983; Carrara et al., 1995; Atkinson and Massari, 1998; Dai and Lee, 2001; Ayalew and Yamagishi, 2005). Limitations with such methods result from data quality such as errors in mapping, incomplete inventory and poor resolution of some data sets as the models are essentially data trained. In addition, the results of such models are

able to predict the occurrence of landslides whose magnitude is similar to that of the (already occurred) phenomena considered in the calibration and validation steps. Finally, the obtained results are not readily transferable from region to region.

Deterministic methods are useful for the analysis of both landslide triggering and propagation stage. In the first case, they apply classical slope stability theory and principles such as infinite slope, limit equilibrium and finite element techniques. These models require soil parameter inputs such as soil thickness, soil strength, groundwater pressures, slope geometry etc. (Cascini et al., 2009). The safety factor may be established in a GIS in pixel cells and the results referred to susceptibility depending on the calculated factor of safety values. Given the complexity of geotechnical conditions in slopes these methods are unreliable unless calibrated via back-analyses of occurred phenomena.

As far as the propagation stage is concerned, deterministic "analytical" methods allow modelling a moving landslide using the physical rules of solid and fluid dynamics. Details on the different models proposed in the scientific literature are provided by the SafeLand Deliverable D1.7 (Pastor, 2010). In particular, the depth-integrated models allow the simulation of run-out processes taking into account the existence of obstructions and pathways such as buildings, roads, channels and bridges. However, owing to the extreme complexity of the processes to be simulated, the parameters comparing in the governing laws are usually obtained by back-analyses of occurred events or by calibration.

In general, where statistical methods are being used heuristic methods will also be used, and for deterministic methods, statistical and heuristic will also be used. It should be noted that the more advanced the characterization method, then the larger scale the mapping and level of detail of information and understanding of slope processes is required (van Westen, 1994; 2004).

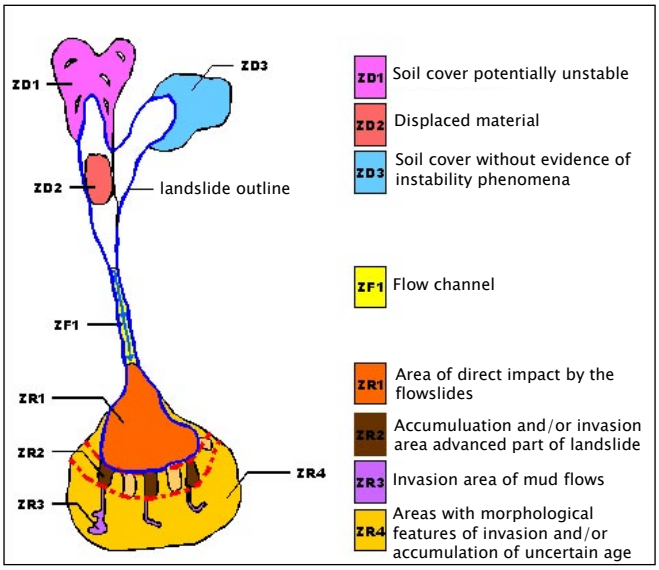


Figure 3.1.1 Area taken into account for the analysis of flowslide phenomena in Campania region, southern Italy (Cascini, 2005)

On the basis of the methods adopted for the analysis of landslide susceptibility and hazard, three different zoning levels (preliminary, intermediate and advanced) can be obtained (Fell et al., 2008a). In this regard, it is important to relate the level of zoning to the required outcomes and to the scale of zoning and, in turn, to relate this to input data and zoning procedures. This means that zoning methods must be selected taking into account level and scale of zoning, quality and accuracy of input data, applicability of zoning procedures. For instance, a detailed susceptibility zoning at large and detailed scales necessarily needs deterministic methods based on geotechnical data; likewise, a detailed hazard map cannot be developed without a dataset allowing the assessment of landslide frequency via statistical methods.

The levels of zoning and zoning map scales are thoroughly discussed within the Deliverable D2.4 (Corominas & Mavrouli, 2011). At small scale, considering that only heuristic methods based on geological data can be used, only a preliminary zoning level can be pursued and obtained. At medium scale, where statistical procedures can be used, two zoning levels may be defined. At large and detailed scales, three zoning levels are possible, respectively based on heuristic, statistical and deterministic methods.

Finally, it is worth to stress that the type, level and scale of zoning also depend on the complexity of the landslide features, the homogeneity of the terrain, the spatial variability of the important causal factors, the geotechnical parameters and the amount of available data and expertise. The existence of related uncertainties should be borne in mind when carrying the information provided by susceptibility zoning maps forward into preparing hazard and risk zoning (Fell et al., 2008b).

3.2 SELECTION OF INDEPENDENT VARIABLES/DATA LAYERS/INPUT DATA USED IN HEURISTIC AND STATISTICAL MODELS

(JRC and GIR)

3.2.1 Data layers for heuristic and statistical landslide susceptibility and hazard modeling at different scales

Tables 5.1 and 5.3 of D2.4 ‘Guidelines for landslide susceptibility, hazard and risk assessment and zoning’ give a schematic overview of the main data layers required for landslide susceptibility, hazard and risk assessment at different scales of analysis (Van Westen el al., 2008). Given its relevance for this section the latter table is also included here (Table 3.2.1). The more detailed the analysis the higher the number of variables that are advised to be included. The list of environmental factors indicated in the table is not exhaustive. Moreover, it is not possible to give a prescribed uniform list of factors to be included, and some expertise is required to select the appropriate set of variables for a specific study. Apart from the scale of the analysis, also for example the type and failure mechanism of the landslides, the local environmental characteristics and the model used determine the set of variables to include in the modeling. Unfortunately most often data availability is the decisive factor. A good analysis of the data sets available at the required scale, and of the datasets that are missing, in combination with appropriate management of the available financial resources can make landslide studies more efficient. The additional acquisition of a fundamental although rather expensive data layer or updating of the landslide

inventory, for example, might be better than the acquisition of a higher number of cheaper data layers.

Data layers influencing landslide occurrence relate to morphometry, geology, soil types, hydrology, geomorphology and land use. Some of the layers are more or less static, while others are dynamic and need regular updating. Geomorphology, geology and soil type are examples of static variables. Information on land use and land cover, population, meteorology and soil hydrology needs updating. The same is true for the landslide inventory. It is important that all the variables included are representative for the time frame for which the modeling is envisaged. We will come back to this further in this section. The update frequency of dynamic data ranges from hours or days (e.g. meteorological data) to months or years (e.g. population data). Van Westen et al. (2008) suggest an update frequency from 1 to 10 years for land use and elements at risk data, depending on the dynamics of the area.

Table 3.2.1 Overview of environmental factors, and their relevance for landslide susceptibility and hazard assessment. Scale of analysis: N=National, R=Regional, L=Local and S=Site Specific. (H= highly applicable, M= moderately applicable, and L= less applicable; Van Westen e al., 2008)

Group	Data layer and types	Relevance for landslide susceptibility and hazard assessment	Scales of analysis			
			N	R	L	S
Digital Elevation Models	Slope gradient	Most important factor in gravitational movements	L	H	H	H
	Slope direction	Might reflect differences in soil moisture and vegetation	H	H	H	H
	Slope length, shape, curvature	Indicator for slope hydrology	M	H	H	H
	Flow direction	Used in slope hydrological modeling	L	M	H	H
	Flow accumulation	Used in slope hydrological modeling	L	M	H	H
	Internal relief	In small scale assessment as indicator for type of terrain.	H	M	L	L
	Drainage density	In small scale assessment as indicator for type of terrain.	H	M	L	L
Geology	Rock types	Based on engineering properties of rock types	H	H	H	H
	Weathering	Depth of profile is an important factor	L	M	H	H
	Discontinuities	Discontinuity sets and characteristics	L	M	H	H
	Structural aspects	Geological structure in relation with slope angle/direction	H	H	H	H
	Faults	Distance from active faults or width of fault zones	H	H	H	H
Soils	Soil types	Engineering soils with genetic or geotechnical properties	M	H	H	H
	Soil depth	Soil depth based on boreholes, geophysics and outcrops	L	M	H	H
	Geotechnical prop.	Grainsize, cohesion, friction angle, bulk density	L	M	H	H
	Hydrological prop.	Pore volume, saturated conductivity, PF curve	L	M	H	H
Hydrology	Water table	Spatial and temporal depth to ground water table	L	L	M	H
	Soil moisture	Spatial and temporal soil moisture content	L	L	M	H
	Hydrologic components	Interception, evapotranspiration, throughfall, overland flow, infiltration, percolation etc.	M	H	H	H
	Stream network	Buffer zones around streams	H	H	H	L
Geomorphology	Physiographic units	First subdivision of the terrain in zones related to overall physiographic setting	H	M	L	L

	Terrain Mapping Units	Homogeneous units of lithology, morphography and processes	H	M	L	L
	Geomorphology	Genetic classification of main landform building processes	H	H	M	L
	Slope facets	Geomorphological subdivision of terrain in slope facets	H	H	H	L
Land use	Land use map	Type of land use/ land cover	H	H	H	H
	Land use changes	Temporal varying land use/ land cover	M	H	H	H
	Vegetation	Type, canopy cover, rooting depth, root cohesion, weight	L	M	H	H
	Roads	Buffers around roads in sloping areas with road cuts	M	H	H	H
	Buildings	Slope cuts made for building construction	M	H	H	H

As topography is one of the major factors in landslide hazard analysis, digital elevation models (DEMs) and their derivatives (slope gradient, slope direction, contributing area, plan curvature, profile curvature, slope length, surface roughness indicators) play a major role in heuristic and statistical models. The suitability of these maps for the analysis is greatly depending on the technique used to produce the DEM and on the resolution. In fact, the scale of the landslide susceptibility and hazard analysis should depend on the resolution of the data layers although many researchers perform what Schmidt and Andrew (2005) called the "let's take a DEM … approach" without paying attention to the quality.

Also the geological map can provide one or more data layers (e.g. chronostratigraphy, lithostratigraphy, faults, tectonic lineaments, tectonic units and earthquake epicentres) for heuristic and statistical landslide susceptibility and hazard assessment. In medium and small scale analysis the subdivision of geological formations into meaningful mapping units of individual rock types often poses a problem, as the intercalations of these units cannot be properly mapped at these scales. Generally also inclusion of the fault map is relevant. Previous studies have used either "distance to faults" (a continuous variable) or "buffers around faults" (a discrete variable). In the latter case, however, the buffer width depends on the activity of the faults. Active faults are included because they are one of the indicators of ongoing seismic activity (earthquakes). Hence, rather wide buffers are appropriate. Inactive faults, on the other hand, indicate weaker zones of more fractured rocks. So buffer size should be small.

For heuristic and statistical models soil information generally consist of soil types, with eventual additional information on associated geotechnical and hydrological properties. Especially in regions with thick soil covered slopes, this data layer should be collected.

Up to now land use and land cover were generally considered as static factors. Only few studies have investigated multitemporal susceptibility or hazards using land use maps of different periods (e.g. Barredo et al., 2000). Meusburger and Alewell (2009) not only used multitemporal land use information but also multitemporal landslide inventory maps. The land use data used should represent the land use conditions that were present at the timing of landslides. Unfortunately, often this is not the case and recent land use maps are used in combination with a historic landslide inventory (i.e. containing all landslides regardless of their age that have been identified in the area). Van Den Eeckhaut et al. (2006), for example, observed that older landslides in central Belgium are currently located under forest, because the hummocky topography hampers other land use types. As for many of these landslides the timing of the landsliding and the associated land use on the site is not known, it was decided not to include land use in the susceptibility modeling.

For landslide hazard analysis, also information of triggering factors is needed. The spatial resolution at which data layers of triggering factors are available determines the scale for which they can be used. For detailed, local analyses the temporal rather than the spatial scale is important. For national and regional analyses this information is generally represented on maps obtained by interpolation of data collected in different measuring stations (precipitation, earthquake magnitude, peak ground acceleration,…). A sufficiently dense network of measuring stations is required to obtain a meaningful distribution map. Their quality will be further determined by the length of the data collection period, the applied interpolation technique and the environmental characteristics (i.e. in mountain areas rainfall characteristics will generally vary over shorter distances than in low-relief areas).

3.2.2 Procedures to assess both the suitability and relevance of the input data

It was already mentioned that the extensive list of data layers in Table 3.2.1 is not available for all regions, nor are the data layers relevant to include in all landslide susceptibility and hazard studies. Important is that often different combinations of environmental factors should be used for landslides with different failure mechanism, resulting in this way in separate landslide susceptibility maps.

Apart from the field expertise, a modeler can also use some (statistical) tests for the selection of suitable and relevant input data. In the first part of this section we focus on the problem that the inclusion of interdependent and correlated input variables in the same model should be avoided to obtain meaningful statistical (and heuristic) models.

In a second part, an overview of procedures previously used in landslide studies to obtain information on the importance of input variables to landslide occurrence is given. When constructing a model always part of the reality is lost. The objective is to find a trade-off between the prediction accuracy and the complexity of the model, or in other words to build a model that contains only the variables that significantly contribute to the presence or absence of landslides.

A. Prior to the modeling: test of conditional dependence and collinearity

Even though all variables in a set can be relevant, the combination of some of them in one model might not be recommended, because of the occurrence of correlation or interdependence among the variables.

Conditional independence asserts that the values taken of two independent input variables are independent of the values of the dependent variable. In other words, each variable should provide independent evidence for the occurrence of future landslides. Conditional dependence may cause the spatial probabilities to be either overestimated or underestimated (Bonham-Carter, 1994).

The presence of conditional dependence can be investigated by pairwise comparison of the input variables using chi-square ($\chi 2$) statistics. This is for example used in weights of evidence (Thiery et al., 2004; Van Den Eeckhaut et al., 2008). Chi-square values can be determined by employing the following equation:

$$\chi^2 = \sum_{i=1}^{i=n} \frac{(O_i - E_i)}{E_i} \quad (3.2.1)$$

where the observed frequencies (Oi) and the expected frequencies (Ei) are determined from a contingency table or cross tabulation. In this context Oi and Ei would be for example soil and lithology types, and the hypothesis tested is that soil and lithology have equal distributions (i.e. are dependent). Knowledge of the degrees of freedom (number of soil and lithology classes minus 2) then allows comparing the calculated $\chi 2$ with the value in the Chi square distribution table. If the critical value is exceeded then the two variables can be used together.

Other studies have quantified the possible correlation between variables by performing a Principal Component Analysis (PCA; Baeza and Corominas, 2001; Santacana et al., 2003; Dewitte et al., 2010). PCA provides some insights on the structure of the overall population, and knowing how the population is structured may allow the identification of variables having similar behaviour and the detection of correlation that is difficult to observe with a simple correlation matrix. It allows a better selection of the most significant variables (Dewitte et al., 2010).

Collinearity and multicollinearity, on the other hand, occurs when a variable is nearly a linear combination of other variables in the model. In the case of collinearity, affected estimates are unstable and have high standard errors. Model fitting via multivariate statistical modeling (such as logistic regression) is sensitive to collinearities among the independent variables (Hosmer and Lemeshow, 1989). Collinearity occurs in many studies. A common mistake is to put too many variables into the model. To find out which variables are nearly collinear with one another, collinearity diagnostic statistics produced by linear regression, i.e. condition indices (collin), Variation Inflation Factors (VIF) and Tolerance (TOL) should be calculated. The condition index is a general index calculated for the complete model (Belsey et al., 1980). Values around 10 indicate that weak dependencies may start affecting the regression estimates. However, when condition indices are above 100, the estimates have a fair amount of numerical error and at least one variable is collinear with some of the other. VIF and TOL are estimated for the individual variables and indicate which ones are collinear. Values above 2 and below 0.4 for VIF and TOL respectively are considered problematic and suggest that at least one variable should be excluded from the logistic analysis (Allison, 2001).

B. Modeling: univariate and multivariate models

For both univariate and multivariate heuristic and statistical models it is up to the expert to decide which available independent input data layers to include in the analysis. Only relevant data layers should be selected. The expertise of the modeler is hence very important. He should have good knowledge of the type of landslides, the characteristics of the study area, the sources of the input layers and the model to be used. This can be illustrated with some examples. If one observes during the production of the landslide inventory map (i.e. during the field survey or aerial photograph analysis) that the effect of road cutting and filling could influence slope stability, then the data layer “distance to roads” should be included if available at the scale of the analysis. For other regions, field experience might show from the beginning that the location of roads is less important for landslide occurrence so that no efforts and costs should be made to collect this information. The

same is true for the variable "distance to faults". It was already indicated that different buffer sizes should be used when faults are active or not. However, incorporation of faults in the model is not appropriate in case of susceptibility modeling of shallow landslides in stable intraplate zones with only faults in the deep bedrock. Fortunately, statistical analysis helps experts in defining the variables most significantly contributing to landslide occurrence. Below we give some examples of analyses often used in univariate and multivariate landslide models.

Univariate test of association using the previously presented chi-square test (χ^2; Eq. 1) allows testing the association between each input variable and the occurrence of landslides. Hence in this context Oi are for example the distribution of the lithology classes in the study area, while Ei are the distribution of landslides over the lithology classes. One can even go a step further and calculate Cramer's V (Kendall and Stuart, 1979) from chi-square to test the strength and type of association:

$$V = \sqrt{\frac{\chi^2}{N \min(R-1, C-1)}} \quad (3.2.2)$$

where N is the sample size, R is the number of the rows in the contingency table (i.e. in this case the number of lithological classes), and C is the number of the columns (i.e. in this case 2, the study area and the landslide affected area). V compares the observed distribution with the one expected under no relationship, and it standardizes this comparison eliminating the effect of N and the size and shape of the contingency table. Values of V range between 0 and 1 (Kendall and Stuart, 1979; Vanacker et al., 2003).

When carrying out weight of evidence modeling significantly contributing variables can be selected by calculated studentized contrasts (StudC). We refer to Raines (1999), Poli and Sterlacchini (2007) and Van Den Eeckhaut et al. (2008) for more information.

Multivariate statistical models determine which variables discriminate between two or more groups. In landslide studies generally two groups are distinguished: the landslides and the non-landslides. The aim is to determine whether the two groups differ with regard to the mean of the variable, and then to use the variable to predict group membership (e.g. probability, discriminant score, favourability index). Irrespective of the number of variables introduced in the model production, the final model should include only those variables that are significantly contributing to landslide occurrence.

Two-group regression models can often in one way or another be written as:

$$Group = \alpha + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n = \alpha + \beta X \quad (3.2.3)$$

Where α is a constant and βi are the regression coefficients, of which the sign indicates whether the variable is positively or negatively correlated to landslide occurrence.

Multivariable models, such a discriminant analysis or logistic regression, allow a stepwise analysis that provides information on the importance or contribution of an 'independent' input variable to landslide occurrence. In a stepwise model the discrimination between landslides and non-landslides is built step-by-step. Distinction is made between forward and backward stepwise analysis. In the

first, at each step all variables are reviewed and evaluated to determine which one will contribute most to the discrimination between landslide and non-landslides. That variable will then be included in the model. In the backward analysis all variables are included in the model, and then at each step the variable that contributes least to the prediction of group membership is eliminated. In both cases the expert can decide about the minimum statistical significance a variable distinguishing between landslides and non-landslides should have so that the resulting model keeps only the variables that contribute most to group discrimination.

Additionally, some parameters provide information on the contribution of the individual (or a set of) input variables to the overall prediction accuracy. One of them is e.g., Akaike's information criterion (AIC; -2 log (L) + 2n), the negative of twice the log likelihood (-2 log L). Increasing the number of input variables might improve the model fitness, but maybe also renders the model unnecessarily complex. To reach a balance between fitness and parsimony, AIC not only rewards goodness of fit, but also includes a penalty that is increasing with the number of estimated parameters. This penalty discourages over-fitting and complexity. In general, when comparing two models with different sets of input variables the model with the lowest AIC and -2 log L corresponds to the most desirable model (Allison, 2001).

Recently some studies have focused on the selection of the appropriate set of input variables using alternative ways. Without being exhaustive we provide some examples. Guzzetti et al. (2006) tested the robustness of statistical models and the importance of the independent variables by producing 50 different models using each time a different sample of 85% of the data. A variable that is included in all or nearly all of the 50 models is important for landslide occurrence. On the other hand, a variable that is included in only part of the models or that has a positive coefficient in some models and a negative in others as occurred in Van Den Eeckhaut et al. (2009), is not reliable, and should therefore be excluded from the final landslide susceptibility model.

Chung et al. (2002) first constructed a model using favourability curves including all significant variables (n). Then they produced n prediction maps, removing each time one of the variables. By differencing these maps and the map obtained from the model with all variables, n difference maps were obtained, and matching rates were calculated. Using the results obtained, variables were subdivided in influent and non-influent, and a model was made including the influent variables only. Prediction rate curves (see section 5) of the models including respectively all variables and the influent variables are nearly identical, and hence the method allowed derivation of a model including the most important independent variables, and the final susceptibility map was the one obtained from the influent variables.

This procedure is quite similar to using analysis of ROC curve (see section 5) of different models including different combination of variables as was for example done by Dewitte et al. (2010).

3.3 SELECTION OF THE PHYSICALLY-BASED MODELS

(AMRA and UNIFI)

Approaches to landslide hazard assessment depend on the purposes of the research, the size of the area to investigate and the availability of input data (Hervàs & Bobrowsky, 2009). Generally, they can be classified as qualitative or quantitative (Aleotti & Chowdhury, 1999); the former include:

- field geomorphological analyses based on the assessment and/or zonation carried out, directly in the fields, by earth scientists on the basis of their experience;
- development of index/parameters maps through which the experts synthetize and, eventually, weight the factors affecting the slope instability.

The latter comprise:

- statistical analyses where, for the study area, the factors affecting the slope instability are mapped and compared with the mapping of the landslide in order to determine the significance/weight of each parameter and to fix statistical correlation between mapped (observed) landslides and landscape attributes; also the approaches relating the intensity and duration of rainfalls triggering landslides can be considered belonging to this class (Caine, 1980; Chelbroad, 2000; Guzzetti, Perruccacci, & Rossi, 2005; Andriola et alii, 2008);
- physically-based models also called geotechnical models, covered by this paragraph, where the main physical properties are quantified and applied to mathematical models to calculate the safety factor (Aleotti & Chowdhury, 1999); usually, these models are used in a grid-based geographic information system (GIS), which partitions topography into regularly celled digital elevation models (DEMs) and allow for rapid spatial analysis of large areas. These models, however, typically require spatially and temporally distributed model parameters (e.g., soil cohesion, root cohesion, soil bulk density, water table level, friction angle, soil depth, and hillslope gradient) and are highly simplified due to difficulty in characterizing parameter variability over large areas. (Haneberg, 2000). Probabilistic approaches allow for uncertainty by assigning probability distributions to model parameters, while deterministic approaches establish invariant or spatially explicit parameter values and lack an element of uncertainty (Stillwater Sciences, 2007).

The most common physically-based models (deterministic or probabilistic) are based on variations of infinite slope idealization (Haneberg, 2007); therefore, they result more accurate where conditions favor planar failure: a) the thickness of the soil mantle is small compared with the slope length (Borga et al., 2002) so that the edge effects can be neglected; b) the slopes are smooth and have uniform cover of material (Baum et al., 2002); c) as result of buildup of pore water pressure, the sliding surface occurs at the interface (slope-parallel) between the soil mantle and bedrock/denser soil mass or develops at the discontinuity surface determined by wetting front during heavy rainfall events (Rosso et al.,2006; Borga et al.,2002); under these assumptions, the failure of an infinite slope can be characterized by the ratio (factor of safety F_s) , along a potential failure surface, between the shear strength τ_f (usually, expressed via the Mohr Coulomb approach) and shear stress τ.

3.3.1 Probabilistic physically-based approaches

Among the probabilistic physically-based approaches, have been widely used the models LISA (Level I Stability Analysis) (Hammond et al.,1992) and PISA.m (Probabilistic Infinite Slope Analysis) (Haneberg, 2007); they perform probabilistic slope stability calculations for topography obtained from digital elevation models (DEMs). They implement, for the factor of safety, the expression:

$$F_s = \frac{c_r + c_s + (q_t + \gamma_m D + (\gamma_{sat} - \gamma_m - \gamma_w) H_w D) \cos^2 \beta \tan \varphi}{(q_t + \gamma_m D + (\gamma_{sat} - \gamma_m) H_w D) \cos \beta \sin \beta} \quad (3.3.1)$$

where:

c_r cohesive strength contributed by three roots (force/area);

c_s cohesive strength of soil(force/area);

q_t uniform surcharge due to weight of vegetation (force/area);

γ_m unit weight of moist soil above phreatic surface (weight/volume);

γ_{sat} unit weight of saturated soil below phreatic surface (weight/volume);

γ_w unit weight of water (weight/volume);

H_w height of phreatic surface above slip surface, normalized relative to soil thickness (-);

D thickness of soil above slip surface (length);

β slope angle (degrees);

φ angle of internal friction (degrees);

A slope-parallel phreatic surface is used to take into account the influence of the groundwater. Comprehensive discussions about the influence of the vegetation on the slope stability (q_t and c_r) are present in Hammond et al. (1992), Sidle (1992) and Schmidt et al. (2001). To incorporate the effects of the parameter uncertainty and variability, PISA.m uses first-order second moment (FOSM) approximations while LISA uses a Montecarlo simulation method (Phoon, 2008; Haneberg, 2004; Hammond et al,1992; Duan & Grant, 2000; Baecher & Christian, 2004). Through Monte Carlo simulation, in LISA, is generated a considerable number of factor of safety values (about 1000) using, as input parameters, random values included in the range of variability defined in advance for each one of them; therefore, the probability of failure P_f is obtained by dividing the total number of simulations into the number of calculated factor of safety values which are less than or equal to one; to generate the probability distribution function (pdf), in this approach, the parameters c_s and φ are assumed experimentally inversely related while, between the dry unit weight γ_d and φ, is supposed existing a positive correlation.

Instead, for the PISA.m, the main output is constituted by the Reliability Index (RI):

$$\boldsymbol{RI} = \frac{\overline{F_s} - 1}{s_f} \qquad (3.3.2)$$

where $\overline{F_s}$ represents the factor of safety calculated using the mean values of each of the independent variables while s_f represents the F_s standard deviation; it is computed as:

$$\boldsymbol{s_f} = \sqrt{\sum_i \left(\boldsymbol{s_{x_i}^2}\right)\left(\frac{\partial F_s}{\partial x_i}\right)^2} \qquad (3.3.3)$$

with $s_{x_i}^2$ variance of the single parameter; RI represents the distance of the $\overline{F_s}$ from the critical value $(= 1)$ for unit standard deviation; with increasing uncertainty about the parameters, s_f increases resulting in the decrease of RI; in the PISA.m , the parameters are assumed uncorrelated.

3.3.2 Deterministic physically-based models

Usually, they couple slope stability equation with a hillslope hydrological model; they use a digital terrain analysis tool to quantify the basic terrain input data, including slope angle and contributing area (§4.2).The coupling term is constituted by the soil pore water pressure or its function (for example, height of phreatic surface above potential slip surface). It can be assumed as time constant (steady-state models: Montgomery & Dietrich,1994; Borga et al.,2002; Rosso et al., 2006; Pack et al.,1998) or variable (transient models: Baum et al., 2008).

In the following, for the sake of brevity, we will consider three of these: SHALSTAB, SINMAP and TRIGRS.

3.3.2.1 SHALSTAB

SHALSTAB (SHAllow Landslide STABility model - Montgomery & Dietrich, 1994) is a distributed model that couples an infinite-slope stability model (Figure 3.3.1a) with a steady-state hydrological model (Figure 3.3.1b). SHALSTAB is implemented as an extension of ArcView GIS.

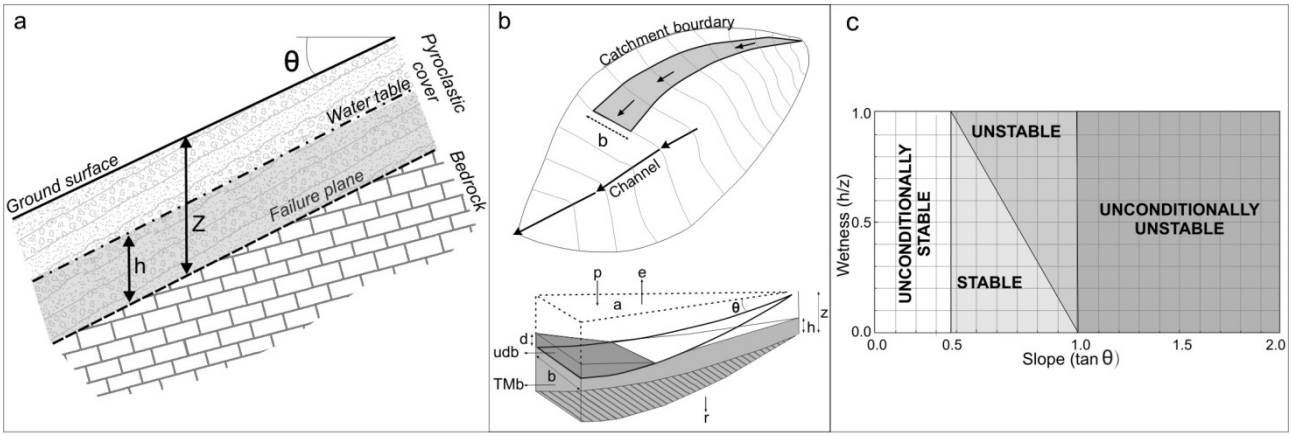


Figure 3.3.1 SHALSTAB conceptual model components (after Montgomery and Dietrich, 1994). Key: a) infinite-slope conceptual scheme; b) hydrological model (p, precipitation; e, evapotranspiration; r, deep drainage; a, drainage area; h, the height of the water table; z, soil thickness; u, mean subsurface flow velocity; h, water level of surface flow; T, transmissivity; θ, slope angle; M, sinθ ; b, contour length); c) definition of stability fields.

The steady-state hydrological model was used to map the spatial variability of a wetness index, given by the ratio of the saturated depth (h) of the soil cover to the total soil depth (z). The wetness index is a function of the slope angle ($\tan\theta$) and of the specific contributing area, expressed by the ratio of the contributing area (a) to the corresponding contour length (b) subtending a (Figure 3.3.1b). Figure 3.3.1c illustrates the relationship between wetness index (h/z) and slope angle ($\tan\theta$), for an angle of internal friction of 45° and a bulk density ratio of 1.6.

The infinite-slope stability model provides an assessment of the susceptibility to flowslide in terms of potential source areas, while it does not provide any prediction of other related phenomena, such as propagation and deposition of the mobilized material. Each location can be plotted in the parametric plane ($\tan\theta$, h/z) (Fig. 3.3.1c), or in the parametric plane ($\tan\theta$, a/b), to identify its susceptibility to landslide triggering. The locations in which stable conditions were never identified, even in case of dry soil, were classified as "unconditionally unstable". These locations generally corresponded to sites with very-steep slope angles and rock outcrops. Conversely, the locations in which unstable conditions never occurred, even in case of saturated soil, were classified as "unconditionally stable". These locations corresponded to areas with slope angles less than 20°, as illustrated in Figure 3.3.1c. In the remaining locations, the landslide susceptibility index was expressed as the upper value of the ratio between i) the effective rainfall (q) to ii) the soil cover transmissivity (T) for which unstable conditions were verified, according to the following relation:

$$\frac{q}{T} = \frac{sen\theta}{(a/b)}\left[\frac{c'}{\rho_w g z \cos^2\theta \tan\varphi'} + \frac{\rho_s}{\rho_w}\left(1 - \frac{\tan\theta}{\tan\varphi'}\right)\right] \quad (3.3.4)$$

where c' is the effective soil cohesion, ρ_s is the soil bulk density, ρ_w is the water density, and φ' is the effective soil friction angle. Table 3.3.1 summarises the default stability classes used in SHALSTAB (after Montgomery & Dietrich, 1994).

Table 3.3.1 Stability classes used in SHALSTAB (after Montgomery & Dietrich, 1994).

Classification	log q/T (1/m)
I	Chron. Unstable
II	log q/T < -3,1
III	-3,1 ≤ log q/T < -2,8
IV	-2,8 ≤ log q/T < -2,5
V	-2,5 ≤ log q/T < -2,2
VI	log q/T ≥ -2,2
VII	Stable

3.3.2.2 SINMAP

SINMAP **(**Stability INdex MAPping - Pack et al., 1998) employs a modelling framework similar to SHALSTAB, i.e. coupling an infinite-slope stability model (Fig. 3.3.2a) with a steady-state hydrological model. The differences between SINMAP and SHALSTAB can be summarised as follows: the two models employ different algorithms for computing the slope angle and the contributing area (Fig. 3.3.2b); SINMAP accounts for uncertainty in soil hydrological and geotechnical parameters through uniform probability-distributions, according to expected or measured uncertainty in model parameters.

The factor of safety can be expressed as follows:

$$FS = \frac{C + \cos\theta\left[1 - \min\left(\frac{R}{T}\frac{a_s}{sen\theta};1\right)\cdot r\right]\cdot \tan\varphi'}{sen\theta} \qquad (3.3.5)$$

where C is the ratio between the soil cohesion and the weight of the soil cover per unit length (Fig. 2c), a_s is the specific contributing area, θ is the slope angle, R is the effective rainfall, T is the soil transmissivity, φ' is the effective soil friction angle, and r is the soil-water density ratio (ρ_s/ρ_w).

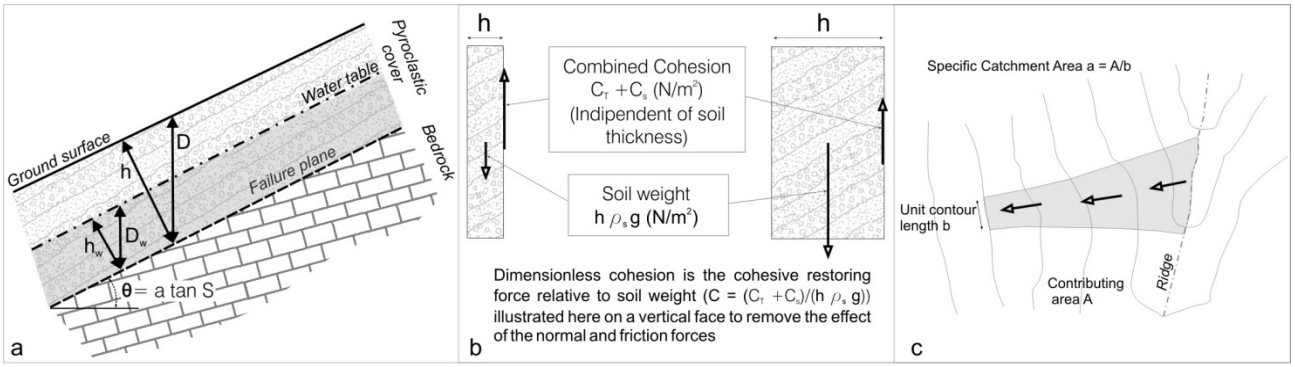


Figure 3.3.2 SINMAP conceptual model components. Key: a) infinite-slope conceptual scheme; b) definition of specific catchment area; c) illustration of the dimensionless cohesion factor concept (after Pack & al, 1998)

The dimensionless parameters R/T, C and $\tan\varphi'$ are treated as random variables with uniform probability-distributions, by assigning upper and lower limits to each of them.

SINMAP computes the susceptibility to flowslides source areas by means of a stability index (SI). Similarly to SHALSTAB, SINMAP does not provide any prediction of propagation or deposition. The safety factor was computed by considering the most conservative combinations of the above dimensionless parameters: it was assumed equal to the computed value if greater than one; otherwise, it was expressed as the probability that the safety factor was greater than 1. Table 3.3.2 summarises the default stability index classes used in SINMAP.

Table 3.3.2 Stability index classifications adopted by SINMAP (Pack et al., 1998). The classes Lower and Upper Threshold refer to two different ranges of the predicted SI, this representing the probability that a location is stable. The areas classified as Lower Threshold are characterised by a probability of being stable larger than 50%, while the areas classified as Upper Threshold are characterised by a probability of being stable smaller than 50%

Classification	Stability Index (SI)
Stable	SI > 1,5
Moderately Stable	1,25 < SI ≤ 1,5
Quasi-Stable	1,0 < SI ≤ 1,25
Lower Threshold	0,5 < SI ≤ 1,0
Upper Threshold	0,0 < SI ≤ 0,5
Defended	SI = 0,0

3.3.2.3 TRIGRS

The hydrological model implemented in TRIGRS (Transient Rainfall Infiltration and Grid-based Slope-stability) is based on solutions of a linearized form of the Richards equation (Iverson 2000; Baum et al. 2002; Savage et al. 2003, 2004). Under the assumptions (Iverson, 2000):

- rainfall of short duration ($t \ll A/D_o$) where D_o is the saturated hydraulic diffusivity, and A is the upslope contributing area above a given location.
- $\varepsilon = H/\sqrt{A} \ll 1$ where H is the depth of the potential failure surface;
- wet initial conditions ($K \approx K_{sat}; D \approx D_o$);

the Richards equation takes the form:

$$\frac{\partial \psi}{\partial t} = \boldsymbol{D_0 \cos^2 \beta} \frac{\partial^2 \psi}{\partial Z^2} \quad (3.3.6)$$

with ψ groundwater pressure head, Z elevation head and β slope angle (Figure 3.3.3).

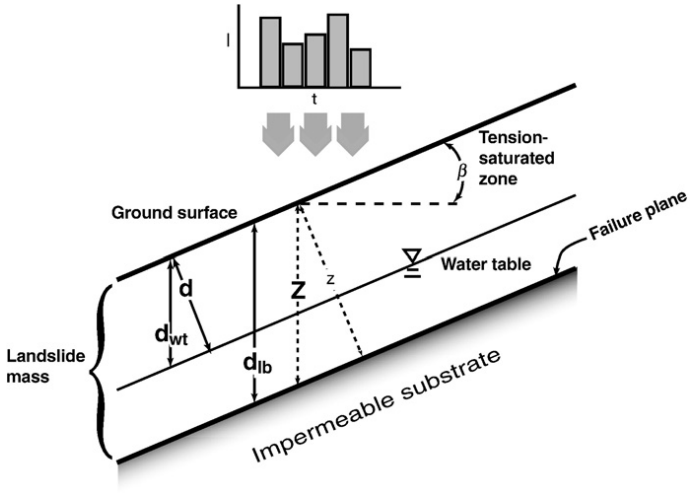


Figure 3.3.3 Sketch showing the coordinate system and groundwater conditions in hillslopes above an impermeable lower boundary at dlb below the ground surface. The depth ofthe water table is dwt, and the slope angle is β. The vertical, Z, and slope-normal, z, coordinates are also shown (Godt et al., 2008).

The solution to 3.3.6, imposing a time-varying specified flux boundary condition at the ground surface and an impermeable basal boundary at a finite depth d_{lb}, given by Savage et al. (2003) and implemented in TRIGRS (Baum et al., 2002), is:

$$\psi(Z,t) = (Z-d)\lambda\cos\beta + 2\sum_{n=1}^{N}\frac{I_{nZ}}{K_s}H(t-tnD1t-tn0.5m=1\infty ierfc2m-1dlb-dlb-Z2D1t-tn0.5+ierfc2m-1dlb+dlb-Z2D1t-tn0.5-2n=1NInZKsHt-tn+1D1t-tn+10.5m=1\infty ierfc2m-1dlb-dlb-Z2D1t-tn+10.5+ierfc2m-1dlb+dlb-Z2D1t-tn+10.5 \quad (3.3.7)$$

The first term on the right-hand side of 3.3.7 describes a steady-state pressure-head distribution in the Z direction for a homogeneous soil in response to long-term rainfall, I_Z; d is the initial depth of the steady-state water table measured in the Z direction; $\lambda\cos\beta = \cos\beta\,(cos\beta - I_Z/K_s)$ is the long-term (steady state) surface flux in the Z direction; The other terms on the right-hand side describe the transient pore pressure distribution in response to a time-varying flux at the ground surface in a layer with a finite depth, d_{lb}, in the Z-direction. $ierfc$ is the first integral of the complementary error function (Carslaw and Jaeger, 1959) and $H(t-t_n)$ is the Heavyside step function (equal to zero until $t = t_n$ and unity thereafter). Iverson's (2000) solution represents a particular case of the general solution (3.3.7), when $d_{lb} = \infty$ and surface flux is applied for a single time interval.

Following Iverson (2000), an additional physical limitation is imposed: the pressure head cannot exceed that which would result from having the water table at the ground surface (TRIGRS):

$$\psi(Z,t) \le Z\lambda\cos\beta \quad (3.3.8)$$

In order to take into account pore-water pressure regime in unsaturated/saturated conditions, it has been added to the original formulation an analytical solution (Srivastava & Yeh, 1991) ; it allows to estimate the soil pore water distributions in unsaturated soil layer due to infiltration processes; the analytical solution of Srivastava &Yeh (1991) is obtained adopting the Gardner model (1958) for estimating the soil water characteristic curve (SWCC) and hydraulic conductivity function:

$$K(\psi) = K_s exp(\alpha\psi) \qquad \theta(\psi) = \theta_r + (\theta_s - \theta_r)exp(\alpha\psi) \quad (3.3.9)$$

with θ volumetric water content, θ_r residual water content, θ_s water content at saturation and α (function of the soil pore-size distribution) representing the desaturation rate of the SWCC; substituting the expressions -9) in one-dimensional vertical Richards equation and posing $K^*(\psi) = K(\psi)/K_s$ and $t^* = \alpha t K_s/(\theta_s - \theta_r)$ the Richards equation assumes the form:

$$\frac{\partial^2 K^*}{\partial Z^2} + \frac{\partial K^*}{\partial Z} = \frac{\partial K^*}{\partial t^*} \quad (3.3.10)$$

it can be analytically solved. The solution to 3.3.10, for sloping surface, imposing a time-varying specified flux boundary condition at the ground surface and water table boundary at a finite depth d, implemented in TRIGRS (Savage et al., 2004) is reported in Savage et al. (2004).

If water reaching the water table exceeds the maximum amount that can be drained by gravity, it is simulated, through a numerical approach, the water table rise while the pressure diffusion below the water table is calculated analytically (Carslaw&Jaeger, 1959).

To estimate the surface runoff a simple method is adopted; where the rainfall intensity and upslope run-off exceed the saturated hydraulic conductivity of the soil, the excess is diverted to adjacent down-slope cells where it can either infiltrate or flow farther downslope. This process is assumed to occur instantaneously.

The factor of safety F_s is expressed as:

$$\boldsymbol{F_s = \frac{c_s - \chi\psi(Z,t)\gamma_w \tan\phi}{\gamma Z \cos\beta \sin\beta} + \frac{\tan\phi}{\tan\beta}} \qquad (3.3.11)$$

with γ soil unit weight and $\chi = \frac{\theta - \theta_r}{\theta_s - \theta_r}$ (Vanapalli & Fredlund,2000) in order to take into account the beneficial effect of the suction on the shear strength.

Calculating 3.3.11 at various depths for transient pressure heads $\psi(Z,t)$, the depth Z where, first, $F_s < 1$, represents the depth of landslide initiation.

4 ROBUSTNESS

4.1 ROBUSTNESS OF SUSCEPTIBILITY, HAZARD MAPS AND RISK MAPS

(UPC)

Most research based on quantitative approaches to landslide susceptibility assessment is focused on data and methods used to prepare the model. However, there is little information on the quality and robustness of the proposed models (Guzzetti et al. 2006).

Given that susceptibility, hazard and risk maps are too influenced by errors in the data, robustness evaluation is essential for land-use spatial planning, management and decision making tasks. Robustness analysis can reduce the uncertainty that leads to risk-ignorant decisions and miscalculation of expected impacts and the costs required to minimize these impacts (Yemshanov, et al. 2010**)**

The term *robustness* is used in statistics to express a technique that performs well even though its assumptions are somewhat violated by the true model from which the data were generated. We use the term more broadly not only for the assumption of the statistical technique but also for any factor that may influence the input data. Hence, the robustness of the model to variations in the data (sensitivity) may come from factors such as: sample (size and design), terrain unit used and statistical technique to construct the model.

4.1.1 Sample selection and size

Any attempt to design a landslide susceptibility model in a region leads to a diagnosis of the goodness of fit of data and prediction capability, and to a validation step. While data fitting and predictability are evaluated with the same data used to obtain the susceptibility model (training, calibration or estimation set), the validation requires landslide information that is not available to construct the model (validation, or test set). To this end, the sampling design to split data into training and test sets and the sample size take on a special importance (Stehman et al, 1998).

Sampling is usually based on spatial (Baeza, and Corominas, 2001) or temporal (Luzi, 1995) landslide criteria. The geological-geomorphological conditions and triggering factors in the occurrence of the past and future movements must be similar in the samples of both the generation model and validation. The method of sampling must utilize some form of random selection regardless of the criteria used to define data sets. Random selection of observations helps to circumvent the potential risks of spatial autocorrelation in statistical analyses (Diniz-Fiho, et al. 2003). This correlation should be avoided. Simple, stratified, systematic, cluster or multi-stage random samplings are some of the random methods that can be used (Trochim, 2006) but few papers apply them (Guisan et al, 2000; Powell, et al. 2004; Hjort, et al. 2008; Baeza, et al. 2010, Frattini et al. 2010). Simple random sampling is not the most statistically efficient method of sampling and cannot yield representative subgroups in a population made up of landslides and not landslide areas. Stratified random sampling ensures that not only overall population, but also key subgroups, especially small minority group (landslides) will be represented. Cluster or area random

sampling may be used across a wide geographic region which has different conditioning variables for landslide susceptibility. When sampling methods are combined, multi-stage sampling is selected.

Mathematical models based on statistical inference are used to assess landslide susceptibility. The statistical inference makes propositions about large information, using data drawn from the database of interest via some form of random sampling (Lohr, 1999). If the *sample size* cannot be accurately justified, the researcher will not be able to make a valid inference, and the reliability of the susceptibility model will be compromised (Israel, 2008). Different commercial software calculates the sample size which depends on the type and variability of the variables to analyze. In the literature, this is not usually taken into account and the sample size and their representativeness are not justified. Only some papers analyze the effect of sample size on the accuracy of susceptibility models. Baeza et al. 2010a, determined the size of the ideal sample for a susceptibility map in the eastern Pyrenees considering nominal scales (presence or absence of landslides). In the light of their findings, a sample correctly represents the population of the study area with 50% of their observations, yielding good results of landslide susceptibility with discriminant analysis. Hijort et al 2008 concluded that the predictive models improved sharply when the sample size grew from 20 to 100 and the level of robust predictions was researched with 200 observations. The robustness of the model to changes in the input data was analyzed in Guzzetti et al 2006. In this paper, discriminant models were constructed using a different number of terrain units, from 268 (30%) to 849 (95% of the 894) units, and the random selection was repeated 50 times for each subset. The fit and correct classification of the discriminant model did not improve significantly when exceding 75% of the terrain units.

It should be noted that, under the usual statistical assumptions, most statistical techniques used to model the susceptibility are consistent, efficient and asymptotically normal. But these properties hold as the sample size approaches infinity (*central limit theorem*, Fischer, H. 2010). The majority of statistical analyses are fairly sensitive to the ratio of sample size and, in general, the larger the sample size, the smaller the sampling error. Most popular susceptibility assessment techniques are not recommended when the sample is less than 100. At least 10 observations per predictor variable should be included in the analysis and more than 10 when the variables are of a categorical nature (Long, et al. 2001). Hence, the reliability of the final susceptibility model depends on the sample selection and size, which challenges the conclusions of some published papers.

4.1.2 Terrain units

Landslide susceptibility mapping requires a terrain unit characterized by geomorphological and geological attributes that differ from the adjacent unit by some limits. These attributes configure the stability conditions (Hansen 1984). Similarly, a dependent variable is attributed to each unit, which corresponds to the observed landslide (failed) or not observed (unfailed). Different types of terrain mapping units (unique conditions terrain units, slope units, grid cells units, administrative divisions, etc.) have been proposed (Meijerink 1988). However, the selection of terrain units for modeling purposes is often undertaken more in accordance with computating facilities than in accordance with the terrain unit and its relation to the objectives of the research (García, et al. 2008). The main

terrain units have been compared in Carrara et al.(1995) to resolve the differences of consistency and physical meaning. The geomorphological unit has an optimal physical meaning for assessing landslide susceptibility. However, assessment is not easy to carry out since the boundaries must be defined manually (Pasuto and Soldati 1999) with the result that the map will have a lower consistency than other terrain units.

In Garcia et al (2008) the influence of the grid-cell units, slope units, and unique conditions resulting from the overlapping of lithology and slope in the assessment of landslide susceptibility is evaluated by computing and comparing the success-rates and areas under the curve for each model. Carrara et al (2008) and Frattini et al (2010) evaluated different terrain units (fine and coarse slope unit and grid cell) using statistically (discriminant and logistic) and physically-based methods. In the light of their research, the coarse slope unit model outperformed the others, whereas the physically-based model was always the poorest.

As stated above, the type of terrain unit affects the final model. The dimension of the unit also exerts an influence on the reliability of the model. In Baeza et al 2010a, an attempt was made to evaluate the error of the DTMs by comparing the values of the variables of different terrain units: grid cells of 15 and 45 m (automatic from GIS) and the slope unit (field database). The results showed that the mean error (using the slope unit as the "gold standard") for the variables based on 15-m grid cell ranged from 19.5% for the slope angle to 33% for the watershed length and, the mean error for 45-m grid cell reached 33.5% for the slope angle and 98% for the watershed length. The values estimated with GIS decreased as grid cell size increased. Given that some of the statistical techniques most commonly applied to evaluate susceptibility (Discriminant and Logistic regression) use mean values of the variables to differentiate between populations (unfailed and failed), the greater the smoothing of the variable when 45-m grid cell is used, the smaller the separation between the populations and the poorer the discriminant model. Hence robustness is susceptible to the type and the size of the terrain unit and should be considered when the quality of landslide susceptibility model is estimated.

Since geostatistical methods are based on the analysis of the relationships between landslide and the spatial distribution of some instability factors, the manner in which the landslide source area is represented can seriously affect the robustness of the susceptibility model. The strategies of its representation can generate changes in the input variables and in the capability of describing relationships between landslide and conditioning factors. Poli and Sterlacchini 2007 analyzed the application of specific strategies of a source area represented: only by the centroid of the polygon rasterized considering a pixel size of 20m, and a pixel size of 50m (Lee et al, 2002). Although there was little influence on the final response map, considering the scale of the work, the interpretation of the relationship between variables and landslide location was evident when using centroid with respect to a set of points. A new approach was followed in the generation of decision rules of landslide occurrence by Süzen and Doyuran (2004), termed ''seed cells''. A buffer zone was added to the crown and flanks of the landslide polygon to reveal the best undisturbed morphological conditions, simulating conditions before the occurrence of a landslide, which was used to extract the critical parameter ranges responsible for the landsliding mechanism.

4.1.3 Statistical tools

4.1.3.1 Model assumption

Not all research takes into account that the characteristics of the data may affect the performance of different statistical methods to assess susceptibility. There are a number of assumptions that underlie most statistical models and research should guarantee these assumptions in order to obtain the best results at the different stages of the construction of the model. These assumptions are readily dealt with through the design of the study. If these assumptions were checked, the results of the study would benefit considerably. Four of the most important assumptions that are not highly robust to violations are as follows:

Multivariate normal distribution: It is assumed that the data represent a sample from a multivariate normal distribution. Whether or not variables are normally distributed can be examined with histograms of frequency distributions, and the Kolmogorov-Smirnov test, which provides inferential statistics on normality (Baeza and Corominas, 2001; Baeza et al, 2010b, Kiang, 2003). When the distribution was not symmetrical but skewed, different transformation (e.g., square root, log, or inverse) can improve normality but can complicate the interpretation of the results. A detailed discussion of the robustness of the *F* statistic used in most of the classification methods applied in assessing susceptibility can be found in Box and Andersen (1955), or Lindman (1974).

Violations of the normality assumptions may lead to a biased and overly optimistic prediction of the performance of rules in the population, and thus limit the usefulness of the model.

Multicolinearity: Independence of the predictor variables is required to create a mathematical equation. High degree of multicolinearity will have adverse effects on the parameters estimated by Disciminant Analysis or Logistic methods. Hence, several analyses (principal component analysis– PCA- and analysis of variance ANOVA/MANOVA tests) can be applied to test independence. Baeza et al 2010 show a statistical design to guarantee this assumption when constructing a model. A detailed review of the performance of the main tests can be found in Park (2008).

Linear relationship: Many statistical analyses that are used to define susceptibility models can only accurately estimate the relationship between dependent and independent variables if the relationships are linear in nature. If they are not linear, the results of the analysis will under- or over- estimate the true relationship, increasing the chance of a error Type II (false negative- under estimation) for independent variables (see section 5.2. of the deliverable) or increasing the risk of error Type I (false positive- overestimation is higher in the area susceptibility) for independent variables (Tabachnick, et al. 2001).

Homocedasticity: The robustness of the mathematical model is higher when the variance/covariance is identical in the two (failed-unfailed) or more groups analyzed. This assumption is called the homogeneity of variances or homocedasticity. When heteroscedasticity (variances not equal) is

marked it can lead to a serious distortion of findings and seriously weaken the analysis, thus increasing the possibility of error Type I (over estimated the susceptibility). If this assumption is violated, a formal test for heteroscedasticity should be performed and probably a simple transformation of the independent variables could reduce this problem.

Sample proportion: Prediction models using Discriminant Analysis require populations (landslide/not landslide) that have a similar number of observations to avoid the bias of the mathematical function (Dillon and Goldstein, 1986). Hence, the equal sample proportion is usually used to train the models. However, the models show that when the sample proportion differs from the true population, the prediction accuracy becomes very poor. If the true population is very different, as in most hazard studies, other non-parametrical mathematical models (Logistic, K-Nearest Neighbor), which are not affected by biased sample proportion, should be chosen.

Sample proportion should also be considered when selecting the accuracy statistics commonly used in classifying and predicting a model. However, accuracy statistics derived from a confusion matrix (efficiency, misclassification rate, odds ratio...) are usually applied incorrectly on an unequal sample in many studies. Other accuracy statistics that do not depend on prevalence (sensitivity, specificity, likelihood ratio…) should be used (Beguería, 2006; Liu, et al. 2007; Frattini et al. 2010). The prevalence of one of the groups has a considerable effect on the nature of the model. As a result of this, the model will be liberal or conservative depending on the dominant group.

4.1.3.2 Type and number of predictor variables

A mixture of *continuous and categorical predictors* (e.g. slope angle, watershed area are continuous, lithology, surficial formation thickness are categorical) is usually used for predicting a categorical dependent variable in landslide susceptibility and hazard assessment. However, the robustness of a model to the type of data depends on the procedures. Hence, the reiterated use of categorical predictors in discriminant analysis is statistically questionable. The robustness of the model is low when categorical variables are used. Caution should be used before accepting the results of tests of statistical significance, and before drawing final conclusions from the analyses.

In such cases, similar analyses (Logistic or Generalized Linear Models) that involve the interaction between categorical and continuous predictors are available.

Some studies use a high *number of variables* and reach good prediction models. Nevertheless, a susceptibility model obtained using a reduced number of predictor variables is less expensive than a model with a large number of variables. Few studies have carried out the robustness analysis for variable selection. Guzzetti et al (2006) and Melchiorre et al (2006) reveal that the use of a stable combination of variables makes for a robust model that can cope with uncertainty in the input data. The former study showed that the higher the variable number, (when the sample is small) the better the model classification at the expense of the variability of the model. The model was less reliable and less useful in practice. The latter study revealed that the reduction of the input variables increased the robustness of the neural model. Hence, the models should apply Occam's razor or law of parsimony, which recommends selecting the simplest model (lower number of variables) that accounts for the susceptibility to equality of variance. In this regard, the reader is referred to Baeza et al (2010b) where a procedure of selecting input variables is explained. The Principal Components

and the analysis of variance test were used to select the predictors that exert most influence on the dependent variable.

As regards the aforementioned assumptions and data characteristics, the statistical power of analysis often increases, giving a desirable outcome. Nevertheless, it should be noted that there are many types non-parametric statistical techniques available when the assumptions of a parametric statistical technique is not satisfied (Gibbons et al. 2003; Wasserman, 2007; Chan, 2003). Although these are often not as powerful as parametric techniques, they provide valuable alternatives. A detailed research in Kiang (2003) provides hypotheses between data characteristics and method performances -Neural Networks, Decision Tree, Multivariate Discriminant Analysis (DA), Logistic Models and Nearest neighbor (kNN)-. In the light of her findings, Logistic method is superior to Discriminant Analysis when model assumptions are violated; the performance of DA is better than other methods when there is an unequal sample proportion, and kNN improves when the sample size increases. In general, Neural Networks and the Logistic methods are superior to other methods. However, Neural Networks contain more parameters to estimate than Logistic Regression, and therefore require larger data sets to reach a comparable level of accuracy. On the other hand, Decision Tree is the lowest method in relative performance. Finally, this study concludes that data characteristics have a considerable impact on the classification performance of the methods and on the final susceptibility and hazard maps.

The effect of violation of basic assumptions and data characteristics is reflected in susceptibility and hazard maps by estimation errors, which can have important implications for land management and decision-making tasks. When the mathematical model underestimates the susceptibility, the final map could lead to loss of life or to the destruction of infrastructures as a result of incorrect classification of hazardous areas. If the susceptibility is overestimated, the map could imply the loss of a potentially safe space, or even the uselessness of investments made for prevention in areas that might not be dangerous.

To sum up, the robustness of the model to change in the sample size, terrain unit and in data characteristics (model assumption and type of variables) is, in general, low and depends on the methods (Decision tree and kNN are more vulnerable). Consequently, the reliability of the estimates of the model may be affected. To obtain a robust model, the mathematical procedure to assess the susceptibility should be chosen taking these factors into account.

4.1.4 Final remarks

Random selection of observations should be employed to avoid the spatial autocorrelation in statistical analyses and to obtain a valid inference. The stratified random sampling is recommended to ensure a good representation of the population.

Most multivariate statistical techniques are sensitive to sample size. Hence, the representative sample size of the population in any study should be justified to ensure the reliability of the landslide susceptibility model.

The type and size of the terrain unit modify the input variables used for susceptibility modeling as well as the ability to discriminate between landslides and not landslide areas. Thus, the most suitable terrain unit should be taken into account for the accuracy of the model.

Data characteristics affect the performance of statistical methods. The violation of the model assumptions can lead to an optimistic prediction and can overestimate susceptibility, restricting the usefulness of the model. The statistical method should be chosen taking into account the type and characteristics of the data that describe the relationship between landslide and conditioning factors. Non parametrical statistical techniques are recommended when data do not satisfy the basic assumptions. Neural Networks and Logistic regression methods provide a better relative performance than other methods.

4.2 SENSITIVITY ANALYSES OF PHYSICALLY-BASED MODELS

(AMRA)

In section **3.3**, the physically-based models are classified as probabilistic or deterministic depending on whether they take into account or not the uncertainties about soil parameters or other attributes, hardly well-known, as the soil thickness above the bedrock or initial hydraulic conditions for transient approaches. Actually, in order to consider them also in the deterministic approaches and to carry out a proper model calibration, have been developed various approaches; in the following, are then shown two examples.

Salciarini et al. (2006), in order to model landslide susceptibility of a study area in the eastern part of Umbria Region (about $100km^2$), carry out back-analyses by TRIGRS model; in these analyses, the thickness of the superficial soil d_{lb} is estimated through an experimental exponential relation (Figure 4.2.1); to calibrate soil geomechanical and hydraulic parameters for which are available ranges of variability, obtained through in situ and laboratory tests (Table 4.2.1), they perform a parametric study on a single "virtual" cell, assuming an average slope angle and water table depth $(d_{wt} - d_{lb})/d_{lb} = 0.25$ and investigate values of strength and permeability that produced stability ($F_s > 1$) in the initial condition and instability after the application of the reference rainfall event.

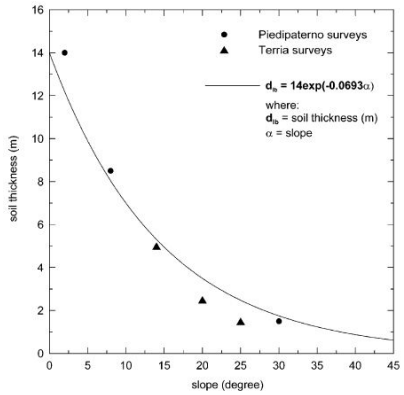


Figure 4.2.1 Exponential soil thickness model (Salciarini et al., 2006)

Table 4.2.1 For each soil parameter, it is shown the range of variability obtained by laboratory and field tests and the value adopted after the calibration

	c' (kPa)		ϕ (°)		γ_s (kN/m^3)		K (m/s)	
	range	value	range	value	range	value	range	value
zone 1	0-10	4	24-34	30	18-19.5	18	1e-5-1e-2	1.00E-04
zone 2	5-10	10	18-22	18	19-19.6	19	-	1.00E-04
zone 3	20-80	30	26-34	30	19.6-22	20	1.00E-08	1.00E-08
zone 4	40-100	40	40	40	21-21.5	21	1.00E-06	1.00E-06
zone 5	350-100	100	35-50	40	22-29	22	1e-6-1e-4	1.00E-06

After the calibration, to evaluate the influence of the initial water depth, are considered other three additional cases $((d_{wt} - d_{lb})/d_{lb} = 0, 0.5, 1$; the model performances are estimated considering the total error due to the sum of "false positives" (prediction of a landslide where none has occurred; it is estimated as the percent of the total study area predicted to be unstable minus the percent of the total area covered by landslides) and "false negatives" (no prediction of a landslide where one has occurred, computed as the number of landslides outside the area predicted to be unstable multiplied by their area); the results are shown in Table 4.2.2.

Table 4.2.2 Summary of the different types of error in the prediction of the TRIGRS model (modified from Salciarini et al.,2006)

$(d_{wt} - d_{lb})/d_{lb}$	False positives (%)	False negatives (%)	Total error
0	1.5	37	38.5
0.25	2.8	20	22.8
0.5	4.5	17	21.5
1	7.6	15	22.6

Godt et al. (2008) describe an application of the TRIGRS (Baum et al.,2002) to an $18km^2$ study area located in southwestern Seattle (USA); in order to define the thickness of colluvium cover, following Schulz's (2003) approach, they distinguish three different hillslope landforms (escarpment, midslope and footslope); furthermore, they consider colluvium thickness can be assumed function of four topographic parameters (1) topographic slope angle of the ground surface, (2) slope angle of the escarpment, (3) height of the escarpment, and (4) distance downslope from the escarpment crest; using a database of 77 borehole, they carry out experimental relationships between the colluvium thickness for each hillslope landform and topographic attributes. For what concern the initial water depth, through the software VS2DI (Hsieh et al., 2000) they have been carried out a back-analysis of water-level information obtained by the boreholes available in the study area; in this way, is obtained a relationship between the saturated thickness of the colluvium and two variables: slope angle at the base of the colluvium and saturated hydraulic conductivity; in this case, hydraulic parameters, obtained by laboratory and field tests, are assumed constant for

each zone identified in the study area while material strength values were assumed spatially invariant and were determined from literature values (Savage et al., 2000; Montgomery et al., 2001) and laboratory tests.

In addition to the considered parameters, other factors can affect the analysis reliability and the predictive capabilities of the physically based models:

- DEM resolution;
- algorithm for specifying flow directions;

In special way, these factors influence two parameters: the local slope angle β and a contributing upslope area per unit width of contour (also called specific upslope area); the contributing area represents the catchment area at any given point in space while the unit flow width in grid based DEM is approximately the length of grid cell (Gallant and Wilson, 2000).

As regards the first problem, generally speaking, coarser DEM resolution leads to lower β values because of smoothening effects whereas the distribution of a tends to shift towards larger values (Wilson et al., 2000); Zhang and Montgomery (1994) consider a DEM resolution equal to 10m sufficient to take properly into account the runoff processes and that, increasing the resolution to 2 or 4 m, no important additional information are added.

Few studies have been done on the identification and quantification of the influence of DEM resolution on landslide hazard assessment. Dietrich and Montgomery (1998) show two example to illustrate the DEM resolution effects on the hydrologic ratio $log(q/T)$ where q is steady-state rainfall and T is transmissivity; comparing the results obtained with SHALSTAB between a 30 m and a 6 m DEM and a 10 m and 2 m DEM, they conclude that, although percentages of the landscape in the moderate landslide hazard classes are similar for coarse and fine resolutions, the spatial patterns in general differ in important ways: finer resolutions lead to patterns of relative slope stability much more strongly defined by local ridge and valley topography; low $log(q/T)$ values (higher slope instability potential) are more concentrated in steep valleys, rather than spread out across the landscape; for these reasons, they conclude that with finer resolution topography, sites with highest instability increase and can be delineated more precisely, rather than mapping broad zones of instability in the case of coarser resolutions (Claessens et al., 2005); similar conclusions can be drawn considering Figure 4.2.2 (Claessens et al., 2005).

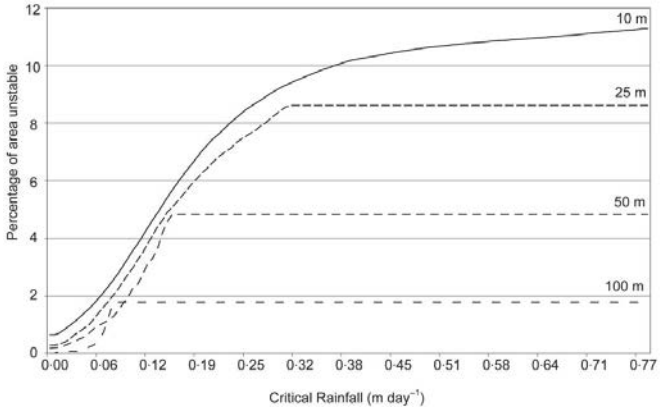


Figure 4.2.2 Percentage of the study area having a certain critical rainfall value for the 4 different DEM resolutions (Claessens et al., 2005)

For the different DEM resolutions, is shown the percentage of the study area having a certain critical rainfall value, interpreted as the relative potential for landsliding or becoming 'unstable' (calculated via the Shalstab model §3.3.2.1); the y-intercept (critical rainfall equal to zero) represents, for each DEM resolution, the percentage of area unconditionally unstable; it is only function of the slope angle and so increases with higher resolutions; along the entire trends, are clear the effects of higher slope angles associated to finer resolutions; instead, the effect of specific catchment area is not very clear; probably, the effect of higher values for coarser resolutions are nullified because those higher values tend to occur in the valley bottom towards the outlet cell, where slopes are generally low enough to be classified as unconditionally stable.

In the physically based models, the three often used algorithms for specifying flow directions are:

- D8 (O'Callaghan & Mark, 1984); it assigns flows from each grid cell to one between its eight neighbors (cardinal or diagonal) with the lowest elevation; it is a very simple approach but it may depend strongly on the orientation of the grid system (Montgomery & Dietrich, 1994);
- FD8 (Quinn et al.,1991; Freeman, 1991); these approaches allocate water flow, fractionally, to each lower neighbor in proportion to the slope; as Tarboton (1997) pointed out, they have the disadvantage that flow from a pixel is dispersed to all neighboring pixels with lower elevation.
- D_∞ (Tarboton, 1997); this approach defines eight planar triangular facets between each grid point and its neighbors; only the two neighboring cells defining the steepest facet receive upslope flow; the fractional flow is divide according the rule referred in Figure 4.2.3;

It is worthwhile to note that these approaches are used, without distinction, for computing subsurface flows (for example, for SHALSTAB & SINMAP models) and surface runoff (for example, for TRIGRS model).

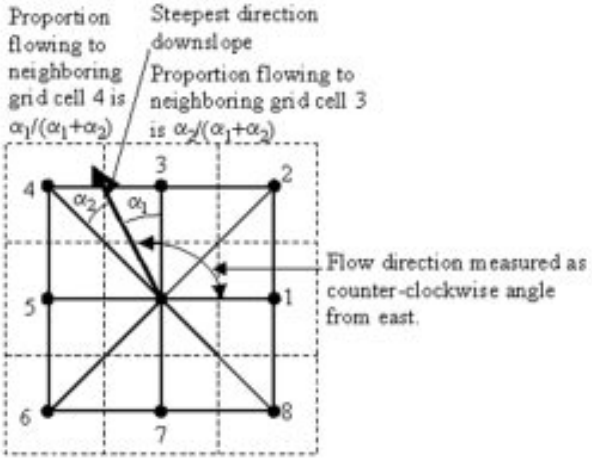


Figure 4.2.3 Flow direction defined as steepest downward slope on planar triangular facets on a block centered grid

An interesting research carried out by Huang et al. (2007) by using SHALSTAB model allows to identify how the differences of algorithm can induce different predictions of FS patterns and FS values; in Table 4.2.3 are referred estimated soil wetness distributions (§ 3.3.2.1) derived from the three algorithms during different reference rainfalls: the more divergence algorithm (FD8) induces the larger saturated area (soil wetness > 0.8); the differences between the algorithms FD8 and D8 are about 30% for the less intense rainfall event and are reduced to 15% for the heaviest rainfall; for this reason, using FD8 approach, the number of predicted unstable cells increases causing an overestimation of the predicted areal extent of landslide initiation (Table 4.2.4).

Table 4.2.3 Soil wetness distributions derived from algorithms during different rainfalls (Huang et al., 2007)

Area (%)	Haitang 144 (mm/day)			Mindulle 248 (mm/day)			Herb 327 (mm/day)		
Soil wetness	D8	Dinf	FD8	D8	Dinf	FD8	D8	Dinf	FD8
<0.2	1.9	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.2–0.4	17.3	9.6	1.3	6.0	1.8	0.0	0.2	0.0	0.0
0.4–0.6	13.2	11.0	4.4	10.5	6.1	0.7	7.5	2.5	0.0
0.6–0.8	9.0	9.1	5.5	9.1	6.9	1.9	7.3	4.4	0.5
>0.8	58.6	69.9	88.8	74.4	85.2	97.3	84.9	93.1	99.5

Table 4.2.4 Simulation results and model performances for three algorithms (Huang et al., 2007); in the parenthesis, near the reference rainfall denomination are referred landslide number and landslide in unit of cell; RPL means Rightly Predicted Landslides; PUC means Predicted Unstable Cells; modified from Huang et al. (2007)

	D8		D_∞		FD8	
	RPL	PUC	RPL	PUC	RPL	PUC
Haitang (24,1116)	18	2465	18	2949	20	4032
Mindulle (39,1245)	31	3600	32	4188	33	5186
Herb(20,2203)	18	5457	18	6133	18	6964

5 ASSESSING THE ACCURACY OF THE PREDICTIONS

(UNIMIB, CNRS and JRC)

The most relevant criterion for quality evaluation is the assessment of model accuracy, which is performed by analyzing the agreement between the model results and the observed data. In the case of landslide susceptibility models, the observed data consist of the presence/absence of landslides within a certain terrain unit used for the analysis.

Ideally, model developers and end-users (local administrators and population) should wait for the occurrence of new landslide events to verify whether they take place in areas classified by the model as unstable. This approach is certainly scientifically valid but inapplicable for end-users.

In practice, accuracy assessment can be performed using either the same dataset used to build the model (i.e., training and validation sets are note separated) or a different dataset, called test or validation set that is separated from the training or calibration set either spatially, temporally or randomly. In the latter approach, the predictive success of the model, built on the training set, is performed on the validation set.

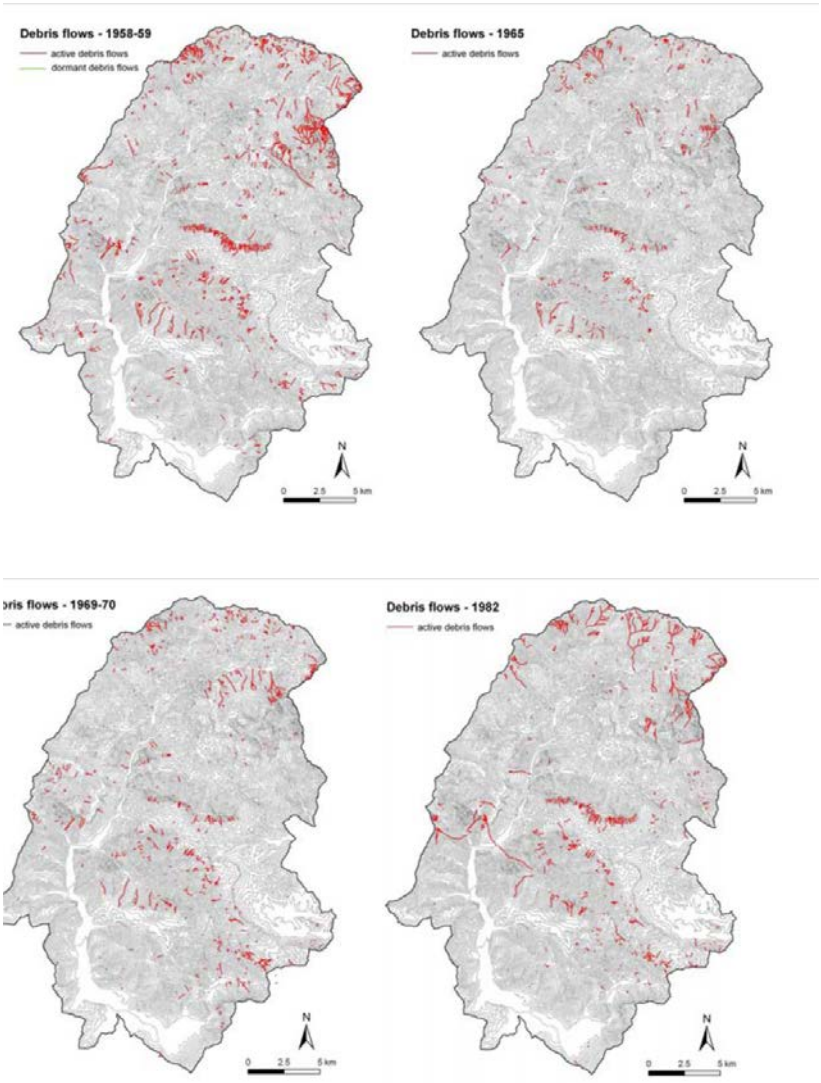
A spatially separated validation set can be located contiguous to the training set, or in other words it can be obtained by splitting the study area in a region used for training and an adjacent region used for validation (Chung and Fabbri, 1999; 2003; Guzzetti et al., 2005). However, because landslides are intrinsically related to local factors (e.g., geological units, structures, land-use, rainfall distribution) that can differ even between contiguous areas, this approach may be conceptually and operationally questionable. The only reliable application can be achieved under the condition that the contiguous areas are actually similar with respect to landslide controlling factors such as geology, climate and morphology.

A temporally separated evaluation/validation set would be the best choice, since it directly tests the quality of the model. However, the time period used is critical for “validation”. It is difficult to obtain a significant time validation for deep-seated landslides since they can be very rare phenomena. On the other hand, most shallow landslides and debris flows are activated during rainfall events, whose distribution, intensity and duration is normally very different from one to another. Hence, it is possible that the processes related to the events used for validation could differ from those that generated the landslides in the training set (Figure 5.1). In other words, temporally separated training and validation sets need a significant period of observation for both datasets in order to avoid the biasing effect of single landslide events.

A strategy commonly adopted in applied statistics to generate separated training and validation sets is based on the random splitting of the data set into two groups. The size of the training and validation sets is defined according to the size of the sample of terrain units. If the sample is large enough, terrain units are split into equal-size sets. If the sample is relatively small, it can be preferable to reduce the size of the validation set in order to build the model on a significant and robust training set (Carrara et al, 2008).

Time partitioning, spatial partitioning and random selection are called data splitting techniques by Mucherino et al. (2009). Apart from that there is Cross-validation where only one dataset is used for both calibration and validation. In a first step the data is divided into k equal subsets. Then, a model is fit on k-1 subsets and validated with the predictions for the one left-out subset. Finally, a measure

of predictive accuracy, like mean-squared prediction error (MSPE), is calculated. These three steps are repeated for each subset and all measures of predictive accuracy are averaged. The leave-one-out method in which k is equal to the number of elements in the dataset is a special case of this k-fold cross validation, but can be time-consuming when having large datasets. Therefore the choice of k should be the trade-off between computation speed and accuracy. Of course adopting a "bootstrapping" re-sampling technique enables automatic repetition of the model calibration and validation steps (e.g. Rossi et al., 2009).



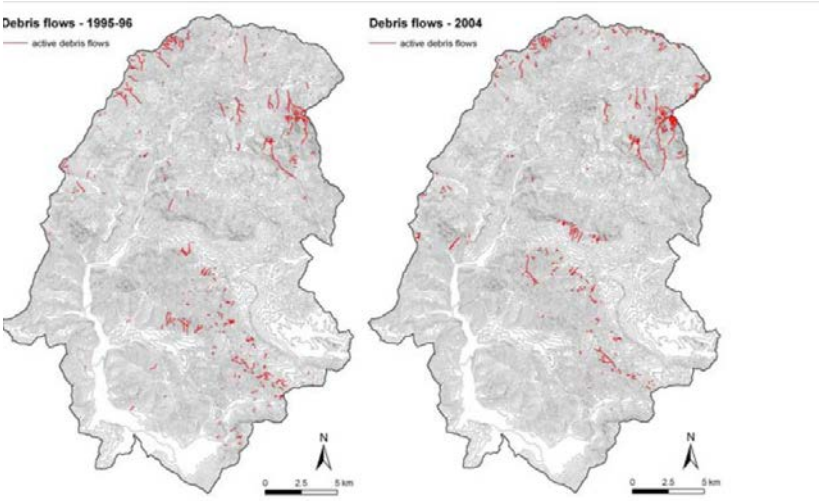


Figure 5.1 Multitemporal debris-flow inventory for a Pre-Alpine Italian valley (Val Trompia, Brescia, 510 km2). The comparison of active debris flows recognized in different aerial photos shows a clustered distribution over the area. A model developed for a single event (e.g., 1965) could result in a susceptibility map that is not reliable for some areas (e.g., in 1965 debris flows were absent in the south-eastern portion of the area) (from Crosta et al, 2007)

5.1 TRADITIONAL VALIDATION METHODS

Whereas landslide susceptibility models have been produced for decades, only recently awareness is raised on the importance to assess their quality. Before, models were often not evaluated. In their 'framework for landslide risk assessment and management', Dai et al. (2002) for example do not include validation. In the pioneering susceptibility models produced beginning in the 1980s, accuracy was evaluated through visual comparison of actual landslides with susceptibility classification (Brabb, 1984, Gökceoglu and Aksoy, 1996), or in terms of efficiency (or accuracy) (Carrara, 1983). In the last decade, different authors have suggested equivalent methods to evaluate the models in terms of landslide density or area within different susceptibility classes. Examples of comparison between the landslide susceptibility classes on the susceptibility map and the number of landslides per km^2 or the total area affected by landslides exist for both statistical (e.g. Lee and Min, 2001; Ercanoglu and Gokceoglu, 2002; Van Den Eeckhaut et al., 2010; Fig. 5.2; Table 5.1) and physically-based (e.g. Montgomery and Dietrich, 1994; Crosta and Frattini, 2003) models. Baeza and Corominas (2001) defined an index of relative landslide density (i.e. degree of fit) as the ratio between the density of landslides of a given susceptibility class (n_i) and the overall landslide density (N_i). The index takes the following form:

$$100\frac{(n_i / N_i)}{\sum(n_i / N_i)} \quad (5.1.1)$$

If evaluated, the predicted susceptibility map was generally confronted to the same landslide dataset as was used for producing the model. Often, authors were aware that the landslide susceptibility map should not perfectly match the mapped landslide distribution as areas currently free of landslides can be affected in the future.

Table 5.1.1The b/a ratio can be used for evaluating landslide susceptibility maps. Here a and b represent the distribution of susceptibility classes over the study area and the distribution of the landslides over the susceptibility classes respectively (after Lee and Min, 2001). Hence, b is similar to n_i in Eq. 5.1.1.

Range	Total		Landslide occurred		*b*/*a*
	Count	Ratio (%), *a*	Count	Ratio (%), *b*	
0.00	157,973	23.98	43	0.37	0.02
0.01	54,750	8.31	131	1.12	0.13
0.02–0.03	75,386	11.44	273	2.33	0.20
0.04–0.05	48,399	7.35	507	4.32	0.59
0.06–0.07	55,795	8.47	841	7.17	0.85
0.08–0.09	58,998	8.96	1,305	11.12	1.24
0.10–0.13	75,926	11.53	1,975	16.83	1.46
0.14–0.19	63,914	9.70	2,423	20.65	2.13
0.20–0.99	67,649	10.27	4,237	36.11	3.52
Total	658,790	100.00	11,735	100.00	1.00

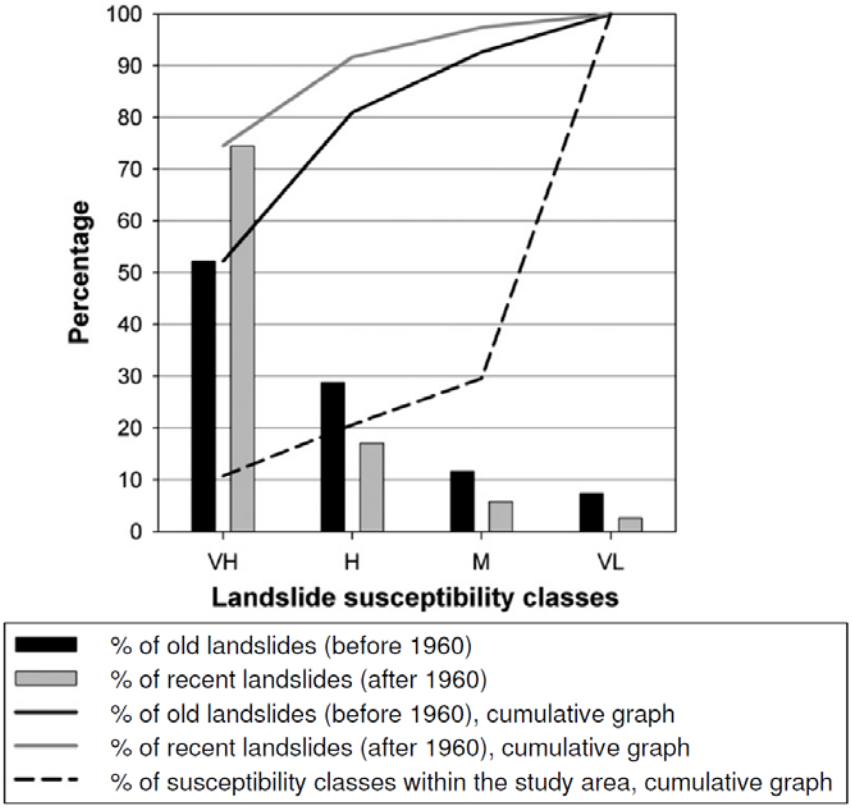


Figure 5.1.1 Distribution (bars) and cumulative distribution (continuous lines) of area of observed landslides on the classified landslide susceptibility map obtained (with the logistic regression modelling; after Van Den Eeckhaut et al., 2010).

5.2 STATISTICAL VALIDATION

As previously mentioned, validation is performed by analyzing the agreement between the model results and the observed data. Since the observed data consist in the present/absence of a landslide within a certain terrain unit, the simpler method to assess the accuracy is to compare these data with a binary classification of susceptibility in stable and unstable units. This classification requires a cut off value of susceptibility that divides stable terrains (susceptibility lower than cut off) and unstable terrain (susceptibility higher than cut off).

The comparison of observed data and model results reclassified into two classes is represented through contingency tables (Table 5.2).

Table 5.2.1 Contingency table used for landslide model evaluation (N_O and P_O: number of stable and unstable observation, N_P and P_P: number of stable and unstable predictions; T: total number of observations

Observed		Predicted	
	Class 0 (-) stable	Class 1 (+) unstable	
Class 0 (-) stable	(-|-) true negative, tn	(+|-) false positive, fp	N_O
Class 1 (+) unstable	(-|+) false negative, fn	(+|+), true positive, tp	P_O
	N_P	P_P	T

By combining correct and incorrect classified positives (i.e., unstable areas) and negatives (i.e., stable areas) it is possible to derive a number of quality measures called binary accuracy statistics, such as is the Efficiency (Finley, 1884) , the Threat score (Gilbert, 1884), the Equitable threat score (Gilbert's skill score; Gilbert, 1884, Schaefer, 1990), the Pierce's skill score (True skill statistic; Pierce, 1884; Hanssen and Kuipers, 1965), the Heidke's skill score (Cohen's kappa; Heidke, 1926), the odd's ratio (Stephenson, 2000) and the odd's ratio skill score (Yule's Q; Yule, 1900). A review of these statistics for landslides have been provided by Beguería (2006); Guzzetti et al. (2006), Frattini et al. (2010).

Accuracy statistics require the splitting of the classified objects into a few classes by defining specific values of the susceptibility index that are called cut off values. For statistical models, a statistically significant probability cut off ($p_{cut\ off}$) exists, equal to 0.5. In different conditions (Van Den Eeckhaut et al., 2006), or for other typologies of landslide susceptibility models, such as physically-based, heuristic, artificial neural networks, and fuzzy logic, the choice of cut off values to define susceptibility classes is not scientifically sound. A solution to this limitation consists in evaluating the performance of the models over a large range of cut off values by using cut off-independent performance criteria such as success-rate curves, ROC curves and Cost curves.

5.2.1 Success rate curves

Success-rate curves (Vazquez-Selem and Zinck, 1994; Zinck et al, 2001; Chung and Fabbri, 2003) represent the percentage of correctly classified objects (i.e., terrain units) on the y-axis, and the percentage of area classified as positive (i.e., unstable) on the x-axis (Fig. 5.3). Chung and Fabbri (2003) distinguish between success- and prediction-rate curves. A success-rate curve is based on the comparison between the susceptibility map and the landslides used in the modeling (i.e. the training set). This is considered by Chung and Fabbri (2003) as a degree-of-fit measure. In contrast, prediction-rate curve provides the validation of the prediction regardless of the prediction model because is based on the comparison between the susceptibility map and a separated (spatially, temporally, or randomly) dataset of landslides.

In the landslide literature, the y-axis is normally considered as the number of landslides, or the percentage of landslide area, correctly classified. In the case of grid-cell units where landslides correspond to single grid cells and all the terrain units have the same area, the y-axis corresponds to true positive rate, in analogy with the ROC space, and the x-axis corresponds to the number of units classified as positive. A success rate curve is better than another if it is closer to the upper left corner.

Success-rate curves are normally obtained with the following procedure:

1. the susceptibility index is reclassified into two classes (stable and unstable) using a cut off value, and the percentage of area classified as unstable is calculated;
2. the resulting classified susceptibility map needs to be overlaid with landslides, in order to calculate either the percentage of landslides (as number or area) correctly laying within the area classified as unstable;
3. the total unstable area and the percentage of landslides within this area are plotted in the success-rate space.

By changing the values of the cut off, it is possible to obtain different points in the success-rate curve.

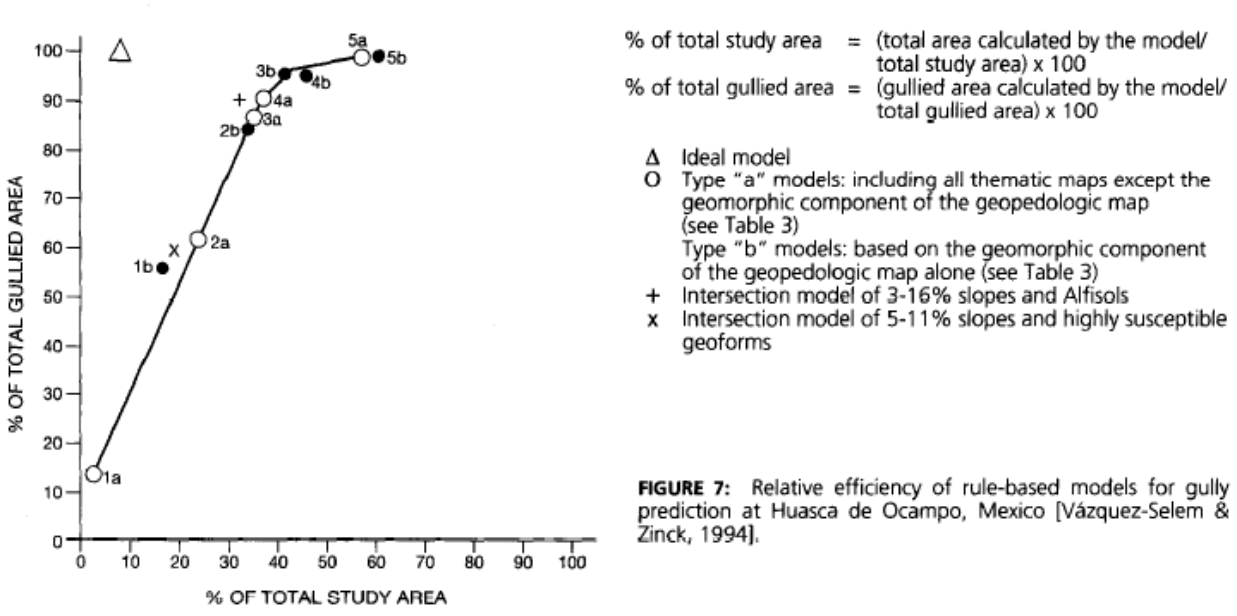


FIGURE 7: Relative efficiency of rule-based models for gully prediction at Huasca de Ocampo, Mexico [Vázquez-Selem & Zinck, 1994].

Figure 5.2.1 Example of success-rate curve for a susceptibility model of gullies (from Zinck et al, 2001)

5.2.2 ROC curves

ROC curves are applied in many fields to test model performance. The points on the ROC curve represent (FP,TP) pairs derived from different contingency tables created by applying different cut offs (TP = true positive rate; FP = false positive rate). Points closer to the upper-right corner correspond to lower cut off values. A ROC curve is better than another if it is closer to the upper left corner, and the range of values for which the ROC curve is better than the trivial model (i.e., the model represented by a straight line joining the lower-left and the upper-right corner) is defined operating range.

The Area Under the ROC Curve (AUC) can be used as a metric to assess the overall quality of a model (Hanley and McNeil, 1982): the larger the area, the best the performance of the model over the whole range of possible cut offs.

Operationally, ROC curves can be developed by extracting True Positive (TP) and False Positive (FP) rates from contingency tables associated to different cut off values for each model, using the following formulas:

$$\text{True positive rate (TP)= sensitivity} = \frac{tp}{tp + fn} = \frac{tp}{P} \qquad (5.2.1)$$

$$\text{False positive rate (FP) = 1-specificity} \frac{fp}{fp + tn} = \frac{fp}{N} \qquad (5.2.2)$$

For statistical models, this operation is automatically performed by commonly used statistical packages, such as SPSS and SAS.

For heuristic and physically-based models it is necessary to follow the following procedures:

1. the susceptibility index (derived from objective or subjective ranking or based on stability measures) is reclassified into two classes (stable and unstable) using a cut off value;
2. the resulting classified susceptibility map needs to be overlaid with landslides, in order to calculate the number of terrain units with or without landslides;
3. the contingency table can be derived for the specific cut off by intersecting the susceptibility classes and the presence/absence of landslides.

By changing the values of the cut off, it is possible to obtain different contingency tables which correspond to different points in the ROC curve (Fig. 5.2.2).

For statistical models, ROC curves are normally obtained considering the statistical units introduced into the analysis as training or validation sets. In case of grid-cell based models, this dataset could be a random sample of the population of grid-cells. In fact, due to the extremely high number of stable grid-cells with respect to landslides, it is a common practice to sample within the entire area a limited number of stable grid cells (e.g., equal or close to the unstable grid-cells). This is done in

order to improve the efficiency of statistical methods. Hence, the ROC curves calculated by for statistical model consider only a small part of the grid cells. This makes these curves conceptually different from those built for other type of models.

probability	0%		
0	1507	fp	1.000
0	1508	tp	1.000

probability	10%		
459	1048	fp	0.695
59	1449	tp	0.961

probability	20%		
834	673	fp	0.447
137	1371	tp	0.909

probability	30%		
1028	479	fp	0.318
203	1305	tp	0.865

probability	40%		
1148	359	fp	0.238
292	1216	tp	0.806

probability	50%		
1250	257	fp	0.171
413	1095	tp	0.726

probability	60%		
1325	182	fp	0.121
546	962	tp	0.638

probability	70%		
1390	117	fp	0.078
712	796	tp	0.528

probability	80%		
1448	59	fp	0.039
946	562	tp	0.373

probability	90%		
1494	13	fp	0.009
1238	270	tp	0.179

probability	100%		
1507	0	fp	0.000
1508	0	tp	0.000

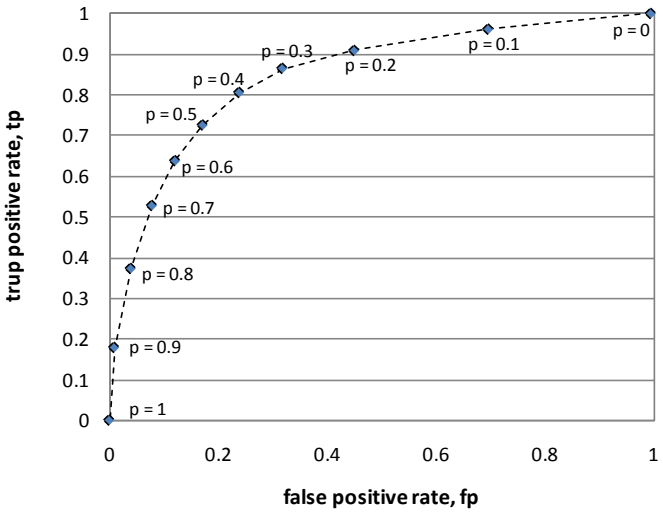


Figure 5.2.2 Example of contingency tables for different values (cut off) of membership probability of a landslide susceptibility discriminant model. For each contingency table the false positive rate (FP) and the true positive rate (TP) are calculated and plotted in the ROC space

5.2.3 Cost curves

The total cost of misclassification of a model depends on (Drummond and Holte, 2000): the percentage of terrain units that are incorrectly classified, the a-priori probability to have a landslide in the area, and the costs of misclassification of the different error types.

In order to explicitly represent costs in the evaluation of model performance, Drummond and Holte (2006) proposed the Cost curve representation. The Cost curve represents the Normalized Expected Cost as a function of a Probability-Cost function.

The Normalized Expected Cost, NE[C] is calculated as:

$$NE(C) = \frac{1 - TP \cdot p(+) \cdot c(-|+) + FP \cdot p(-) \cdot c(+|-)}{p(+) \cdot c(-|+) + p(-) \cdot c(+|-)} \qquad (5.2.3)$$

where the expected cost is normalized by the maximum expected cost, that occurs when all cases are incorrectly classified, i.e. when FP and FN are both one. The maximum normalized cost is 1 and the minimum is 0.

The Probability-Cost function, PC(+) is:

$$PC(+) = \frac{p(+)\cdot c(-|+)}{p(+)\cdot c(-|+) + p(-)\cdot c(+|-)} \quad (5.2.4)$$

which represents the normalized version of p(+)C(−|+), so that PC(+) ranges from 0 to 1. When misclassification costs are equal, PC(+) = p(+). In general, PC(+) = 0 occurs when cost is only due to negative cases, i.e., positive cases never occur (p(+) = 0) or their misclassification cost, C(−|+), is null. PC(+) = 1 corresponds to the other extreme, i.e., p(−) = 0 or C(+|−) = 0.

A single classification model, which would be a single point (FP, TP) in ROC space, is a straight line in the Cost curve representation (Fig. 5.5). A set of points in ROC space, the basis for an ROC curve, is a set of Cost lines, one for each ROC point.

Taking into consideration that a single point (FP, TP) in ROC space results in a straight line, with coordinates (0, FP) and (1, FN), in Cost space, Cost curves are implemented through the following steps:

1. a straight line is calculated for each point of the ROC curves, with function NE(C) = FP + (FN – FP) PC(+), where NE(C) is the Normalized Expected Cost and PC(+) is the Probability-Cost function;
2. for small increments of PC(+), the values of the Normalized Expected Cost are calculated for all the linear functions;
3. for each increment, the minimum expected cost is selected;
4. the selected minima are joined to trace the cost curve, which represents the lower envelope of the straight lines.

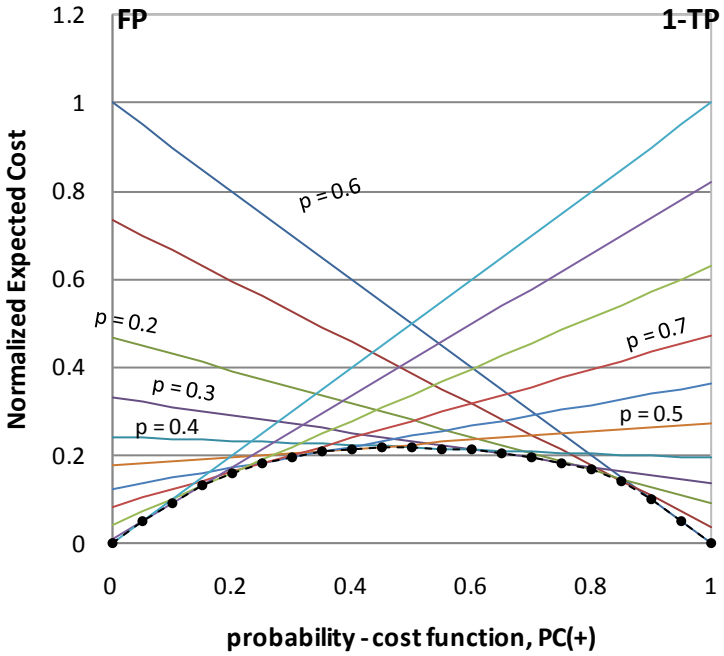


Figure 5.2.3 Example of cost curve construction starting from ROC data of Figure 5.2.2.The cost curve is represented by the black thick line, and represents the minimum envelope of cost lines associated to different cut offs

Once the cost curve has been prepared, the quality of the susceptibility model is assessed in terms of Normalized Expected Cost, given a specific value of the Probability-Cost function. In general, the lower the cost curve, the better the performance, and the difference between two models is simply the vertical distance of the curves. As mentioned, the value of the Probability-Cost function depends on both the a-priori probability and the misclassification costs. For landslide studies, given the uncertainty in the observed distribution of population, a condition of equal-probability can be assumed. Regarding the misclassification cost, a simple analysis of costs based on land-use maps can be performed.

Once a value of Probability-Cost function is assigned, the optimal cut off is calculated for each model by selecting the lower straight line in the Cost space (Fig. 5.5). In fact, this straight line corresponds to the contingency table of the optimal cutoff, which minimizes the Normalized Expected Cost for that particular combination of a-priori probability and misclassification costs.

Success rate curves, ROC curves and Cost curves can be created for both the training and validation dataset. A good model should not only have a ROC curve with high AUC and a low cost curve for the training set. Curves obtained for the validation set should be located close to those of the training set.

Finally, as mentioned in section 2.3, success rate curves and ROC curves were already suggested as statistical methods to select the best model out of a set of models using different combinations of thematic variables or different modeling approaches.

5.3 OTHER VALIDATION METHODS

Guzzetti et al. (2006) suggested the use of success-rate and ROC curves. However, they further proposed a framework for evaluating the model reliability and prediction skill of statistical landslide susceptibility models. They prepared an ensemble of 350 landslide susceptibility models using the same landslide and thematic information but for different numbers of mapping units in the training and validation sets. More specifically, for training sets with 30%, 45%, 55%, 65%, 75%, 85% and 95% mapping units, each time 50 models were trained. They were validated using the remaining 70%, 55%, 45%, 35%, 25%, 15% and 5% mapping units. Analysis of this ensemble contributed to understanding the role and stability of the thematic variables used to construct the model. The most important thematic layers should be included in (almost) all 350 models. It also provided insight in the model sensitivity to changes in the input data. For each mapping unit the average probability and standard deviation was calculated for the models calibrated from the 50 training sets with 85% mapping units. The standard deviation of a mapping unit can be used as an estimate of the error associated with the susceptibility assessment determined for this mapping unit. It was therefore suggested to show this error estimate on a map that complements the landslide susceptibility map. This is illustrated in Figure 5.2.4.

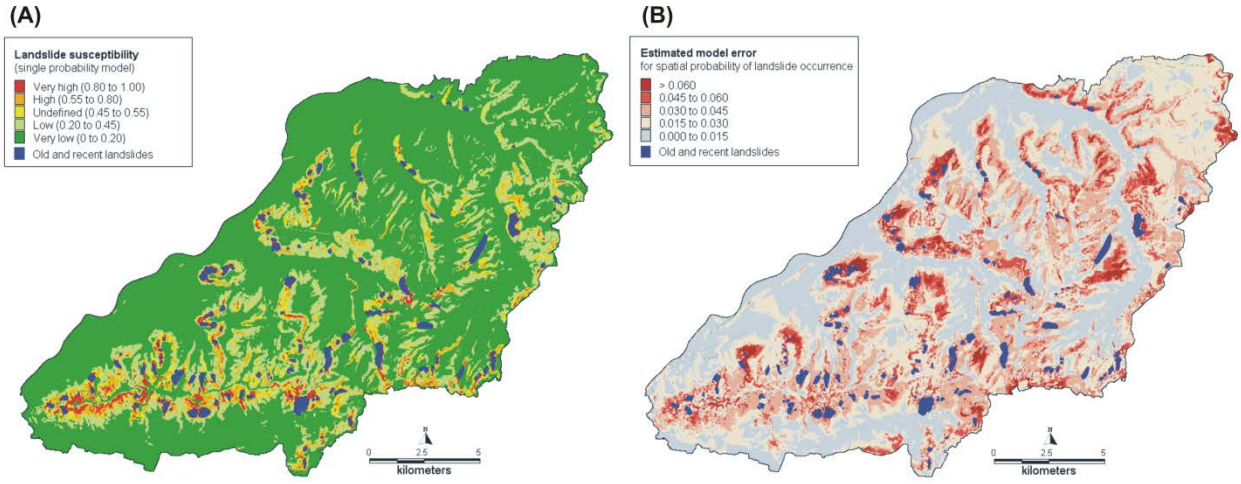


Figure 5.2.4 (A) Landslide susceptibility map of a catchment in Northern Belgium; (B) Distribution of model error (standard deviation) in the landslide susceptibility map calculated from an ensemble of 50 models using 80% of the dataset for model calibration and the 20% remaining data for validation (Van Den Eeckhaut et al., 2009).

Finally, in the absence of landslide inventory maps to evaluate the quality of the produced model, some studies on national or continental landslide susceptibility modelling have carried out local evaluations the produced landslide susceptibility map. Malet et al. (subm.), for example, evaluated the quality of the nation-scale landslide susceptibility map for France with regional or local susceptibility maps available in the literature or created for planning purposes, and Van Den Eeckhaut et al. (subm.) evaluated a landslide susceptibility map of Europe by visual comparison

with published national landslide inventory and landslide susceptibility maps of e.g. Germany, Italy, Ireland, Norway, Poland, Romania and UK. Of course, a more detailed validation is needed, but currently this is not possible due to the lack of a European landslide database.

REFERENCES

Adler RF, Huffman GJ, Bolvin DT, Curtis S, Nelkin EJ (2000) Tropical Rainfall Distributions Determined Using TRMM Combined with Other Satellite and Rain Gauge Information, Journal of Applied Meteorology, 39(12), 2007-2223

Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. Bull. Eng. Geol. Env., 58, 21-44.

Allison, P.D., 2001. Logistic Regression Using the SAS system: Theory and Application. Wiley Interscience, New York

American Institute of Aeronautics and Astronautics, (1998). Guide for the Verification and Validation of Computational Fluid Dynamics Simulations, AIAA-G-077-1998, Reston, VA

Anderson MG, Lloyd DM (1991) Using a combined slope hydrology/stability model to develop cut slope design charts. Proc. Inst. Civil Engineers, 91(2): 705-718

Andriola P., Chirico G. B., De Falco M., Di Crescenzo G., Santo A. (2008) - Assessing flowslides triggering susceptibility in the pyroclastic deposits of the carbonatic contest: a comparison between physically based models and a statistic-semiquantitative methodology. Volume speciale "Innovative approaches for evaluating landslide hazard and risk" della rivista Geografia e Dinamica Quaternaria, vol. 32/2.

Apel H., Thieken A.H., Merz B., Blöschl G. (2004). Flood risk assessment and associated uncertainty. Natural Hazards and Earth System Sciences, 4, pp. 295-308.

Apostolakis G. (1990). The Concept of Probability in Safety Assessments of Technological Systems. Science, 250, pp.1359-1364

Atkinson, P.M., Massari, R., (1998). Generalised linear modelling of susceptibility to landsliding in the Central Apennines. Computers & Geosciences 24, 373-385.

Ayalew, L., Yamagishi, H., (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65, 15–31.

Baecher G B, Christian J T (2003) Reliability and statistics in geotechnical engineering John Wiley and sons

Baeza, C., Corominas, J., (2001). Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. Earth Surface Processes and Landforms 26 (12), 1251–1263.

Baeza, C.; Lantada, N and Moya, J. (2010a). Influence of sample and terrain unit on landslide susceptibility assessment at La Pobla de Lillet, Eastern Pyrenees, Spain. Environmental Earth Science 60:155-167.

Baeza, C; Lantada, N and Moya, J.(2010b). validation and evaluation of two multivariate statistical models for predictive shallow landslide susceptibility mapping of the Eastern Pyrenees (Spain). Environmental Earth Science 61:507-523.

Baillifard F., Jaboyedoff M., Sartori M. (2003). Rockfall hazard mapping along a mountainous road in Switzerland using a GIS-based parameter rating approach. *Natural Hazards and Earth System Sciences*, **3**, pp. 431–438.

Barredo, J., Benavides, A., Hervas, J. and van Westen, C.J.,(2000). Comparing heuristic landslide hazard assessment techniques using GIS in the Tirajana basin, Gran Canaria Island, Spain. International journal of applied earth observation and geoinformation 2, 9-23

Baum R L, Savage W Z, Godt J W (2002) TRIGRS – A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis. U.S. G.S. Open-File Report 02-0424, 27

Baum R L, Savage W Z, Godt J W (2008) TRIGRS-A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0. US Geological Survey Open-File Report 2008–1159. Available via: http://pubs.usgs.gov/of/2008/1159/

Belsley, D.A., Kuh, E., Welsch, R.E., (1980). Regression Diagnostics: Identifying influential data and sources of collinearity. John Wiley, New York.

Bonham-Carter, G.F., (1994). Geographic Information System for Geoscientists: Modelling with GIS. Computer Methods in the Geosciences 13. Pergamon, New York.

Borga M, Dalla Fontana G, Gregoretti C, Marchi L (2002) Assessment of shallow landsliding by using a physicallybased model of hillslope stability Hydrol. Process. 16: 2833–2851Beguería, S., 2006 Validation and evaluation of predictive models in hazard assessment and risk management. Natural Hazards 37(3), 315-329.

Box, G. E. P. and Andersen, S. L. (1955). Permutation theory in the derivation of robust criteria and the study of departures from assumption (with discussion). J. Roy. Statist. Soc. Ser. B **17:** 1–34.

Brabb, E.E., Pampeyan, E.H., Bonilla, M.G., (1972). Landslide susceptibility in San Mateo County, California. U.S. Geol. Surv., Misc. Field Studies, Map MF-360. Scale 1:62,500.

Brabb, E.E., (1984). Innovative approaches to landslide hazard mapping. 4th International Symposium on Landslides, Toronto, Vol 1, pp. 307–324.

Briggs, W.M., Ruppert, D., (2005). Assessing the skill of yes/no predictions. Biometrics 61, 799-807.

Budetta P. (2004). Assessment of rockfall risks along roads. *Natural Hazards and Earth System Sciences*, **4**(1), pp. 71-81.

Burton A, Arkell TJ, Bathurst JC (1998) Field variability of landslide model parameters. Environ Geol 35(2–3):100–114

Caine N (1980) The rainfall intensity duration control of shallow landslides and debris flow. Geogr. Annaler, 62:23-37.

Canuti P, Casagli N (1996) Considerazioni sulla valutazione del rischio di frana. Estratto da “Fenomeni franosi e centri abitati”, Atti del convegno di Bologna del 27 maggio 1994

Carrara, A.,(1983). Multivariate models for landslide hazard evolution. Mathematical Geology 15, 403–427.

Carrara, A., Cardinali, M., Guzzetti, F., Reichenbach, P.,(1995). GIS technology in mapping landslide hazard. In: Carrara, A., Guzzetti, F. (Eds.), Geographical Information Systems in Assessing Natural Hazards. Kluwer Academic Publishers, pp. 135–175.

Carrara A., Crosta G.B**.,** Frattini P., (2008). Comparing models of debris-flow susceptibility in the alpine environment. Geomorphology, 94, 353-378.

Carrara A, Cardinali M, Guzzetti F, Reichenbach P. (1995). GIStechnology in mapping landslide hazard. In: Carrara A, GuzzettiF (eds) Geographical information systems in assessing natural hazards. Kluwer, Dordrecht, pp 135–175

Carslaw H S, Jaeger J C (1959) Conduction of heat in solids. Oxford Univ Press, Oxford, UK

Cascini, L. (2005). Risk assessment of fast landslide–From theory to practice. General Report. In: Picarelli, L. (Ed.), Proceedings of the International Conference on “Fast Slope Movements – Prediction and Prevention for Risk Mitigation”. Pàtron Editore, Bologna, Vol. 2, pp. 33-52.

Cascini L., Cuomo S., Ferlisi S., Sorbino G.(2009). Detection of mechanisms for destructive landslides in Campania region – southern Italy. In: Picarelli, L., Tommasi, P., Urciuoli, G., Versace, P. (Eds.) “Rainfall-induced landslides: mechanisms, monitoring techniques and nowcasting models for early

warning systems". Proceedings of the first Italian workshop on Landslides. Eds. Naples, 8-10 June 2009, vol. 1, pp. 43-51.

Catani, F, Casagli N, Ermini L, Righini G, Menduni G (2005), Landslide hazard and risk mapping at catchment scale in the Arno River basin, Landslide, 2(4), p329-342, DOI: 10.1007/s10346-005-0021-0

Catani, F., S. Segoni, and G. Falorni (2010), An empirical geomorphology-based approach to the spatial prediction of soil thickness at catchment scale, Water Resour. Res., 46, W05508, doi:10.1029/2008WR007450.

Chan, Y.H. (2003). Bioestatistics I02: quantitative data-parametric & Non-parametric tests. Singapore Med J. 44(8):391-396.

Chelbroad A F (2000) A method for anticipating the occurrence of precipitation-induced landslides in Seattle, Washington U.S. Geological Survey Open-File Report 00-469Chung, C.F., Fabbri, A.G., 1999. Probabilistic Prediction Model for Landslide Hazard Mapping. Photogrammetric Engineering and Remote Sensing 65, 1389–1399.

Cheng, K, Wei, C, & Chang, S (2004) Locating landslides using multi- temporal satellite images. Advances in Space Research, 33, 296 – 301

Chung, C.F., Fabbri, A.G., (2003). Validation of spatial prediction models for landslide hazard mapping. Natural Hazards 30 (3), 451-472.

Chung, C.J.F., Kojima, H., Fabbri, A.G., (2002). Stability analysis of prediction models for landslide hazard mapping. In: R.J. Allison (Ed.), Applied Geomorphology: Theory and Practice. John Wiley and Sons, Ltd, New York.

Claessens L, Heuvelink G B M, Schoorl J M, Veldkamp A (2005) DEM resolution effects on shallow landslide hazard and soil redistribution modelling Earth Surf. Process. Landforms 30: 461–477

Coles S. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer, London.

Coles S., Pericchi L.R., Sisson S. (2003). A fully probabilistic approach to extreme rainfall modeling. Journal of Hydrology, 273(1), pp. 35-50.

Corominas, J. and Mavrouli, O. (Editors) (2011). Guidelines for landslide susceptibility, hazard and risk assessment and zoning. Deliverable D2.4. SafeLand Project.

Crosta G (1998) Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. Environmental Geology, 35: 131-145

Crosta G, Frattini P (2001) Rainfall thresholds for triggering soil slips and debris flow. Proc. of the 3rd Plinius Conference 2001.

Crosta GB, Dal Negro P (2003) Observations and modelling of soil slip-debris flow initiation processes in pyroclastic deposits: the Sarno 1998 event. Natural Hazards and Earth System Sciences, 3: 53-69

Crosta, G.B., Frattini, P., (2003). Distributed modeling of shallow landslides triggered by intense rainfall. Natural Hazards and Earth System Sciences 3 (1–2), 81–93.

Crosta G.B., Frattini P., Valbuzzi E. (2007) Historical datasets - phase II. DEliverable 90 of the LESSLOSS EU project. www.lessloss.org.

Crozier MJ (1986) Landslides: Causes, Consequences and Envirorment. Croom Helm, London, pp. 252.

Dai, C.F., Lee,C.F. (2001). Terrain based mapping of landslide susceptibility using a geographic information system: a case study. Canadian Geotechnical Journal 38, 911–923.

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

Dewitte, O., Chung, C.J., Cornet, Y., Daoudi, M., Demoulin, A., 2010. Combining spatial data in landslide reactivation susceptibility mapping: A likelihood ratio-based approach in W Belgium. Geomorphology 122, 153-166.

Dillon WR, Goldstein M. 1986. Multivariate Analysis. Methods and Applications. John Wiley: Chichester.

Diniz-Fiho, J.A.F.; Bini, L.M. and Hawkins, B.A. 2003. Spatial autocorrelation and red herrings in geographical ecology. Global ecology & Biogeografy 12:53-64.

Drummond, C., Holte, R.C., 2000. Explicitly representing expected cost: an alternative to ROC representation. Proc. of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 198–207.

Drummond, C., Holte, R.C., 2006. Cost Curves: An Improved Method for Visualizing Classifier Performance. Machine Learning 65 (1), 95-130.

Dietrich W E, Montgomery D R (1998) SHALSTAB: A digital terrain model for mapping shallow landslide potential, http://socrates.berkeley.edu/_geomorph/shalstab.

Duan J, Grant G E (2000) Shallow Landslide Delineation for Steep Forest Watersheds Based on Topographic Attributes and Probability Analysis, in: Terrain Analysis – principles and applications,edited by: Wilson, J. P. and Gallant, J. C., Wiley, 311–329.

Dussauge-Peisser, C., Helmstetter, A., Grasso, et al. (2002). Probabilistic approach to rock fall hazard assessment: potential of historical data analysis, *Natural Hazards and Earth System Sciences*, **2**, pp. 15-26.

Duzgun H.S.B., Yucemen M.S., Karpuz, C., (2003). A Methodology for Reliability-Based Design of Rock Slopes. Rock Mechanics and Rock Engineering, 36(2), pp. 95-120.

Ercanoglu, M., Gokceoglu, C., 2002. Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. EnvironmentalGeology 41 (6), 720–730.

Faber MH, Stewart MG (2003). Risk assessment for civil engineering facilities: critical overview and discussion. Reliability Engineering and System Safety, 80, pp. 173-184.

Fell, R., Corominas, J., Bonnard, Ch., Cascini, L., Leroi, E., Savage, W.Z. on behalf of the JTC-1 Joint Technical Committee on Landslides and Engineered Slopes. 2008a. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology, 102, pp. 85-98.

Fell, R., Corominas, J., Bonnard, Ch., Cascini, L., Leroi, E., Savage, W.Z. on behalf of the JTC-1 Joint Technical Committee on Landslides and Engineered Slopes. 2008b. Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning. Commentary. Engineering Geology, 102, pp. 99-111.

Finley, J.P., 1884. Tornado predictions. American Meteorological Journal 1, 85–88.

Frattini P., Crosta G.B., Carrara A. (2010) Techniques for evaluating the performance of landslide susceptibility models. Engineering Geology, 111(1-4), 62-72.

Fischer, H. 2010. A History of the Central Limit Theorem: From Classical to Modern Probability Theory, Springer.

Freeman G T (1991) Calculating catchment area with divergent flow based on a regular grid, Comput. Geosci., 17, 413–422.

Gallant J C Wilson J P (2000) Primary Topographic Attributes, in:Terrain Analysis – principles and applications, edited by: Wilson,J. P. and Gallant, J. C., Wiley, 51–85.

Garcia, R.A.C.; Zêzere, J.L and Oliveira, S.C. 2008. The influence of terrain units in landslide susceptibility assessment: a case study in the Abadia Basin (Portugal). Geophysical Research Abstracts, Vol. 10, EGU2008-A-07486.

Gardner W R (1958) Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. Soil Sci 85:228–232.

Gibbons, J. D. and Chakraborti, S. 2003. Nonparametric Statistical Inference. 4th Ed. CRC. 680 pp.

Gilbert, G.F., 1884. Finley's tornado predictions. American Meteorological Journal 1, 166-172.

Godt J W, Schulz W H, Baum R L, Savage, W Z, (2008) Modeling rainfall conditions for shallow landsliding in Seattle, Washington, in Baum, R.L., Godt, J.W., and Highland, L.M., eds., Landslides and Engineering Geology of the Seattle, Washington, Area: Geological Society of America Reviews in Engineering Geology, v. XX, p. 137–152, doi: 10.1130/2008.4020(08).

Gökceoglu, C., Aksoy, H., 1996. Landslide susceptibility mapping of the slopes in the residual soils of the Mengen region (Turkey) by deterministic stability analyses and image processing techniques. Engineering Geology 44, 147–161.

Guisan, A. and Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147-155.

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., Cardinali, M., 2005. Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. Natural Hazards and Earth System Sciences 6, 115–131.

Guzzetti F, Perruccacci S, Rossi M (2005) Definition of critical thresholds for different scenarios. RISK-Advanced Wheather Forecast System to Advise on Risk and Management

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology 81 (1–2), 166–184.

Hall J.W. (2003). Handling uncertainty in the hydroinformatic process. *Journal of Hydroinformatics*, **5**(4), pp. 215-231.

Hammond C, Hall D, Miller S, Swetik P (1992) Level I Stability Analysis (LISA) documentation for version 2.0, Gen. Tech. Rep. INT-285, For Serv. U.S. Dep. Of Agric., Ogden, Utah.

Haneberg W C (2000) Deterministic and probabilistic approaches to geologic hazard assessment Environmental & Engineering Geoscience, 6:209–226.

Haneberg W C (2004) A rational probabilistic method for spatially distributed landslide hazard assessment Environmental & Engineering Geoscience, 10:23-47, (doi: 10.2113/10.1.27).

Haneberg W C (2007) PISA-m Map-Based Probabilistic Infinite Slope Analysis Reference Manual Haneberg Geoscience

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143 (1), 29–36.

Hansen MJ. 1984. Strategies for classification of landslides. In Slope Instability, Brunsden D (ed.). John Wiley: Chichester; 1–25.

Hanssen, A.W., Kuipers, W.J.A., 1965. On the relationship between the frequency of rain and various meteorological parameters. Mededelingenen Verhandelingen 81, 2–15.

Heidke, P., 1926. Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungdienst (Calculation of the success and goodness of strong wind forecasts in the storm warning service). Geografika Annaler 8, 301–349.

Helton J.C., Burmaster D.E. (1996). Guest editorial: treatment of aleatory and epistemic uncertainty in performance assessments for complex systems. *Reliability Engineering & System Safety*, **54**(2-3), pp. 91-94.

Hervas J, Barredo J, Rosin P, Pasuto A, Mantovani F, Silvano S (2003). Monitoring landslides from optical remotely sensed imagery: The case story of Tessina landslide, Italy. Geomorphology, 54, 63 – 75.

Hervàs J, Bobrowsky P (2009) Mapping inventories, susceptibility, hazard and risk In: Sassa K & Canuti P Landslides: Disaster Risk Reduction

Hjort,J and Marmion, M. 2008. Effects of sample size on the accuracy of geomorphological models. Geomorphology 102:341-350.

Hong, Y, Adler R, Huffman G (2007b), An Experimental Global Monitoring System for Rainfall-triggered Landslides using Satellite Remote Sensing Information, IEEE TGRS, DOI: 10.1109/TGRS.2006.888436

Hosmer, D.W., Lemeshow, S., 1989. Applied Regression Analysis. Wiley, New York.

Hsieh P A., Wingle W Healy R W (2000) VS2DI—A Graphical Software Package for Simulating Fluid Flow and Solute or Energy Transport in Variably Saturated Porous Media: U.S. Geological Survey Water-Resources Investigation Report 99-4130, 16 p.

Huang J C, Kao S J, Hsu M L Liu Y A Influence of Specific Contributing Area algorithms on slope failure prediction in landslide modeling Nat. Hazards Earth Syst. Sci., 7: 1–12,

Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB (2007), The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. J. Hydrometeor., 8(1), 38-55

Huggel C, Kaab A, Haeberli W, Teysseire Paul F (2002). Remote sensing based assessment of hazards from glacier lake outbursts: A case study in the Swiss Alps. Canadian Geotechnical Journal, 39, 316 – 330.

Hungr, O. Evans, S. G., and Hazzard, J., (1999). Magnitude and frequency of rock falls and rock slides along the main transportation corridors of southwestern British Columbia, Can. Geotech. J., 36, pp. 224-238.

Hungr, O., Corominas, J., Eberhart, E – Estimating landslide motion, travel distance and velocity - State of the Art report. In: Hungr, O., Fell, R., Couture & Eberhardt, E (Eds.), Landslide Risk Management. Proceedings of the International Conference on Landslide Risk Management, pp. 99-128, A.A. Balkema Publishers.

Israel GD.2008. Determining sample size. University of Florida IFAS extension. Publication #PEOD6.

Iverson R M (2000) Landslide triggering by rain infiltration. Water Resour Res 36(7):1897–1910

Jimenez-Rodriguez R., Sitar N., Chacón J. (2006). System reliability approach to rock slope stability, *International Journal of Rock Mechanics and Mining Sciences*, **43**(6), pp. 847-859.

Kaab A (2002). Monitoring high-mountain terrain deformation from repeated air- and spaceborne optical data: examples using digital aerial imagery and ASTER data ISPRS, Journal of Photogrammetry and Remote Sensing, 57(1-2), 39-52

Kaab A, Wessels R, Haeberli W, Huggel C, Kargel J, Khalsa, S. (2003). Rapid ASTER Imaging facilities timely hazard assessment of glacier hazards and disasters. EOS, Transactions American Geophysical Union, 84(13), 117 – 124.

Kendall, M., Stuart, A., 1979. The Advanced Theory of Statistics: Inference and Relationship. Griffin, London.

Kiang, M, 2003. A comparative assessment of classification methods. Decision suport systems 35:441-454.

Kniveton, D, de Graff P, Granicas K, Hardy R (2000). The development of a remote sensing based technique to predict debris flow triggering conditions in the French Alps. International Journal of Remote Sensing, 21(3), 419 – 434.

Lee S, Choi J, Min K. 2002. Landslide susceptibility analysis and verification using the Bayesian probability model. Environ Geol 43:120–131.

Lee, S., Min, K., 2001. Statistical analysis of landslide susceptibility at Yongin, Korea. Environmental Geology 40, 1095–1113.

Liu, Canran; Frazier, P and Kumar Lalit. 2007. Comparative assessment of the measures of thematic classification accuracy. Remote sensing of environment 107:606-616.

Long, S. And Freese, J. 2001. Regrssion models for categorical dependent variables using Stata. Stata Press Pub. Texas, 311pp.

Luzi L. 1995. GIS for slope stability zonation in the Fabriano area, central Italy PhD thesis, International Institute for Aerospace Survey and Earth Sciences (ITC), Enschede.

Malet, J.-P., Puissant, A., Mathieu, A., Van Den Eeckhaut, M., Fressard, M., subm. Landslide susceptibility assessment at 1:1M scale for France. Landslides.

Mann C.J. (1993). Uncertainty in geology. In: Davis J.C. ,Herzfeld U.C. (eds.), Computers in geology – 25 years of progress, Oxford Univ. Press, p. 241–254.

Meijerink AMJ. 1988. Data acquisition and data capture through terrain mapping units. ITC J 7:23–44.

Mercogliano P. , Schiano P., Picarelli L., Olivares L., Catani F., Tofani V., Segoni S., Rossi G. (2010). Short term weather forecasting for shallow landslide prediction. In "Mountain Risks: bringing Science to Society", Cerg Editions, pp. 525 - 530. Proceedings of the Mountain Risks International Conference, Firenze, Italy 24-26 November 2010

Metternicht, G, Hurni L, Gogu R (2005) Remote sensing of landslides: An analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments. Remote Sensing of Environment 98 (2005) 284 – 303.

Meusburger, K., Alewell, C., 2009. On the influence of temporal change on the validity of landslide susceptibility maps. Nat. Hazards Earth Syst. Sci. 9, 1495-1507.

Montgomery, D.R., Dietrich, W.E., 1994. A physically based model for the topographic control on shallow landsliding. Water Resources Research 30 (4), 1153–1171.

Montgomery D R, Greenberg H M, Laprade W T, Nashem W D. 2001. Sliding in Seattle: Test of a model of shallow landsliding potential in an urban environment, in Wigmosta, M.S., and Burges,S.J., eds., Land Use and Watersheds: Human Influence on Hydrology and Geomorphology in Urban and Forest Areas: Washington, D.C.,American Geophysical Union, Water Science and Application, v. 2,p. 59–73.

Mucherino, A., Papajordgji, P.J., Pardalos, P.M., 2009. Data mining in agriculture. Springer-Verlag.

Murphy, A.H., 1996. The Finley affair: A signal event in the history of forecast verification. Weather Forecasting 11, 3-20.Nichol J, Wong MS (2005) Satellite remote sensing for detailed landslide inventories using change detection and image fusion. International journal of Remote Sensing, 26:9, 1913-1926. DOI:10.1080/01431160512331314047.

O'Callaghan J F, Mark D M (1984) The extraction of drainage networks from digital elevation data, Computer Vision, Graphics and Image Processing, 28, 323–344..

Pack R T, Tarboton D G, Goodwin C N (1998) A Stability Index Approach to Terrain Stability Hazard Mapping. SINMAP User's Manual , 68 pp.

Pastor, M. 2010. Landslide runout: Review of analytical/empirical models for subaerial slides, submarine slides and snow avalanche. Numerical modelling. Software tools, material models, validation and benchmarking for selected case studies. Deliverable D1.7. SafeLand Project.

Pasuto, A. and Soldati, M. 1999. The use of landslide units in geomorphological mapping: an example in the Otalian Dolomites. Geomorphology 30:53-64.

Peirce, C.S., 1884. The numerical measure of the success of predictions. Science 4, 453-454.

Phoon K K (2008) Reliability-based design in geotechnical engineering: computation and application Taylor Francis (London & New York) 530 pp.

Poli, S., Sterlacchini, S., 2007. Landslide Representation Strategies in Susceptibility Studies using Weights-of-Evidence Modeling Technique. Natural Resources Research 16, 121-134.

Powell R. L. Matzke, N. de Souza, Jr.; Clark, M.; Numata, I.; hess, L.L. and Roberts, D.A. 2004. Source of error in accuracy assessment of thematic land-cover maps in Brazilian Amazon. Remote sensing of environment 90:221-234.

Quinn P, Beven K, Chevallier P, Planchon O (1991) The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. Hydrological Processes 5(1), 59–79.

Raines, G.L., 1999. Evaluation of weights of evidence to predict epithermal-gold deposits in the great basin of the western United States: Natural Resources Research 8, 257-276.

Reinartz, P., Lehner, M., Müller, R., Schroeder, M., 2004. Accuracy for DEM and Orthoimages derived from SPOT HRS Stereo Data without using GCP. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 35 (Part B1), pp. 433-438.

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A., Peruccacci, S., 2009. Optimal landslide susceptibility zonation based on multiple forecasts. Geomorphology 114, 129-142.

Rosso R, Rulli M C, Vannucchi G (2006) A physically based model for the hydrologic control on shallow landsliding Water Resour. Res., 42, W06410, doi:10.1029/2005WR004369

Salciarini D, Godt J W, Savage W Z, Conversini P, Baum R L, Michael J A (2006) Modeling regional initiation of rainfall-induced shallow landslides in the eastern Umbria region of central Italy. Landslides 3:181–194

Santacana, N., Baeza, B., Corominas, J., De Paz, A., Marturia, J., 2003. A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet area (Eastern Pyrenees, Spain). Natural Hazards 30, 281-295.

Savage W Z, Morrissey M M Baum R.L (2000) Geotechnical Properties for Landslide-Prone Seattle-Area Glacial Deposits: U.S.Geological Survey Open-File Report 00-228, 5 p., http://pubs.er.usgs.gov/usgspubs/ofr/ofr00228.

Savage W Z, Godt J W, Baum R L (2003) A model for spatially and temporally distributed shallow landslide initiation by rainfall infiltration. In: Proceedings of 3rd international conference on debris flow hazards mitigation: mechanics, prediction, and assessment, Davos, Switzerland, 10–12 September 2003:179–187.

Savage W Z, Godt J W, Baum R L (2004) Modeling time-dependent areal slope stability. In: Lacerda WA, Erlich M, Fontoura SAB, Sayao ASF (eds) Landslides-evaluation and stabilization, Proceedings of 9th International symposium on Landslides, vol 1. Balkema, Rotterdam, pp 23–36

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather Forecasting 5, 570-575.

Schubert M., Straub D., Faber M.H. (2005). Reliability of rock fall protection galleries – A case study with a special focus on the uncertainty modelling. *Safety and Reliability of Engineering Systems and Structures* (Proc. ICOSSAR 05, Rome), Augusti et al. (eds), Millpress, pp. 1333-1340.

Schmidt, J., Andrew, R., 2005. Multi-scale landform characterization. Area 37, 341-350.

Schmidt K M, Roering J J, Stock J D, Dietrich W E, Montgomery D R, Schaub T (2001) The variability of root cohesion as an influence on shallow landslide susceptibility in the Oregon Coast Range Can. Geotech. J. 38: 995–1024

Schulz W H (2003) Modeling distribution of colluvium and shallow groundwater on landslide-susceptible slopes for parts of Seattle, Washington: Geological Society of America Abstracts with Programs, v. 35, no. 6,p. 216.

Segoni, S., Leoni, L., Benedetti, A. I., Catani, F., Righini, G., Falorni, G., Gabellani, S., Rudari, R., Silvestro, F., and Rebora, N.: Towards a definition of a real-time forecasting network for rainfall induced shallow landslides, Nat. Haz-ards Earth Syst. Sci., 9, 2119-2133, doi:10.5194/nhess-9-2119-2009

Sidle R C (1992) . A theoretical model of the effects of timber harvesting on slope stability. Water Resources Research, 28: 1897–1910.

Singhroy V (1995) Sar integrated techniques for geohazard assessment. Advances in Space Research, 15(11): 67-78

Singhroy V (2002). Landslide hazards: CEOS, The use of earth observing satellites for hazard support: Assessments and scenarios. Final report of the CEOS Disaster Management Support Group, NOAA, p. 98.

Singhroy V, Mattar K, Gray A (1998). Landslide characterisation in Canada using interferometric SAR and combined SAR and TM images. Advances in Space Research, 21(3), 465 – 476.

Singhroy, V. and Molch, K. (2004) Characterizing and monitoring rockslides from SAR techniques. Advances in Space Research, Vol. 33 #3, p. 290-295

Soeters, R., van Westen, C.J., 1996. Slope instability recognition, analysis and zonation. In: Turner, A.K., Schuster, R. (Eds.), Landslides Investigation and Mitigation. Transportation Research Board Special report, 247. National academy press, Washington DC.

Srivastava R, Yeh T C J (1991) Analytical solutions for one-dimensional, transient infiltration toward the water table in homogeneous and layered soils. Water Resour Res 27:753–762

Stehman, S.V. and Czaplewski R L. 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. Remote Sensing of Environmental 64:331-344.

Stephenson, D.B., 2000: Use of the “Odds Ratio” for Diagnosing Forecast Skill. Weather Forecasting 15, 221–232.

Stillwater Sciences. (2007) Landslide Hazard in the Elk River Basin, Humboldt County, California. Final report. Prepared by Stillwater Sciences, Arcata, California for the North Coast Regional Water Quality Control Board.

Straub D., Der Kiureghian A. (2007). Improved Seismic Fragility Modeling from Empirical Data. Structural Safety, in print, doi:10.1016/j.strusafe.2007.05.004.

Süzen, M and Doyuran, V. 2004. Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. Engineering Geology 71:303-321.

Tabachnick, B.G., and Fidell, L.S. 2001. Using Multivariate Statistics. 4th Ed. New York: Harper & Row. 582 – 633.

Tarboton D G (1997) A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models Water Resour. Res., 33(2), 309–319.

Terlien MTJ (1998) The determination of statistical and deterministic hydrological landslide-triggering thresholds. Environmental geology , 35(2-3): 124-130.

Tofani V, Dapporto S , Vannocci P, Casagli N (2006) Infiltration, seepage and slope instability mechanisms during the 20–21 November 2000 rainstorm in Tuscany, central Italy, Natural Hazards and earth System Sciences, 6, 1025-1033.

Toutin T, (2003) Error tracking in IKONOS geometric processing using 3D parametric model, Photogrammetric Engineering and Remote Sensing 69, 43-51.

Toutin T, (2004) Comparison of stereo extracted DTM from different high resolution sensors: SPOT-5, EROS-A, IKONOS-II and Quickbird. IEE Transactions on Geoscience Remote Sensing 42, 2121-2129.

Toutin T., 2006. Generation of DMs from Spot-5 in track HRS and across-track HRG stereo data using spatiotriangulation and autocalibration. ISPRS Journal of Photogrammetry and Remote Sensing, 60(3), pp. 170-181.

Vanacker, V., Vanderschraeghe, M., Govers G., Willems, E., Poesen, J., Deckers, J., De Bievre, B., 2003. Linking hydrological, infinite slope stability and land-use change models through GIS for assessing the impact of deforestation on slope stability in high Andes watersheds. Geomorphology 52, 299-315.

Vanapalli S K, Fredlund, D G, (2000) Comparison of different procedures to predict unsaturated soil shear strength in Shackelford, C.D., Houston, S.L., and Chang N.-Y., eds., Advances in unsaturated geotechnics. (Proceedings of Geo-Denver 2000, Denver, Colo., August 5-8, 2000): Reston, Va., American Society of Civil Engineers, Geotechnical Special Publication 99, p. 195–209.Trochim, W.M.K. 2006. Probability sampling In: Research methods. Knowledge base. http://www.socialresearchmethods.net/kb/sampprob.php

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., Vandekerckhove, L. (2006). Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium). Geomorphology 76, 392-410.

Van Den Eeckhaut, M., Moeyersons, J., Nyssen, J., Amanuel Abraha, Poesen, J., Mitiku Haile, Deckers, J., 2008. Spatial patterns of old, deep-seated landslides: a case-study in the northern Ethiopian highlands. Geomorphology 105, 239-252.

Van Den Eeckhaut, M., Poesen, J., Verstraeten, G., Vanacker, V., Moeyersons, J., Nyssen, J., and van Beek, L. P. H. (2005). The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides, Geomorphology, 67(3–4), 351–363.

Van Den Eeckhaut, M., Reichenbach, P., Guzzetti, F., Rossi, M., Poesen, J., 2009. Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium. Natural Hazards and Earth System Sciences 9, 507-521.

Van Den Eeckhaut, M., Marre, A., Poesen, J., 2010. Comparison of two landslide susceptibility assessments in the Champagne-Ardenne region (France). Geomorphology 115, 141-155.

Van Den Eeckhaut, M., Jaedicke, C., Hervás, J., Malet, J.P., Nadim, F., subm. Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. Landslides.

van Westen, C.J., 1994. GIS in landslide hazard zonation: a review, with examples from the Andes of Colombia. In: Price, M.F., Heywood, D.I. (Eds.), Mountain Environments and Geographic Information Systems. Taylor and Francis Publishers, pp. 135–165.

van Westen, C.J., 2004. Geo-information tools for landslide risk assessment: an overview of recent developments. In: Lacerda, W.A., Ehrlich, M., Fontoura, S.A.B., Sayao, A.S.F. (Eds.), Proceedings 9th International Symposium on Landslides, Rio de Janeiro, Brasil, 1. Balkema, pp. 39–56.

van Westen, C.J., Castellanos, E., Kuriakose, S.L. 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology, 102, 112-131.

van Westen, C, Rengers, N, Terlien, M, Soeters, R (1997). Prediction of the occurrence of slope instability phenomena through GIS-based hazard zonation. Geologische Rundschau, 86, 404 – 414.

Vázquez-Selem, L. y Zinck, A.J., 1994. Modeling gully distribution on volcanic terrains in the Huasca area, central Mexico. ITC Journal, 1994-3: 238-251

Wasserman, L. 2007. All of nonparametric statistics. Springer, 268pp.

Weirich F. & Blesius L. (2007) Comparison of satellite and air photo based landslide susceptibility maps Geomorphology 87 (2007) 352 – 364

Wieczorek GF (1987) Effect of rainfall intensity and duration on debris flows in central Santa Cruz Mountains, California. In: (Costa and Wieczorek editors) Debris flows avalanches: processes, recognition and mitigation. Reviews in Engineering Geology, Gelogical Society of America, 7:23-104.

Wieczorek GF, Gori PL, Jager S, Kappel WM, Negussy D (1996) Assessment and management of landslide hazards near Tully Valley landslide, Syracuse, New York, USA. In: Proceedings of the VII International Symposium on Landslides, Trondheim, June 1996, vol 1. Balkema, Rotterdam, The Netherlands, pp 411–416

Wilson J P, Repetto, P L, Snyder R D (2000) Effect of data source, grid resolution, and flow-routing method on computed topographic attributes. In Terrain Analysis – principles and applications, Wilson, J. P. and Gallant, J. C.,Wiley, 133–161.

Wu W, Sidle RC (1995) A Distributed Slope Stability Model for Steep Forested Basins, Water Resources Research, 31, 8, 2097–2110

Yemshanov, D.; Koch, F.; Ben-Haim, Y and Smith, W.D. 2010. Robustness of Risk Maps and Survey Networks to Knowledge Gaps About a New Invasive Pest. Risk Analysis 30(2):261.276.

Yule, G.U., 1900. On the association of attributes in statistics. Philos. Trans. Roy. Soc. London 194A, 257–319.

Zhang W H, Montgomery D R (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations Water Resources Research 30 (4), 1019–1028.

Zinck, J.A., López, J., Metternicht, G.I., Shrestha, D.P. and Vázquez-Selem, L., 2001. Mapping and modelling mass movements and gullies in mountainous areas using remote sensing and GIS techniques. International Journal of Applied Earth Observation and Geoinformation 3(1), 43-53.

6 CASE STUDIES

6.1 CASE-STUDY ON ROBUSTNESS (INFLUENCE OF TERRAIN UNIT) OF SUSCEPTIBILITY MODELS

By: Cristina Baeza, Nieves Lantada and José Moya (UPC)

6.1.1 Abstract[1]

The reliability of susceptibility maps depends largely on the quality of the information used for its evaluation. This study seeks to analyze the influence of sample size and type on the results of discriminant analysis applied to shallow landslide susceptibility assessment. The study also assesses the role of the terrain unit in Discriminant analysis. To this end, two databases based on fieldwork (slope unit) and GIS with 15m and 45m grid cells (grid cell-based unit), were compared in the same zone at La Pobla de Lillet, Spanish Eastern Pyrenees. The results show that although there is no significant influence of the type of sample, it is necessary to use at least half of the individuals of the sample in order to obtain good results from Discriminant analysis. It is the terrain unit that exerts the biggest influence on the result of susceptibility. Some morphometric parameters related to landslides were compared in the databases. The slope unit of the fieldwork database better reflects the land characteristics than the regular grid used by GIS. The values of the variables obtained by GIS procedures are smooth, obtaining mean errors for the slope angle variable of 19.5% and 33.5% for the grids of 15m and 45m, respectively, in the study area. One-way and T tests demonstrate that the smooth of the values exerts a decisive influence on the discriminant results. Kappa's analysis shows that there is no significant equivalence between some of the categorical variables used in both databases. The use of these variables demand the application of clearly defined criteria. The cell size should match the dimensions of the phenomenon analyzed given the unsuitability of the grid of 45m in this study.

6.1.2 Introduction

Nowadays the use of Geographic Information Systems (GIS) technology has allowed the application of sophisticated quantitative techniques to landslide susceptibility assessment. The ability of GIS tools to store, manage, analyze, and display large amounts of spatial and tabular data enhances our capability to analyze large regions. It has proved to be an essential tool in data analysis because it provides automatic procedures for deriving relevant parameters such as slope gradient, slope aspect, slope convexity, upslope contributing area and drainage network order for slope stability. The quality of susceptibility maps depends even more on the quality of the input data (errors in data acquisition, processing and analysis), specifically the landslide inventory. The errors will also depend on the terrain unit used.

A terrain unit is a portion of land surface that contains a set of ground attributes whose values will configure the stability conditions. Different types of terrain mapping units have been proposed (Meijerink 1988). The main terrain units (geomorphologic, grid-cell, unique condition and slope)

[1] This case study is a summary of the paper entitled Influence of sample and Terrain Unit on Landslide Susceptibility Assessment at La Pobla de Lillet, Eastern Pyrenees, Spain. Environ Earth Sci (2010) 60:155–167

have been compared (Carrara et al. 1995) in order to resolve the differences of consistency and physical meaning. The geomorphologic unit has an optimal physical meaning for assessing landslide susceptibility. However, the assessment is not easy to carry out since the boundaries must be defined manually with the result that the map will have lower consistency in front of other terrain units.

Parameters derived from a DEM (digital elevation model) have an uncertainty level that depends on the quality of the data source and the interpolating algorithms (Carla et al. 1997) among others. Some investigators (Van Westen 1993; Wise 1998; Bonin and Rousseaux 2005) have attempted to quantify these errors derived from different algorithms for calculating slope and aspect as the most common geomorphologic parameters used in stability analysis in large regions. Errors, which differ up to ±5° and ±70° for slope and aspect, respectively, have been calculated from DEM (Skidmore 1989). Instability parameters derived from DEM, the terrain unit selected and the approach adopted to assess susceptibility exert a considerable influence on the feasibility of the final results. The type of approach (geomorphologic, empirical and statistical) and the terrain unit are conceptually and operationally interrelated (Carrara et al. 1995). Although type of unit and approach have advantages and disadvantages, their selection depends more on the characteristics of the GIS available, the methodology and the techniques applied to the evaluation of the susceptibility than on scientific and technical requirements of the study.

The foregoing suggests that the uncertainty level in the quality of the final susceptibility maps should be resolved by incorporating the estimated error in each phase of the process. This case study analyses the error in the first phase (data acquisition). It undertakes a comparative analysis between some morphometric parameters related to slope stability obtained manually in the field and derived automatically from DEM (15m cell and 45m cell). An analysis of the distribution of the error of these parameters was performed. The error was considered to be the difference between the automatic value from the DEM and the field value.

The study also assesses the role of the terrain unit (slope unit and pixel-based) in Discriminant analysis. In addition, the influence of the sample type and size used by GIS on susceptibility analysis and on the reliability of the map was analyzed. The sample zone was located at La Pobla de Lillet, Eastern Pyrenees, Spain. This area is composed of sandstones, limestones, marls and flysch formations of Devonian to middle Eocene age and was affected by intense landslide activity during an exceptional rainfall event in November 1982 (Gallart and Clotet 1988; Corominas and Alonso 1990). The spatial distribution of failures (Fig. 6.1.1) in the area is controlled by the lithology, geological structure and by the morphological and hydrological characteristics of the slopes. Most failures were developed on marly and clayey formations (Keuper, Upper Cretaceous, Lower Paleocene and Lower Eocene) and especially on colluvial deposits although some also affected underlying weathered clayey formations. These formations appear mainly in the southern area of the study where failures are concentrated. However, failures are rare in the northern area, which is characterized by bedrock with steep slopes and the absence of surficial deposits. Slope failures are not randomly distributed in the southern area, being common in hollows filled with colluvium, where groundwater flow converges (Baeza 1994, Santacana 2001).

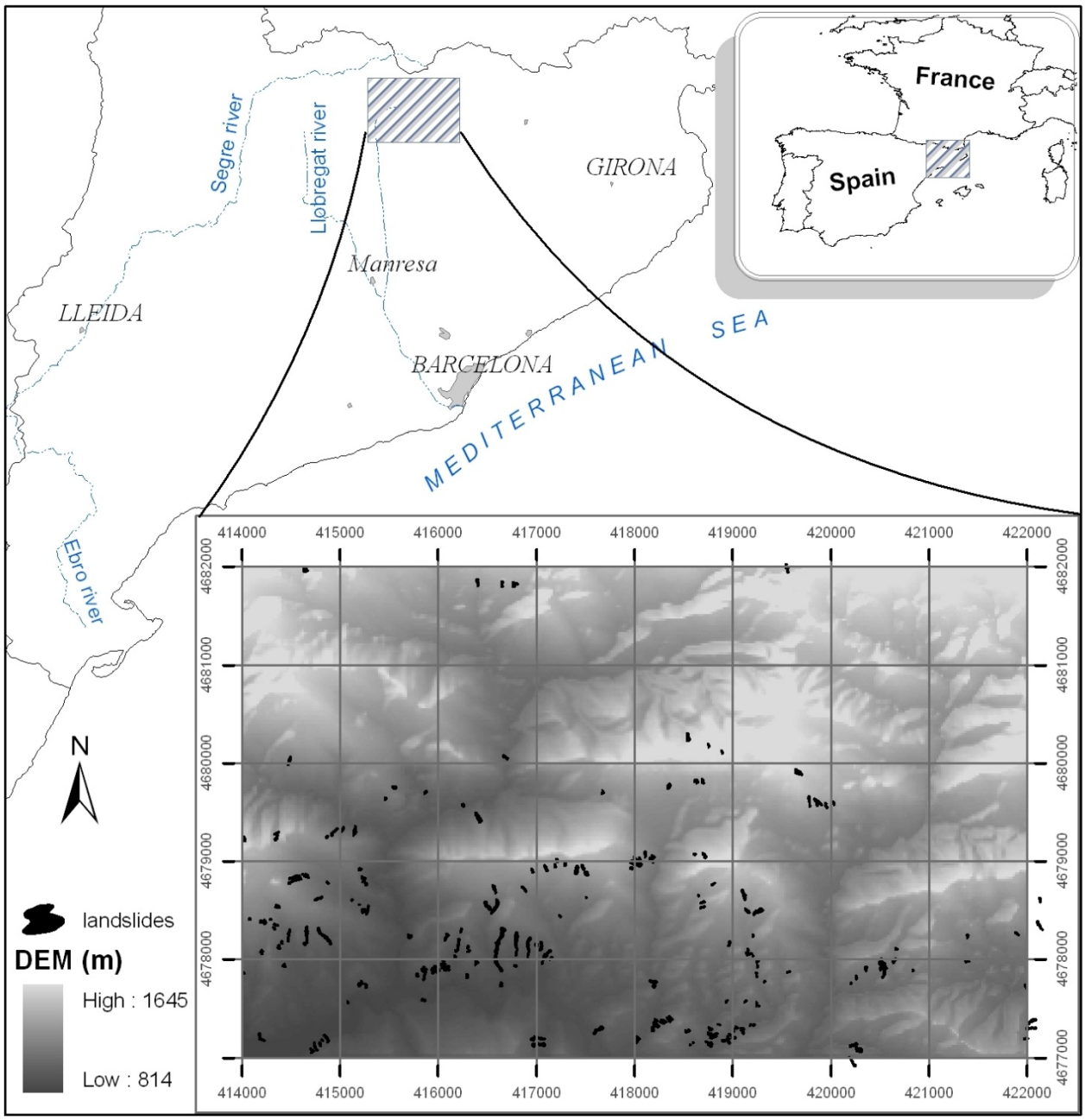


Figure 6.1.1 Geographical location and landslide inventory of the study area

The slope failures considered in this study belong to different shallow landslide types (translational slides and debris flows). These landslides have an average length of 70 meters with small mobilized volumes (less than 10.000m) and failure surfaces that are usually located at a depth up to 2 meters. Shallow landslides were selected because of their simplicity and the knowledge available of the conditioning factors (Baeza and Corominas 2001; Santacana et al. 2003).

6.1.3 Data collection

6.1.3.1 Database used

Two databases of La Pobla de Lillet in the Eastern Pyrenees were used to carry out this study. The first one corresponds to the classic procedure based on variables obtained by interpreting aerial photos and orthophotos and fieldwork. This database forms part of a study to assess the shallow landslide susceptibility over large regions undertaken by Baeza and Corominas (2001). These procedures are costly and time consuming with the result that the extent of the area analyzed is often restricted. The second database corresponds to data simplified in subsequent works (Santacana et al. 2003) that used GIS to automatically capture most of the parameters related to slope failures. Many variables were derived directly from digital elevation models (DEM) or from simple algorithms by GIS (ArcGis). GIS techniques allow us to extract variables from large areas and regionalize the mapping information. The second database was extracted from two DEMs (15m and 45m regular grids) supplied by the Cartographic Institute of Catalonia. Both DEMs were generated by interpolation from a TIN (triangle irregular network) using the topographic information at 1:5000 scale. In both databases landslides were identified by means of aerial photographs at 1:22.000 scale (July 1983) and mapped on the orthophoto at 1:5.000 scale.

The selected parameters were related to the characteristics of the slope and the upslope contributing area. These provide indicators of the geomorphologic evolution of the slope and valuable information for the analysis of stability. The method of selecting these variables, based on a simple failure mechanism, and the significance of these with respect to slope stability are discussed in Baeza and Corominas (1997and 2001) or Hürlimann and Baeza 2002.

The following variables have been analyzed: slope angle, watershed length, watershed angle, upslope contributing area, aspect, transverse curvature of terrain, thickness of surficial formation and land use. The method of obtaining them by GIS and in the field are discussed below.

6.1.3.2 Terrain units

Landslide susceptibility mapping requires a terrain unit that is part of the land surface that contains a group of conditions that differ from the adjacent unit by means of some defined limits (Hansen 1984). A landslide susceptibility level is attributed to the terrain unit according to the stability characteristics defined by geomorphologic and geological parameters. Similarly, a group variable is attributed to each unit, which corresponds to the observed (failed) or to the locations where landslides were not observed (unfailed). The group variable allows us to establish the statistical relationships with the conditioning variables and to make predictions based on the evidence of earlier landslides. The characteristics of these populations (failed/unfailed) could be analyzed by different statistical techniques in order to model landslide susceptibility.

The selection of the terrain unit for the analysis influences the final definition of the model and consequently the reliability of the final result. The difference in treatment of the data in the two databases resides in the terrain unit type (slope unit/grid cell-based). The terrain unit that is considered in the field database is the slope unit. If the clear relationship between landslides, drainage and water divide is borne in mind, it seems reasonable to use these units (Carrara et al. 1995). The analysis based on the grid cell is the one that is the most commonly used by GIS. Two

regular grids (15m and 45m cell) were compared. These results were also compared with those obtained from the field database.

Thus, in the "field-based" approach the variables are arranged in accordance with the terrain units mapped in the field, whereas in the "GIS-based" method all variables are mapped in raster grids in a continuous way. In the latter approach the whole territory is covered. The two types of databases have advantages and disadvantages. The slope unit mapped in the field is defined by units with geomorphologic boundaries that frequently reflect geomorphologic and geological differences. This facilitates the interpretation of the results. However, a problem arises when converting these data to a digital format and storing them in a GIS in order to show landslide susceptibility in a continuous way in a map. The grid cell presents a clear advantage at analytic level. However, their boundaries show no geological and geomorphologic relationships. The disadvantages decrease because of the precision of DEM with increasingly smaller cell dimensions. This implies a larger number of cells to cover an area, and demands more time for processing.

6.1.3.3 Landslide susceptibility assessment method: discriminant analysis

One of the most widely used multivariate techniques for landslide susceptibility assessment is discriminant analysis. Two populations (failed and unfailed cells or slopes) and a group of variables for characterization are necessary. This technique allows us to separate the populations by means of the attribution of objects (slopes or cells) to a group, depending on landslide susceptibility. The procedure is carried out by selecting a group of variables to minimize the erroneous classification (of the previously identified slope or cells as failed or unfailed), and to contribute to the discriminant process. The independent variables selected combine linearly as follows:

$$D = d_1 V_1 + ... + d_p V_p \quad (6.1.1)$$

where $V_{1...p}$ are the independent variables that contribute to instability according to their statistical significance; $d_{1...p}$ are the classification coefficients and D is the discriminant score. The discriminant value is obtained for each slope or cell in a given area. This allows us to obtain a score in terms of their proneness to slope failure, and to define the susceptibility classes. As with other methods of susceptibility analysis, there are a number of assumptions that underlie the discriminant function. The optimum discriminant technique is conditioned upon the independence of the variables and their multivariate normal distribution with the result that the variables analyzed are usually quantitatively continuous. A detailed description of the discriminant analysis can be found in (Dillon and Goldstein 1986).

Discriminant analysis was only applied to the evaluation of the influence of sample size and type on susceptibility but not to the terrain unit.

6.1.4 Data analysis

6.1.4.1 Sample size analysis

When the population is very large, the analysis of the number of elements is often costly and not viable. The solution is to take a sample that contains the information of the population represented (Spiegel 2002; Guzzetti et al. 2006). When carrying out random sampling (Rodríguez 1993), it is necessary to determine the minimum number of elements that defines a sample that ensures a previously fixed standard error in a given population of elements. It is possible to use the following function in order to determine the size of the ideal sample when nominal scales (presence or absence of landslides) and a finite population are used:

$$n = \frac{Nz_\alpha^2 p(1-p)}{e^2(N-1) + z_\alpha^2 p(1-p)} \quad (6.1.2)$$

where n is the size of the sample, N is the size of the population, z_α^2 is the confidence level selected, p is the proportion in which the variable is studied in the population and e is the estimated error. A detailed description of sample size analysis can be found in Prat et al. (1997) and Israel (2008).

Applying this equation, 6.1.1 shows the results for different levels of error and a population (failed and unfailed cells) of 400 and 500. It is assumed that the population follows the Normal (Laplace-Gauss) distribution that the confidence level is 0.05 and that the same probability of belonging to each subpopulation is 0.5. Below an estimated error of 4%, sample sizes are not indicated given that this considerably exceeds half the population. In this case it would be advisable to take the whole population. According to Table 6.1.1, it would be possible to analyze the whole population with one sample of between 200 and 222 cells.

Table 6.1.1 Size of the sample results for different levels of error

Population (failed +unfailed)	Sample (failed + unfailed)		
	Error		
	4%	5%	10%
400	244	200	80
500	278	222	84

Different samples of the grid inventory of 15m were extracted (179,225 unfailed cells and 272 failed cells) to confirm the effect produced by the sample size on the susceptibility results. The susceptibility methodology of shallow landslide described by (Baeza and Corominas 2001) was

employed to determine the function for each sample. The samples were randomly selected automatically by means of a script implemented to this end by GIS. Table 6.1.2 compares the discriminant results (Wilks' Lambda Centroid classification) obtained from different samples with a total population of 50, 100 and 200 cells. Of these original variables, slope angle, slope aspect, thickness of the surficial deposits and elevation were selected in the function, achieving a correct general classification that ranged from 82.0% to 85.6% depending on the sample (Table 6.1.2). Lambda values less than 0.556 and the distance between centroids, in all samples, show that the separation between the two groups by the discriminant functions is successfully achieved. The three samples yield satisfactory results and are similar as regards descriptive values of the function. However, equation 2 demonstrates that the best results are obtained with 200 individuals (out of population of 400 with an error of 5%). Finally, in this case, it may be confirmed that one sample correctly represents the population with at least half of its individuals.

Table 6.1.2 Sample size analysis. Results from discriminant function

Discriminant Function variables: Slope angle + Thickness of surficial deposits + Slope aspect – Dem*					
	Correctly classified (%)				
Sample size: (failed+unfailed)	unfailed cells	failed cells	General	Wilks' Lambda	Group centroids: unfailed / failed
50	75.0	92.0	82.0	0.505	-0.970 / 0.970
100	73.5	93.9	83.7	0.556	-0.885 / 0.885
200	71.7	100.0	85.6	0.501	-0.972 / 1.013

* DEM: Height above mean sea level (m)

6.1.4.2 Sample Type analysis

The random sample must have the same behavior as the global population, taking into account not only the dispersion of the sample but also its variability. Therefore, the influence of the different samples of same size on the results of the discriminant analysis was analyzed. To this end, from a grid of 15m of the study area, 51 random samples with the same probability of individuals (failed and unfailed cells) were extracted, obtaining a discriminant function of each sample ((Baeza and Corominas 2001)).

The statistical parameter of Wilks' Lambda was used to determine the effect of sampling on the final results. Wilks' Lambda is a direct measure of the proportion of variance in the combination of dependent variables that is unaccounted for by the independent variable (grouping variable).Thus, with a range of values of 0 to 1, the smaller the Lambda, the better the statistical significance of the discriminant function. The results (Figure 6.1.2) show that Lambda moves within a very small range of values (from 0.479 to 0.622) with a standard deviation of 0.035. These data indicate the low influence of the sample type on this analysis. Assuming a normal distribution of the values

obtained, the confidence bands of 68% ($\overline{X} \pm SD$) and 95% ($\overline{X} \pm 2SD$) were defined. The mean values of Lambda as well as the mean values of classification of the cells are practically equal (Table 6.1.3). From the foregoing, it follows that the selection of the sample in this study does not significantly influence the final results of the discriminant function. It seems that the random samples correctly reflect the variability of the population and hence the landslide distribution.

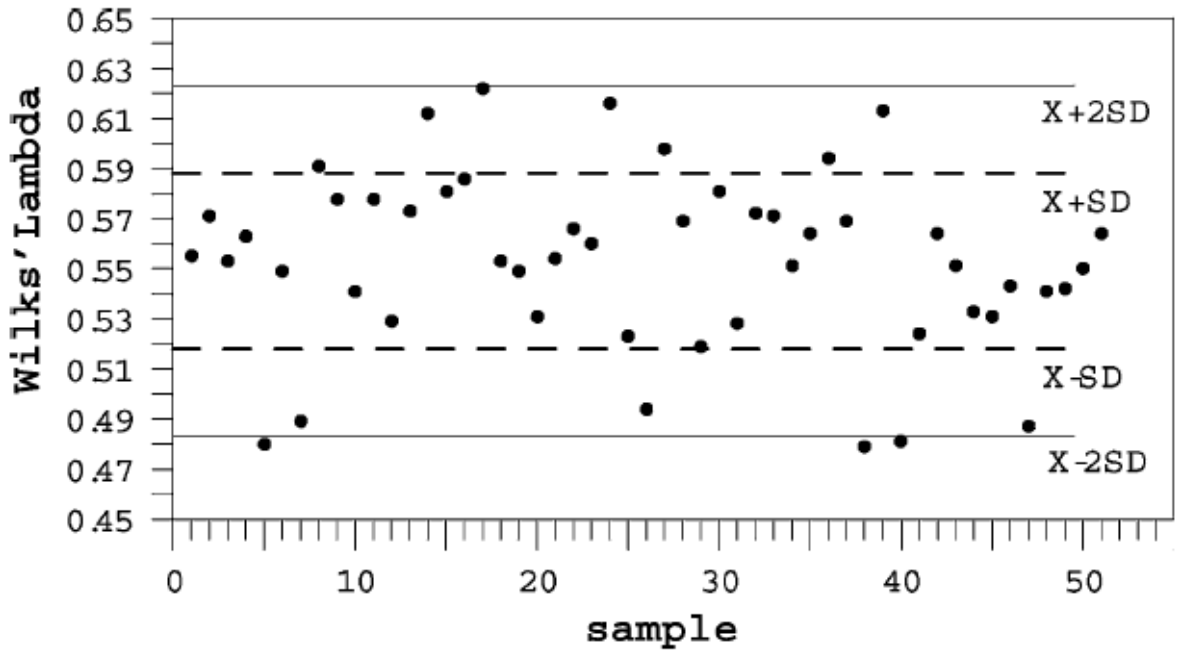


Figure 6.1.2 68% and 95% confidence bands for the Wilks' Lambda values for the samples.

Table 6.1.3 Sample type analysis. Results from discriminant analysis for different confidence levels

Confidence bands	Lambda Mean Value	Correctly classified % (mean value)		
		Unfailed slopes	Failed slopes	General
$\overline{X} \pm SD$ (68%)	0.554	71.81	92.89	82.35
$\overline{X} \pm 2SD$ (95%)	0.557	71.36	92.86	82.11

6.1.4.3 Terrain Unit Analysis

A comparative analysis between terrain factors derived from DEM with 15m and 45m grid cell (pixel-based) and field work (slope unit) used in landslide susceptibility assessment at La Pobla de Lillet was undertaken. An attempt was made to determine the influence of the terrain unit on the discriminant technique. The evaluation was made by contrasting the real values obtained from orthophotos and field checking and the values derived automatically from DEM by ArcGis software, native ESRI ((ESRI 2008)). For the sake of comparison, units from the two databases (slope units and pixel-based) were taken in accordance with UTM coordinates. The field values were regarded as a source of high quality data which enabled us to provide an orientation of the error of the DTMs (Digital Terrain Model). Parametric tests (T-test and One-way) were also applied on continuous variables, in addition to the error analysis, in order to assess the influence of terrain unit on susceptibility. These tests (Chan and Walmsley 1997; Bewick et al. 2004) are used in a preliminary application of discriminant analysis that maximizes the differences between means and variances of failed and unfailed groups. The T-test ascertains whether the means of the two groups differ statistically from each other by calculating Student's t and testing the significance of the differences between the means. One-way analysis of variances uses Fisher's F- distribution as a part of the tests of statistical significance. These tests provide information about the terrain unit that best discriminates between failed and unfailed groups and that gives the most reliable explanation of susceptibility.

6.1.4.4 Continuous variables

The values of the following continuous variables were contrasted: slope angle, mean watershed angle, watershed length and upslope contributing area.

Slope angle

GIS database (pixel-based unit): slope angle is derived directly from DEM and identifies the maximum change in elevation over the distance between each cell and its eight neighbors. Field database (slope unit): values were measured by inclinometer around the landslide scar prior to failure. The range of its values in degrees is 0 to 90.

The figures below (Figure 6.1.3) show the error distribution defined as the difference between field values and the values obtained for each grid by GIS. Table 6.1.4 gives the descriptive values of the sample analyzed for each terrain unit. The values of the slope angle obtained directly in the field are higher than those obtained automatically from the regular grids of 15m and 45m, respectively. From the table it follows that the mean error of dtm15 and dtm45 with respect to the field values is 6.56^{o} and 11.32^{o} respectively, which is equivalent to an error of 19.5% and 33.5%. For small values, the slope is very similar for all supports whereas for high slopes, the values differ considerably in accordance with the support.

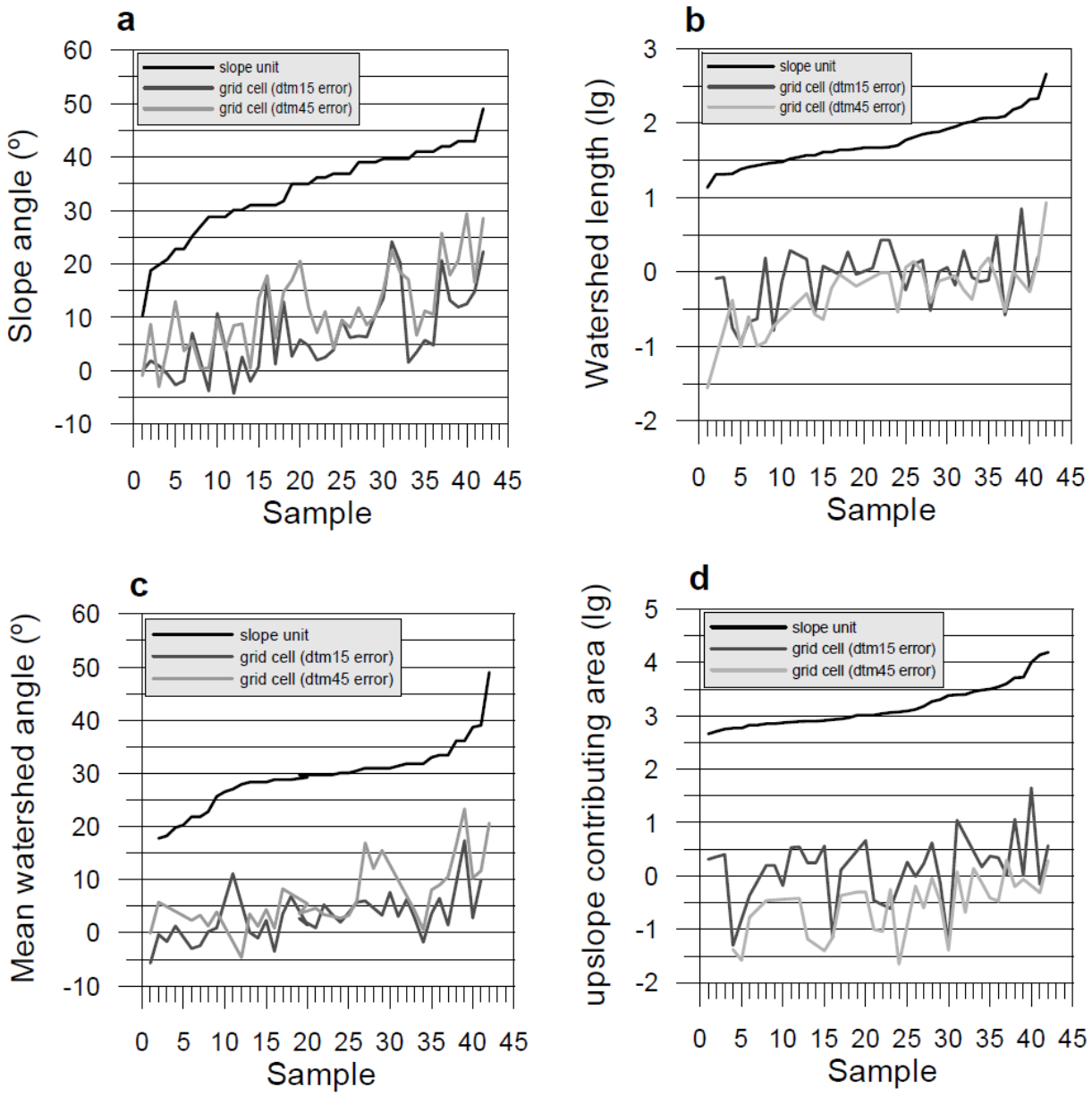


Figure 6.1.3 Error distribution of the digital terrain models of 15m and 45m grid cell for slope angle (a), watershed length (b), mean watershed angle (c) and upslope contributing area (d)

Figure 6.1.3a shows the three curves: the field value, the error of the slope of the grid of 15m and that of 45m. The correlation between slope angle error for the two grids and slope angle for slope unit (field) was quantified using a regression function (Table 6.1.5). The model fit of the regressions of the two grids are very similar but not identical. Thus, Figure 6.1.3a shows that the error of the slope of the grid of 45m exceeds that of 15m. This is confirmed by the standardized coefficients (Sβ) of the regression function (Table 6.1.5), which expresses the rate of change in the slope angle

error from dtm, when the field slope angle varies one unit, i.e. the error of dtm 15 and 45 increases by 0.66 and 0.71, respectively, when the field slope angle increases by one degree.

Table 6.1.4 Descriptive values of the variables for different terrain units

Terrain unit	Minimum (°)	Maximum (°)	Mean (°)	Standard deviation
Slope angle				
slope unit	10.20	48.99	33.76	8.08
grid cell (dtm15)	10.22	38.15	27.20	6.86
grid cell (dtm45)	9.82	34.38	22.44	6.32
Watershed length				
slope unit	13.60	460.00	77.00	78.28
grid cell (dtm15)	16.92	471.61	102.93	100.33
grid cell (dtm45)	47.11	498.41	152.82	114.84
Mean watershed angle				
slope unit	17.74	48.99	29.49	5.83
grid cell (dtm15)	15.92	35.82	25.74	4.58
grid cell (dtm45)	11.42	32.58	23.03	5.50

Table 6.1.5 Regression function of slope angle error, watershed angle and length error of the two grids (dtm15 and dtm45)

	slope angle error (grid cell -dtm15)				
	β	SE	Sβ	T (α=0.05)	Sig T
constant	-14.542	3.896		-3.732	0.001
slope angle (slope unit)	0.599	0.111	0.663	5.391	0.000
	R = 0.66		R^2 = 0.44		
	slope angle error (grid cell-dtm45)				
	β	SE	Sβ	T (α=0.05)	Sig T
constant	-11.384	3.772		-3.018	0.004
slope angle (slope unit)	0.679	0.109	0.71	6.219	0.000
	R = 0.71		R^2 = 0.50		
	watershed angle error (grid cell-dtm15)				

	β	SE	Sβ	T (α=0.05)	Sig T
constant	-9.471	2.106		-4.493	0.000
watershed angle (slope unit)	0.417	0.073	0.717	5.730	0.000
	R = 0.72		R^2 = 0.51		
	watershed angle error (grid cell-dtm45)				
	β	SE	Sβ	T (α=0.05)	Sig T
constant	-9.528	3.130		-3.044	0.005
watershed angle (slope unit)	0.534	0.103	0.707	5.192	0.000
	R = 0.71		R^2 = 0.48		
	watershed length error (grid cell-dtm45)				
	β	SE	Sβ	T (α=0.05)	Sig T
constant	-2.139	0.299		-7.160	0.000
watershed length (slope unit)	1.033	0.164	0.754	6.293	0.000
	R = 0.75		R^2 = 0.57		

Moreover, this is corroborated by the results of a univariate regression analysis of the slopes between the grids of 15m and 45m (Table 6.1.6). The standardized coefficient of the function (Sβ) shows that the rate of change in the value of the slope angle of the grid of 45m is 0.88 when the grid of 15m varies one unit. Although the statistical T (Student's T test) at a significance level of 95% shows a correlation between the error of the grids and the values observed in the field (reference value), the variance (R^2) of the regressions does not exceed 50%. These results suggest that some parameters were not taken into account in the function and should be considered when explaining the error.

The results displayed by *t*-test (Table 6.1.7) show that slope angle is statistically significant in all terrain units (*t*-test sig.<0.05). However, according to the F value, the variable that best characterizes the stability of the slope is represented by the slope unit (F=55.162) followed by the grid cell dtm15 and the grid celldtm45, which is the lowest value. According to these results, the variable would be included in the discriminant function from any terrain unit although the susceptibility would be best represented by data from the slope unit (field database).

Overall, the slope angle values are smoothed compared with the field values, and the sensitivity of the values estimated with GIS decreases as the grid cell size increases. The smoothing of the values derived from DEM reduces the separation between populations, which diminishes the significance of the variable in the discriminant function.

Table 6.1.6 Regression function of slope angle and watershed angle between the two grids (dtm15 and dtm45)

	Slope angle (grid cell-dtm45)				
	β	SE	Sβ	T (α=0.05)	Sig T
constant	3.604	1.929		1.868	0.071
slope angle (dtm15)	0.758	0.071	0.881	10.705	0.000
	R = 0.88		R^2 = 0.78		
	Mean watershed angle (grid cell-dtm45)				
	β	SE	Sβ	T (α=0.05)	Sig T
constant	-3.679	4.047		-0.909	0.372
mean watershed angle (dtm15)	0.992	0.150	0.792	6.619	0.000
	R = 0.79		R^2 = 0.63		

Table 6.1.7 One-way and T-tests to compare unfailed and failed slope groups

Variable	Slope unit (field)				Grid cell (dtm15)				Grid cell (dtm45)			
	One-way test		T-test		One-way test		T-test		One-way test		T-test	
	F	Sig. *	t	Sig. *	F	Sig. *	t	Sig. *	F	Sig. *	t	Sig. *
Slope angle	55.162	0.000	7.427	0.000	27.800	0.000	4.431	0.000	4.730	0.030	2.176	0.030
Watershed length	17.092	0.000	5.179	0.000	6.000	0.015	-0.587	0.558	1.700	0.190	-0.790	0.430
Watershed angle	5.535	0.020	2.353	0.020	5.279	0.022	2.298	0.022	2.400	0.119	1.579	0.116
Upslope contributing area	31.104	0.000	5.577	0.000	6.000	0.015	0.819	0.413	1.800	0.176	1.375	0.170

* Significance level α=0.05

Watershed length

GIS database (pixel-based unit): this defines the terrain distance from each cell to the watershed line in accordance with the slope. Therefore, the variable is derived from DEM, dividing the flowlength value of each cell by the cosine of the slope in the cell.

Field database (slope unit): this variable was interpreted from orthophotos 1:5.000 and then obtained in the field by homogeneous sections from the scar to the watershed line where slope angle and length were measured.

Table 6.1.5 shows an increase from the slope unit to dtm15 and dtm45 in the minimum, maximum and mean of the watershed length values. The mean error of the dtm15 is 26m and that of the dtm 45 is 76m. It should be pointed out that there is a considerable standard deviation, which limits the results. Thus, the error of the grid of 45m shows a clearer trend (Fig. 6.1.3b). Table 6.1.6 gives the regression function obtained from the grid of 45m. The statistical Student's *t* indicates for α=0.05 that the function is statistically significant, establishing a correlation coefficient of R_{dtm45}= 0.75, which accounts for the variance (R^2) of 57%. The standardized coefficient (Sβ) shows an increase of 0.75 for the dtm45 error for each slope unit increase.

Parametric tests (Table 6.1.7) show that watershed length does not have the same significance for each terrain unit. The results demonstrate that the variable is statistically significant for both tests (mean and variance) exclusively for the slope unit. Grid cells (dtm15 and 45) with *t*-test sig.=0.558 and *t*-test sig.=0.430, respectively, do not show differences in the mean of the failed and unfailed groups when discriminating. This means that the variable from the grid cells would not be selected by discriminant analysis to define the susceptibility function. However, the variable from the slope unit would be included in the function.

Mean watershed angle

GIS database (pixel-based unit): this is obtained with the arc cosine function of the horizontal distance of the downslope flowpath to each cell (obtained from DEM with flowdirection and flowlength functions) divided by the flowpath.

Field database (slope unit): this was interpreted from orthophotos and then measured in the field.

The values of the mean slope (Table 6.1.4) of the watershed obtained using dtm 15m and dtm 45m are mostly lower than those obtained from the slope unit. The mean values correspond to an error of 12.7% and 21.9%, respectively. Although the trend is not as marked as the slope angle, the same correlations between the error from the DTMs and the slope unit are observed. The correlation coefficients of the regression functions (Fig. 6.1.3c) with values of R_{dtm15}= 0.72 and R_{dtm45}= 0.71 only account for 51% and 48% of the variance for the grids of 15m and 45m, respectively (Table 6.1.5). The standard coefficient (Sβ) of each function suggests a similar rate of change of 0.7 units for the dependent variable. Despite the low variance, the T parameter for a level of significance of 95% reflects a statistical significance between the error of measure of the grid cell and slope unit based values. Although the rate of change (Sβ) is similar for both grids, it may be observed that the error is generally greater for the values obtained from dtm45. This suggests that the values of the watershed angle are smoothed with respect to those of the slope unit and those obtained from dtm 15. A regression function between the values of each grid was obtained (Table 6.1.6). The statistics

show that the rate of change in the value of the watershed slope angle of dtm 45 is 0.79 for a variance of 63%.

The results displayed by T-test and One-way (Table 6.1.7) show that the slope unit and grid cell (dtm15) are very similar as far as differences of means and variances are concerned. Watershed angle obtained from the grid cell (dtm45) is not statistically significant (*t*-test sig.= 0.116; F test sig.=0.119). The variable from the slope unit and the grid cell (dtm15) could be selected by discriminant analysis for the function, though this is not the case of the variable from dtm45.

Upslope contributing area

GIS database (pixel-based unit): this is the area of the basin upslope of a specific cell. This variable is derived from DEM, and from two direct functions, Flow direction and Flow accumulation, and is the number of cells that flow into each cell. This variable is multiplied by the area of the cell.

Field database (slope unit): this variable was interpreted from aerial photographs at 1:22.000 and orthophotos at 1:5.000 scale.

Figure 6.1.3d shows the error of the value of the watershed of the dtm of 15m and 45m. The figure shows that the mean value of the watershed (3.04) for dtm15 is lower than the mean value of slope unit (3.16). The value obtained for dtm 45 (3.78) is much greater. Moreover, a trend in the error is observed in the values of dtm45. Despite this trend, there is no statistically significant correlation between the error of the values based on the grid cell and their corresponding slope unit values, which enables us to define a regression function. However, the parametric tests provide more information (Table 6.1.7). As usual, the variable from the slope unit stands out for its statistical significance with high values of F (31.104) and T (5.577). The populations (failed/unfailed) show no difference in the means (*t*-test sig>0.05) for the variable from dtm15 or for the one from dtm45. Given that the discriminant analysis works by maximizing the differences between means and variances of failed and unfailed groups, the variable could be included in the discriminant function for data from the slope unit.

6.1.4.5 Categorical variables

A number of tests have been devised to evaluate the quality of the interval, ordinal and nominal data. Both parametric and nonparametric statistical tests can be employed to compare true values (those observed on the ground) and those recorded in data set automatically by GIS. Cohen's kappa coefficient (Cohen 1960) is a statistical measure of inter-rater agreement or concordance. It yields a score of homogeneity when observing or coding qualitative/categorical variables. This coefficient is useful for ascertaining whether a particular scale is appropriate for measuring a given variable. It is considered to be a more robust measure than a simple percent agreement calculation. Kappa has a range from 0-1.00 with larger values indicating better concordance. The Kappa statistic is derived from the following contingency table:

$$Kappa = \frac{(d-q)}{(N-q)} \quad (6.1.3)$$

where *d* is the diagonal total of cells from the contingency table; *N* is the total of columns or rows which should be equal and *q = n(rows)*n(columns)/N.* Thereafter, the Kappa statistic (3) is used to analyze the concordance of the variables obtained automatically from GIS (grid cell unit) and the one obtained by field checking (slope unit). This enables us to obtain a reliability value of the former variables.

The following categorical variables were analyzed by the Kappa coefficient: aspect, transverse curvature, thickness of surficial deposits and land use.

Aspect

GIS database (grid cell unit): aspect is the steepest downslope direction from each cell to its neighbors. It is calculated as the compass direction of the horizontal projection of the normal vector to the surface in a point. Its values range from 1 to 360, and are -1 in flat areas with no downslope direction.

Field database (slope unit): the aspect of the scar of the landslide was measured by a compass in the field. In both databases the variable was transformed from 0-360 to 0-180 (North- South) in order to avoid using the circular statistic and better define the N and S slopes that predominate in the study area.

The results obtained from the aspect values taken in the field and those extracted automatically from the grids of 15m and 45m do not show a big difference. The mean values, despite typical deviations, are very similar: slope unit = 107º (sd=63.8), dtm 15 =115º (sd=59.5) and dtm 45= 116º (sd=56.7). This indicates slopes that are mainly oriented to the S-SE in the study area. The Kappa statistic is used to categorize these variables into four intervals so that it reflects a sunny or shady orientation (N, NE-NW, SE-SW, S). Table 6.1.8 shows the Kappa statistic of the variables in addition to the concordance value between the results of the two grids. The Kappa indexes show that there is a good concordance between the field classification and that obtained automatically from the two grids. As expected, the Kappa statistic obtained after contrasting the two grids is larger than that obtained from the contrast of these and the field classification. However, this statistic does not exceed 0.7, which is considered satisfactory.

Transverse curvature

GIS database (grid cell unit): this is one of the two optional variants of curvature, which influences convergence and divergence of flow. It is generated by adjusting a fourth polynomial function for the immediate 3 x 3 neighbourhood cells. Its variable takes values from -20 to 20 units. Values close to 0 indicate flat areas, and negative and positive values are concave and convex areas, respectively.

Field database (slope unit): transverse curvature was defined by field work in the scar area as a categorical variable (0: rectilinear; 1: convex; -1: concave).

Out of 42 samples, the values obtained for transverse Curvature for the grid of 15m range from -4.11 to 3.02 with a mean value of -0.27. The values for the grid of 45m range from -1.21 to 1.29

with -0.12 as the mean value. The widest range of values of the grid of 15m reflects the relief more satisfactorily. This model approximates more to reality than the grid of 45m, which is smoothed. However, the mean values show that the two grids have slightly concave reliefs near to the flat surfaces with values that are negative and close to 0.

Only 36.5% of the positive values correspond to the convex reliefs in the field classification. There is 55% of correspondence in the classification of concave reliefs (negative values). The Kappa indexes were obtained after redefining the variables in three categories according to the histograms of the value obtained for each grid: convex (values > 0.3) rectilinear (values of 0.3 to - 0.3) and concave (values < -0.3) and contrasting them with the field classification (Table 6.1.8). All the results confirm the very low correspondence between the values identified in the field and those defined by the grids. The correspondence between the two grids is also low due to the fact that the Transverse Curvature variable acquires different value ranges depending on the zone and the DEM used.

Table 6.1.8 Kappa index values from the contingency tables

	Kappa value	T (0.05%)	Sig T
Slope aspect			
grid cell (dtm15) – slope unit	0.578	6.231	0.000
grid cell (dtm45) – slope unit	0.580	6.190	0.000
grid cell (dtm15) – grid cell (dtm45)	0.640	6.863	0.000
Transverse curvature			
grid cell (dtm15) – slope unit	-0.091	-0.834	0.404
grid cell (dtm45) – slope unit	0.086	0.833	0.405
grid cell (dtm15) – grid cell (dtm45)	0.214	2.046	0.041
Thickness of surficial dep.			
grid cell (dtm15) – slope unit	0.553	3.645	0.000
grid cell (dtm45) – slope unit	0.245	1.818	0.069
grid cell (dtm15) – grid cell (dtm45)	0.532	4.261	0.000
Land use			
grid cell (dtm15) – slope unit	0.111	1.371	0.170
grid cell (dtm45) – slope unit	0.026	0.372	0.710
grid cell (dtm15) – grid cell (dtm45)	0.758	8.426	0.000

The thickness of surficial formation

GIS database (grid cell unit): a zoning of the thickness was carried out by photo aerial interpretation at 1: 22.000 scale, orthophotos at 1:15.000 and measures obtained in the field. Thereafter, digitalization and construction of a vector cover of polygons in Arc/Info were undertaken. This

variable is grouped into three categories weighted according to the following values:1(without surficial formation), 2 (from 0 to 0.5m) and 3 (>0.5m of surficial formation).

Field database (slope unit): this variable is continuous with measures performed at road and cut slopes where the whole surficial formation is observed.

The Kappa indexes for thickness of surficial formation were determined by a contingency table (Table 6.1.8) between the categorized variable from GIS and the continuous field values. These reflect a reasonable concordance between the field classification and that of DTM 15 with 0.553 but not for DTM 45 with a value of 0.245.The concordance value between the two grids continues to be relatively satisfactory because of the limitations of cell size.

Land use

GIS database (grid cell unit): this is a qualitative variable obtained by aerial photographs at 1:22.000 scale and orthophotos 1:10.000. This information is digitized and topology is constructed. Different categories of land use were defined (bedrock, grassland, shrubs, clear forest, dense forest and scree deposits).

Field database (slope unit): the variable was defined in same way as the one from GIS. Measures were taken around the scar to characterize the type and density of vegetation of the area prior to failure.

A Kappa index value for of 0.111 for DTM 15 and 0.026 for DTM 45 were obtained from the contingency table (Table 6.1.8). The two values of the index show no concordance with field classification (real correspondence of approximately 35%). However, the concordance between the grids is very high (0.758) given that the dimensions of the polygons correspond to the different categories. The possible error of these polygons does not exert much influence because of the size of these polygons (much larger than the cell).

6.1.5 Concluding Remarks

Two databases based on the slope unit and grid cells, at La Pobla de Lillet (Spanish Eastern Pyrenees), were contrasted to determine the influence of the sample and terrain units on landslide susceptibility. In the light of our findings, a sample correctly represents the population with 50% of its individuals, yielding good results of landslide susceptibility with discriminant analysis. The sample type exerts no significant influence on the final results of the discriminant function in susceptibility evaluation. It seems that the random samples correctly reflect the variability of the population as well as the landslide distribution.

In order to analyze the influence of the terrain units on the susceptibility model by discriminant techniques, two different analyses were performed on some variables. First, an attempt was made to evaluate the error of the DTMs by comparing the values of the variables of different terrain units: grid cells of 15 m and 45m (automatic from GIS) and the slope unit (field database). Second, T and One-Way tests were applied to evaluate the differences between means and variances of failed and unfailed groups used for discriminant function. The results show that the mean error for the variables based on 15m grid cell ranges from 19.5% for the slope angle to 33% for the watershed

length. The values estimated with GIS decrease as grid cell size increases. Thus, the mean error for 45m grid cell reaches 33.5% for the slope angle and 98% for the watershed length. The correlation between variables based on the two grids and the slope unit was quantified using regression functions. This correlation coefficient (R) of the error ranges from 0.66 to 0.75, the error between 15m and 45m grid cells being smaller. Although the correlations are significant at a level of 95%, the variance that accounts for the error (R^2) of the regressions does not exceed 50%.

Given that the discriminant model uses mean values of the variables to differ between populations (unfailed and failed), the greater the smoothing, the smaller the separation between the populations and the poorer the discriminant model. T and One-way tests were performed to account for this effect and provide information about the terrain unit that best discriminates between the groups. The results show that the slope angle is the only variable that is statistically significant in all the terrain units analyzed. However, it is the slope unit that achieves the best separation between failed and unfailed groups. In the variables derived from DEM, which require slope models, the possible error of the models is accumulated. The accumulated errors do not allow a clear separation. Thus, the results show that watershed length, watershed angle and upslope contributing area do not have the same significance for the terrain units. According to the T and F statistics, the separation between failed and unfailed groups of the watershed variables based on grid cells is smaller than that based on the slope units. Using these statistics, the values of the 45m grid cell are smaller than those of the 15m cell. Considerable discrepancies in the variables derived from the watersheds result from a decreased resolution. Hence, watershed boundaries derived from the DEM of 45m have proved to be unreliable. The reason for this could be that the slope and the aspect errors give rise to smooth areas resulting in errors of flow direction and flow accumulation, which are used to obtain the watershed variable. Consequently, the smoothing of the values could result in a minor separation of the populations.

These tests demonstrate that the smoothing of the variables based on grid cells with respect to the slope unit values could exert a decisive influence on the discriminant result. The results indicate that the discriminant model would be better employing the slope unit variables than using those derived from DEM. When employing the grid, the model derived from the grid of 45m would be less reliable than a grid of 15m with the result that the susceptibility would not be suitably defined.

Overall, the results show that the type and size of the terrain unit used exert a significant influence on the results of the susceptibility model. The cell size influences the values of the variables, smoothing them in the case of the grid of 45m. As the cell size is increased, information is lost and modeling of the relief is impaired. The cell size should match the dimensions of the phenomenon to be analyzed. In this study, the grid of 45m is unsuitable given that the cell is too big with respect to the landslide dimensions under study. The use of the grid cell as a terrain unit, especially from dtm45, in discriminant analysis does not seem to be very effective in this area. Thus, another terrain unit using GIS, which better characterizes the slope, should be used in future studies.

According to the Kappa analysis, there is in general no equivalence between the slope unit variables and those obtained from grid cell (dtm15 and dtm45). Only in the case of aspect is there a good equivalence, yielding similar results for any grid size. The variable obtained from the grid cell unit could be replaced by the slope unit variable in the discriminant function without losing much reliability. The Transverse Curvature variable has a relative correspondence with the slope unit variable. This variable can be used in the model. However, it should be noted that the results vary considerably according to DTM. In any case, the physical meaning of the DTM variable is not the

same as that of the qualitative slope unit variable. The error derived from the process of rasterizing of the vector information should be added to the possible errors of interpretation and regionalization of qualitative variables (thickness of surficial formation, land use). The automatic conversion is undertaken so that the final cell contains the value of the polygon with the largest cell surface. The polygon size of the categories of the variable and the cell size exert a decisive influence on the final value. The criteria of categorization of these variables should be clearly defined prior to analysis in order to limit the subjectivity of the qualitative variables in the discriminant function

References

Aleotti, P. and Chowdhury, R. (1999) Landslide hazard assessment: summary review and new perspectives. B Eng Geol Environ 58:21-44

Ayalew, L. and Yamagishi, H. (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology 65:15-31

Baeza, C. (1994) Evaluación de las condiciones de rotura y de la movilidad de los deslizamientos superficiales mediante el uso de técnicas de análisis multivariante. Ph.D.thesis. Departament of Geotechnical Engineering and Geosciences, Technical University of Catalonia. Barcelona, Spain

Baeza, C. and Corominas, J. (1997) Susceptibility analysis of shallow landsliding by multivariate techniques. In:TheThird annual conference of the international association for mathematical geology, Barcelona 2, 928-936

Baeza, C. and Corominas, J. (2001) Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. Earth Surf Proc Land 26:1251-1263

Bewick, V., Cheek, L. and Ball, J. (2004) Statistics review 9: One-way analysis of variance. Critical care 8 (2):130-136

Bonin, O. and Rousseaux, F. (2005) Digital Terrain Model Computation from contour lines: how to derive quality information from artifact analysis. Geoinformatica 9 (3):253-268

Carla, R., Carrara, A. and Bitelli, G. (1997) Comparison of techniques for generating digital terrain models from countour lines. Int J Geogr Inf Sci 11 (5):451-473

Carrara, A., Cardinali, M., Guzzetti, F. and Reichenbach, P. (1995) GIS technology in mapping landslide hazard.In: A. Carrara and F. Guzzetti (ed) Geographical Information Systems in Assessing Natural Hazards. 135-175

Chan, Y. and Walmsley, R. P. (1997) Learning and Understanding the Kruskal-Wallis One-Way Analysis-of-Variance-by-Ranks Test for Differences Among Three or More Independent Groups. Physical therapy 77 (12):1755-1761

Chung, C. and Fabbri, A. (2003) Validation of spatial prediction models for landslide hazard mapping. Nat Hazards 30:451-472

Chung, C., Kojima, H. and Fabbri, A. (2002) Stability analysis of prediction models for landslide hazard mapping.In: R. J. e. Allison (ed) Applied Geomorphology: Theory and Practice. 3-19

Clerici, A., Guzzetti, F., Tellini, C. and Vescovi, P. (2002) A procedure for landslide susceptibility zonation by the conditional analysis method. Geomorphology 48 (4):349-364

Cohen, J. (1960) A coefficient of agreement for nominal scales. Educ Psycho Meas 20 (1):37-46

Corominas, J. and Alonso, E. (1990) Geomorphological effects of extrem floods (November 1982) in the Southern Pyrenees. In:2nd symposium of hidrology in Mountainous Regions Artificial Reservoirs: water and slopes, Lausanne 194, 295-302

Dai, F. C. and Lee, C. F. (2002) Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42:213-228

Dai, F. C. and Lee, C. F. (2003) A spatiotemporal probabilistic modelling of storm-induced shallow landsliding using aerial photographs and logistic regression. Earth Surf Proc Land 28:527-545

Dillon, W. R. and Goldstein, M. (1986) Multivariate analysis. Methods and applications. Wiley & Sons (eds), p 587

Ermini, L., Catani, F. and Casagli, N. (2005) Artificial neural networks applied to landslide susceptibility assessment. Geomorphology 66:327-343

ESRI (2008) ESRI: Gis and Mapping Software.

Gallart, F. and Clotet, N. (1988) Some aspects of the geomorphic processes triggered by an extrem rainfall event: The November 1982 flood in The Eatern Pyrenees. Catena 13:79-85

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M. and Galli, M. (2006) Estimating the quality of landslide susceptibility models. Geomorphology 81:166-184

Hansen, A. (1984) Landslide Hazard Analysis.In: D. Brunsden and D. B. Prior (ed) Slope Instability. 523-602

Hürlimann, M. and Baeza, C. (2002) Analysis of debris-flow events in the Eastern Pyrenees, Spain. In:First European Conference on Landslides, Praga 213-220

Israel, G. D. (2008) Determining Sample Size.University of Florida IFAS Extension. Publication #PEOD6

Lee, S., Ryu, J.-H., Won, J.-S. and Park, H.-J. (2004) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Eng Geol 71:289-302

Meijerink, A. M. J. (1988) Data acquisition and data capture through terrain mapping units. ITC Journal 7:23-44

Muñoz, J. A., Martínez, A. and Vergés, J. (1986) Thrust sequences in the Spanish Eastern Pyrenees. J Struct Geol 8 (3/4):399-405

Ohlmacher, G. and Davis, J. (2003) Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng Geol 69:331-343

Prat, A., Tort-Martorell, X., Grima, P. and Pozueta, L. (1997) Métodos estadísticos. Control y mejora de la calidad. Barcelona, p 300

Puigdefábregas, C., Serrat, D. and Vilaplana, J. M. (1979) El medi geològic.In: R. e. Folch (ed) Patrimoni natural d'Andorra. 15-39

Remondo, J., González-Díez, A., Díaz de Terán, J. R. and Cendrero, A. (2003a) Landslide susceptibility models using spatial data analysis techniques. A case study from the Lower Deba Valley, Guipúzcoa (Spain). Nat Hazards 30:267-279

Rodríguez, J. (1993) Métodos de muestreo. Casos prácticos. Madrid, p 181

Santacana, N. (2001) Análisis de la susceptibilidad del terreno a la formación de deslizamientos superficiales y grandes deslizamientos mediante el uso de sistemas de información geográfica. Aplicación a la cuenca alta del río Llobregat. Ph.D.thesis. Departament of Geotechnical Engineering and Geosciences, Technical University of Catalonia. Barecelona, Spain

Santacana, N., Baeza, C., Corominas, J., De Paz, A. and Marturià, J. (2003) A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet area (Eastern Pyrenees, Spain). Nat Hazards 30:281-295

Skidmore, A. K. (1989) A comparison of techniques for calculating gradient and aspect from gridded elevation data. Int J Geogr Inf Syst 3:323-334

Spiegel, M. R. (2002) Estadística. McGraw-Hill, Madrid, p

Wise, S. M. (1998) The effect of GIS interpolation errors on the use of digital elevation models in geomorphology, in Lnadforms monitoring, modelling and analysis. Wiley, p

Yesilnacar, E. and Topal, T. (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Eng Geol 79 (3-4):251-266

6.2 CASE-STUDY ON THE COMPARISON OF THE PERFORMANCE AND VALIDATION OF DIFFERENT SUSCEPTIBILITY MODELS. APPLICATION TO THE BARCELONETTE REGION SOUTH EAST FRANCE.

By: Jean-Philippe Malet (CNRS)

Probabilistic assessment of landslide susceptibility has become a major topic of research in the last decade. Most progress has been accomplished on producing susceptibility maps at large scales (> 1:25,000) and on evaluating their performance at both a spatial and a temporal scale. At the 1:10,000 scale, which is the scale of production of most of the regulatory risk maps in Europe, few tests on the performance of these methods have been performed, and few information on the accuracy of the produced maps is available.

6.2.1 Objective of the assessment

The objective of this work is twofold:

1) First a workflow is presented in order to identify the best set of variables (e.g. conditioning factors) using quantitative measures to produce an accurate susceptibility assessment. The

workflow discusses the best way to minimize Conditional Independence (CI) between the predictive variables which can severely bias the simulated maps by over or under-estimating the probabilities, and discusses the necessity to potentially combine series of variables or series of classes in neo-predictive variables or classes to increase the predictability power of the model.

The workflow is presented using the bivariate Weight-of-Evidence model but can be applied to any other types of bivariate or multivariate statistical models (Thiery et al., 2007).

2) Second two different types of statistical models (Weight-of-Evidence, Fuzzy Logic) are applied to produce a susceptibility assessment and are compared to an expert map produced according to the methodology used for the implementation of the Risk Prevention Plans (*Plan de Prévention des Risques 'Mouvements de Terrain'*, PPR; MATE/MATL, 1999) in France. The outputs are quantitatively compared in terms of confusion matrices (Thiery et al., 2006).

6.2.2 Description of the study area

The Barcelonnette Basin is representative of climatic, lithological, geomorphological and landcover conditions observed in the South French Alps, and is highly affected by landslide hazards (Flageollet *et al*., 1999). The test site extends over an area of about 100 km^2 and is limited to the north-facing hillslope of the Basin (Figure 6.2.1). The Ubaye River depicts the northern boundary while the Sauze torrent delimits the western boundary; the southern and eastern boundaries are represented by high elevation crests of limestones, sandstones and flyschs. The test site can be subdivided in two geomorphological units separated by a major fault in the North/South direction. The Eastern Unit is dominated by allochtonous flysch outcrops, while the Western Unit is composed of autochtonous Callovo-Oxfordian marls.

The Eastern Unit (about 40 km^2) is drained by the Abriès torrent which cuts a dissymmetric valley in highly fractured flyschs. The gentle slopes (8-27°) are covered by moraine deposits (of thickness ranging from 2 to 15 m) and by coniferous forests or grasslands; these slopes are affected by shallow rotational or translational slides triggered by the torrent undercutting. At the opposite, the steep slopes (30-70°) are characterized by bare grounds and affected by rock-falls in the flyschs.

The Western Unit (60 km^2), drained by four main torrents, presents an irregular topography alternating steep convex slopes, planar slopes and hummocky slopes. The steepest convex slopes (> 35°) are carved in black marl outcrops and are very often gullied in badlands or affected by rock-block slides or complex slides (Malet *et al*., 2005). The planar slopes (8°-17°) are composed of thick moraine deposits (from 10 to 20 m), are very often cultivated and affected by rotational or translational slides. The hummocky slopes are generally covered by forests and/or by natural grasslands and affected by large relict landslides and/or by surficial soil creep. Most of the landslides within the Western Unit are located along the streams or on the gentle slopes where the contact of the moraine deposits on the black marl bedrock creates a hydrological discontinuity favourable to slope movements.

The test site is situated in the dry intra-Alpine zone, characterized by a mountain climate with a Mediterranean influence. Highly variable rainfall amounts (400 to 1300 $mm.year^{-1}$) are observed, with intense storms during summer and autumn. However, as pointed out by

Flageollet *et al*. (1999), the landslides are not controlled only by the climatic conditions; slope instability can occur after relatively dry periods whether or not preceded by heavy rainfalls (van Asch and Buma, 1997; Buma, 2000; Maquaire *et al*., 2003).

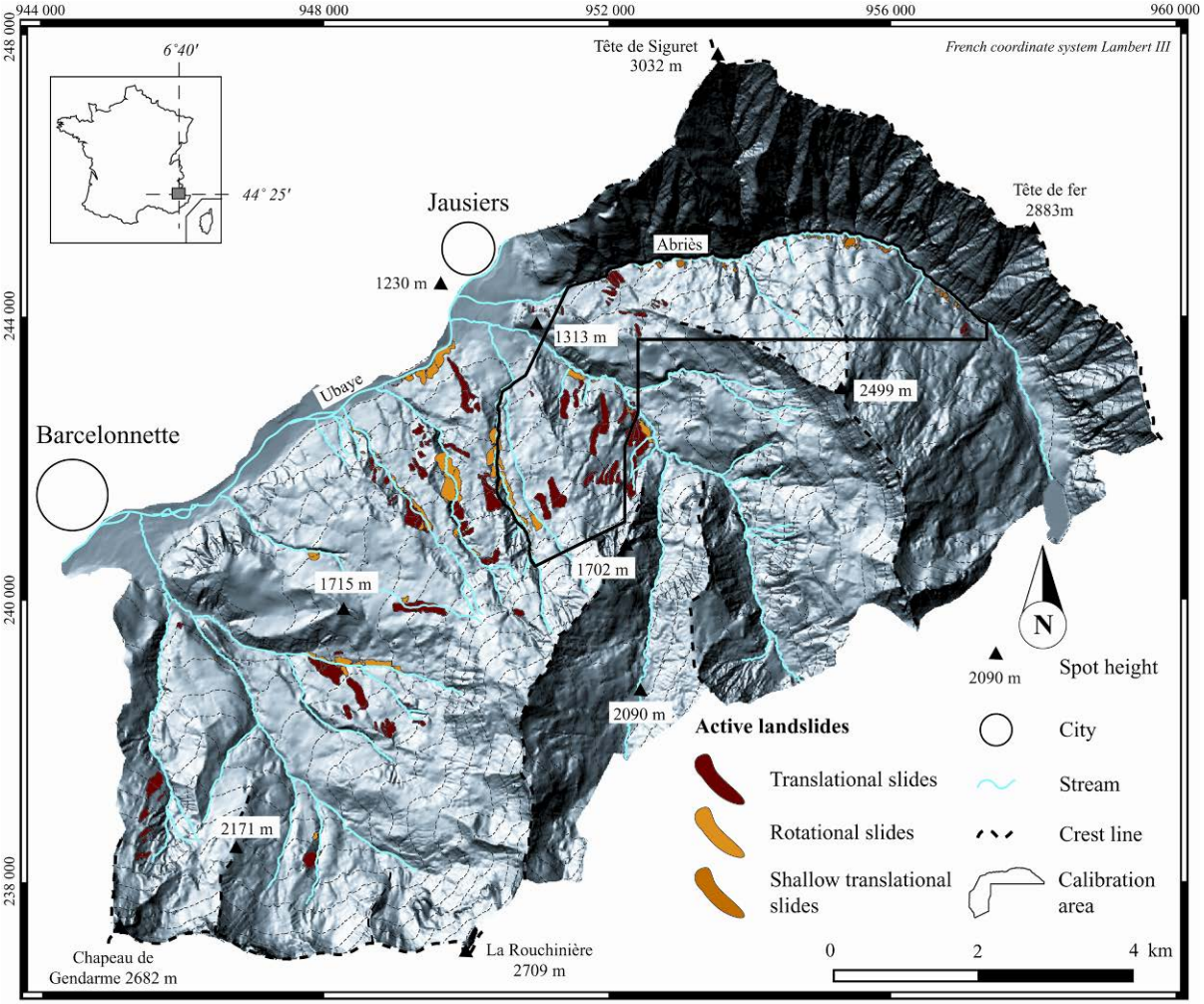


Figure 6.2.1 Relief of the North-facing hillslope of the Barcelonnette Basin and observed distribution of landslides.

6.2.3 Input data

a) Landslide inventory

An inventory of active and relict landslides has been compiled at 1:10,000 scale by associating aerial photo-interpretation (API), fieldwork survey and analysis of past available references and all documents were compared to reduce the uncertainty lied to the expert in charge of mapping.

Aerial photo-interpretation was carried out on photographs from IGN (year 2000) at a 1:25,000 scale. Fieldwork was carried out between the months of July 2002 and July 2003 to complete the photo-interpretation. To reduce uncertainty in the mapping process, two levels of confidence (landslide recognition or not) were defined for the photo-interpretation and three levels of confidence (high, medium, low) were defined for the field work survey. These coefficients were combined in a Mapping Confidence Index (MCI) in three levels (high, medium or low confidence). 392 landslides were recognized and digitized in a GIS environment by this method, with respectively 241 landslides classified with a high MCI (61%), 118 landslides with a medium MCI (30%) and 33 landslides with a low MCI (9%). Among the 241 landslides mapped with a high MCI, 8% are considered as relict, 7% are considered as latent and 85% are considered as active. The active landslides can be grouped in 6 types according to the typology of Dikau *et al*. (1996).

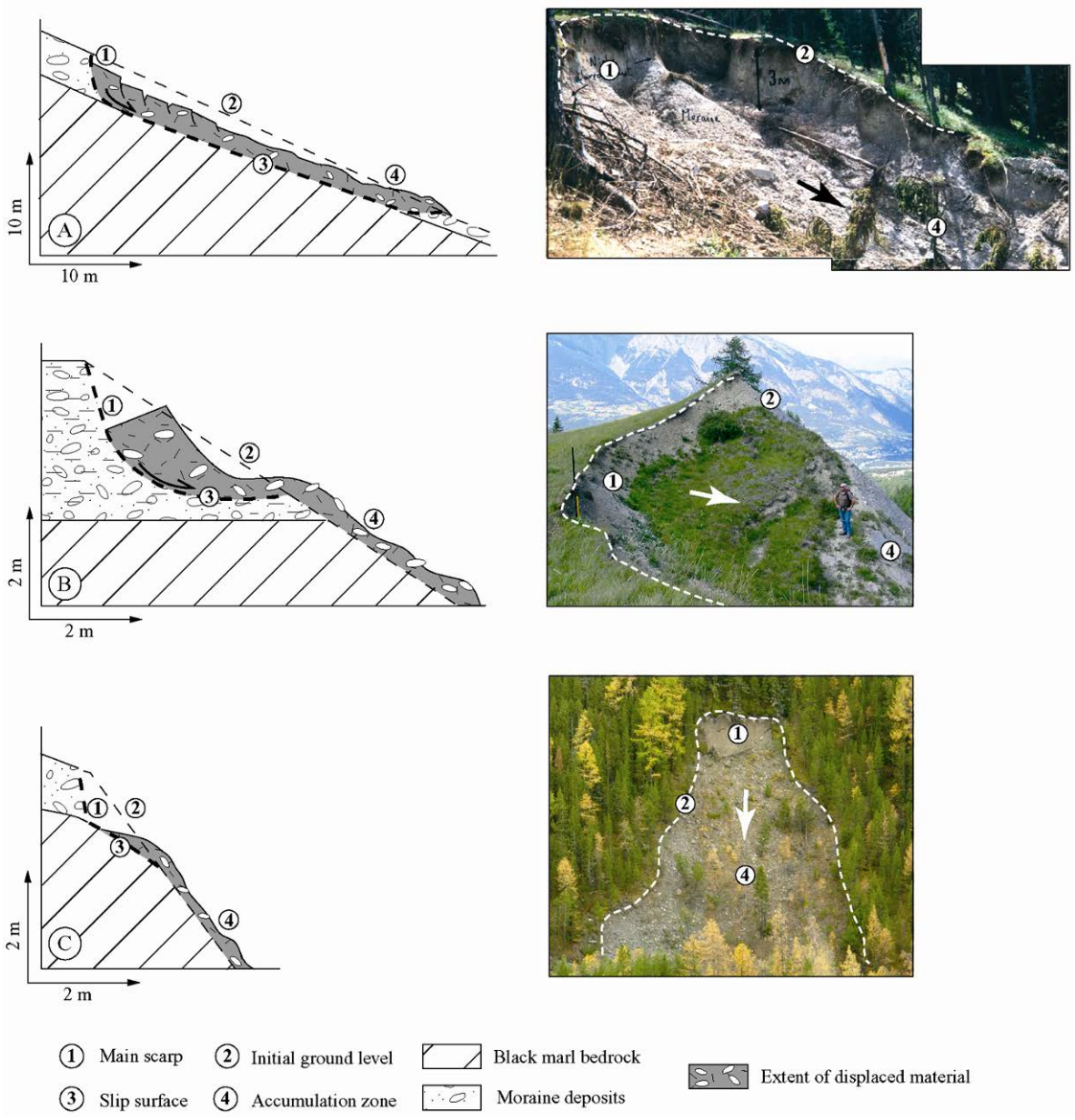


Figure 6.2.2 Main types of landslide observed in the Barcelonnette Basin. A: Translational slide in the Bois Noir catchment; B: Rotational slide in the Poche Torrent; C: Shallow translational slide in the Abriès Torrent

To analyse landslide susceptibility only the active landslides were taken into account. As well rock-falls, rock-block slides and complex slides were not used because of their low statistical significance. Figure 6.2.2 presents the main characteristics of the landslides selected for this study. Shallow translational slides are mainly located on steep slopes along streams, are relatively small, and occur on the weathered bedrock or in moraine deposits. Rotational slides are located along streams but on more gentle slopes than the shallow translational slides. They occur principally in moraine deposits or at the contact with the bedrock. Translational slides are located on more gentle slopes at the contact with the bedrock, and their sizes are very variable.

b) Landslide predisposing factors

The statistical analysis of the landslide inventory has allowed to outline the main predisposing factors (*PV*) to introduce in the probabilistic model. The thematic maps (Table 6.2.1) are derived (i) from available national databases, (ii) from a photo-interpretation analysis, (iii) from satellite imageries, and (iv) from fieldwork surveys. The relief maps are derived from a handmade DTM obtained by the digitisation of elevation lines from 1:10,000 scale topographic maps; the very dense grid of triplets has been interpolated by kriging at a 10m resolution. Three *PVs* (slope gradient, slope profile curvature, slope plan curvature) are derived from the DTM. The *PVs* related to geology are the lithology, the type of superficial formations and the thickness of superficial formations. The lithological map has been derived from the main lithological units observed on a geological map produced by the French Geological Survey completed by fieldwork. The typology of superficial formations is derived from the work of van Westen (1993) funded on the segmentation of the landscape in homogeneous geomorphological macroareas associating closely facies and shape. The thickness of superficial formations is derived from direct observations of cross-sections. Finally, the landcover map has been produced by processing Landsat ETM+ image fused with a SPOT P image, and the boundaries of the homogeneous landcover units were corrected by photo-interpretation.

Table 6.2.1 Input data for assessment of landslide susceptibility. API: Aerial Photo-Interpretation; SII: Satellite Imagery Interpretation

Themes	Variables	Source of information and methods used
Landslide map	LA	API + field survey + available documents
Relief	Slope gradient (SLO) Slope curvature (CUR)	DTM elaborated by digitization and interpolation of elevation lines extracted from topographical maps (1:10,000)
Geology	Lithology (LIT)	Interpretation of geological map + field survey
	Superficial formation (SF)	Interpretation of geological and geomorphological maps + field survey
	Thickness of superficial formation (SFT)	Field survey
	Bedding (BED)	Interpretation of geological map + field survey
Hydrology	HYD	API + Topographical maps
Landcover	LAND	SII + API + field survey

6.2.4 A workflow for the identification of the best set of variables for landslide susceptibility assessment

a) Weights-of-Evidence (WofE): background

Weights-of-Evidence (WofE) is a quantitative '*data-driven*' method used to combine datasets. The method, first applied in medicine and in geology (Bonham-Carter, 1994), uses the log-linear form of the Bayesian probability model to estimate the relative importance of evidences by statistical means. This method was first applied to the identification of mineral potential (Bonham-Carter et

al., 1990) and then to landslide susceptibility mapping (van Westen *et al*., 2003; Süzen and Doruyan, 2004).

The concepts of Prior Probabilities (*PriorP*) and Posterior Probabilities (*PostP*) are the most important concepts in the Bayesian approach. *PriorP* is the probability that a Terrain Unit (TU) contains the *RV* before taking into account the *PVs* and its estimation is based on the *RV* density for the study area. This initial estimate can be modified by the introduction of other evidences. *PostP* is then estimated according to the *RV* density for each class of *PV*. The model is based on the calculation of positive W_+ [1] and negative weights W_- [2] which magnitude depends on the observed association between *RV* and *PV*.

$$W_+ = \ln \frac{P(B|Rv)}{P(B|\bar{R}v)} \quad (6.2.1)$$

$$W_- = \ln \frac{P(\bar{B}|Rv)}{P(\bar{B}|\bar{R}v)} \quad (6.2.2)$$

In equations (6.2.1) and (6.2.2), B is the class of *PV* and the sign'¯' represents the absence of *PV* and/or *RV*. The ratio is called odds and is a measure of the probability that the *RV* is present to the probability that the *RV* is absent (Bonham-Carter, 1994). The weights-of-evidences for all *PVs* are then combined using the natural logarithm of odds (called logit) in order to estimate the conditional probability of landslide occurrence. When several *PVs* are combined, areas with respectively high or low weights correspond respectively to high or low probabilities of presence of *RV*.

b) Workflow

The strategy used to develop the workflow is a three steps procedure, and consists in the:

i. Identification of the best way to calculate landslide prior probabilities based on the characteristics of the landslide inventory (*RV*);
ii. Identification of the most relevant combination of predisposing terrain factors (*PVs*) avoiding the CI violation;
iii. Evaluation of the performance of the statistical simulations and comparison of the simulated susceptibility maps.

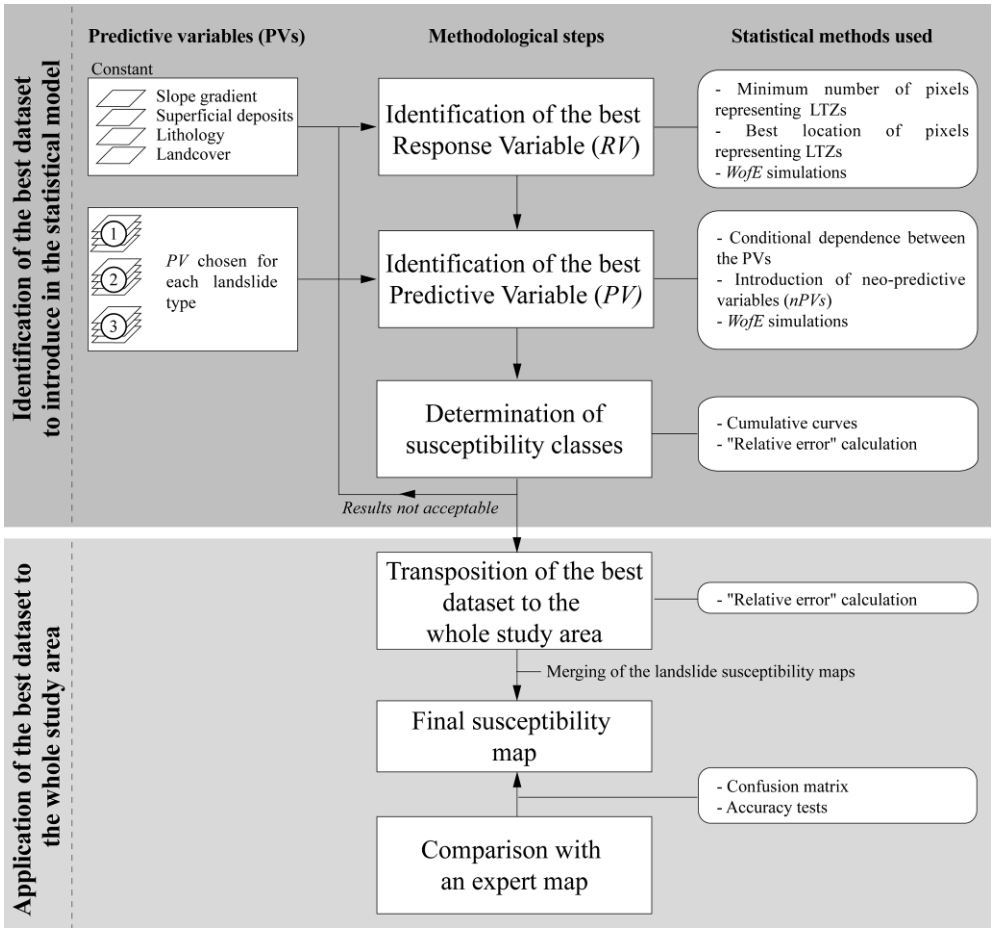


Figure 6.2.3 Flow-chart of the proposed strategy for landslide susceptibility assessment with the WofE model.

To evaluate the robustness, the model was tested on a representative area of the Barcelonnette Basin characterized by the occurrence of the three main types of landslides. This test area extends over about 11 km^2 and is representative of the Western and Eastern terrain units described previously.

Each landslide type has been incorporated separately to take into account their specific predisposing factors and a susceptibility map has been simulated for each landslide type. Only the Landslide Triggering Zones (*LTZs*) are introduced in the calculation of *PriorP*. The statistical model is then validated on the whole area, and the results compared to an expert map (Fig. 6.2.3).

The first step of the procedure is (i) to identify the minimum number of pixels able to represent the variability of the predisposing factors within the Landslide Triggering Zones (*LTZs)* and (ii) to identify the best spatial location of pixels to represent the variability of the predisposing factors within the *LTZs*. The minimal number of pixels to introduce in the model is estimated through a random sampling (10 to 100%) of the pixels representing the *LTZs* of each landslide type. The best spatial location of pixels is estimated through the selection of several locations of pixels within the *LTZs* and successive tests. The computations are performed with a set of four *a priori* 'constant' *PVs* (slope gradient, superficial deposits, lithology, landuse) which are necessary to any landslide susceptibility assessments (Catani *et al*., 2005). A landslide susceptibility map is then created for

each computation with these *PVs*. The obtained *PostP* are very low because very few pixels are introduced in the model and are therefore considered as 'scores'. The scores are analysed on a cumulative curve representing the cumulative area vs. *PostP*. The scores are represented in four classes (null, low, moderate, and high susceptibility) as recommended by the French Official Method of Landslide Risk Zoning (MATE/METL, 1999). The relative error ξ is computed to evaluate the performance of the different tests:

$$\xi = \frac{o_v - p_v}{o_v} \quad (6.2.3)$$

where ξ is the relative error, o_v the observed value (number of pixels representing the *LTZ*) and p_v the predicted value (simulated number of pixels classed in the high susceptibility class). If the relative error decreases with the introduction of a *RV*, the simulation is accepted and compared with the following tests (Figure 6.2.3).

The second step of the procedure deals with the identification of the best set of *PVs* to introduce in the statistical model for each landslide types. Each *PV* is introduced successively in the model, and their performance in terms of violation of Conditional Independence (*CI*), of estimation of *PostP* and of the predictive power of the simulations is tested. Computations are performed with the best *RV* dataset. A contingency table and a χ^2-test are generally used to identify *CI* violation. As the χ^2-test is very sensitive to the density of *RV* introduced in the model (Thiart *et al*., 2003), the test may de facto increase the measure of the dependence between two PVs of at least 25 to 30% (Saporta, 1990; Pistocchi *et al*., 2002). Therefore, to decrease *CI* violation, the dependent *PVs* are combined by detailed statistical analyses combined to expert judgements. Susceptibility maps are computed at each step by introducing iteratively each *PV* or *nPV*, and the relative error associated to each simulated maps is calculated.

The third step consists in the evalaution of the performance and the validation of the model. The best combination of *RV* and *PVs* is identified on the calibration area, and then is applied on the total area (Figure 6.2.1, Figure 6.2.4). First, additional χ^2-tests and calculation of ϕ_c coefficients are performed to validate statistically the model with the best combination of *RV* and *PVs*. A *NOT*-test (Thiart *et al.*, 2003) is also calculated in order to measure the *CI* of the simulations by taking into account the variance of *PostP* (Agterberg and Cheng, 2002). Then, the confidence of the simulated susceptibility maps is evaluated by a Student test. This test uses the variance of *PostP* to create a normalized value *NV* to estimate the certainty of the calculation with the null hypothesis H_0: *PostP* = 0. The value of *NV* has to be in the range [1.64-2.33] to have a certainty calculation ranging between 95% and 99% (Bonham-Carter, 1994; Davis, 2002).

Second, the *PostP* values obtained on the calibration area are then applied on the whole study area, and relative errors ξ are calculated for all the *LTZ*s of the North-facing hillslope. If the relative errors are low, the statistical model is considered as robust.

Finally, the susceptibility maps of each landslide type are merged in one unique susceptibility map which is compared to a qualitative susceptibility map produced by applying the French Official Method of Landslide Risk Zoning (MATE/MATL, 1999). A confusion matrix is calculated between the susceptibility classes of both maps (Table 6.2.3; Fielding and Bell, 1997) and four statistical tests as well as the Kappa coefficient are calculated (Table 6.2.4). A *K* value comprised between 0.4 and 0.75 will generally signify a good agreement between the maps (Fielding and Bell, 1997).

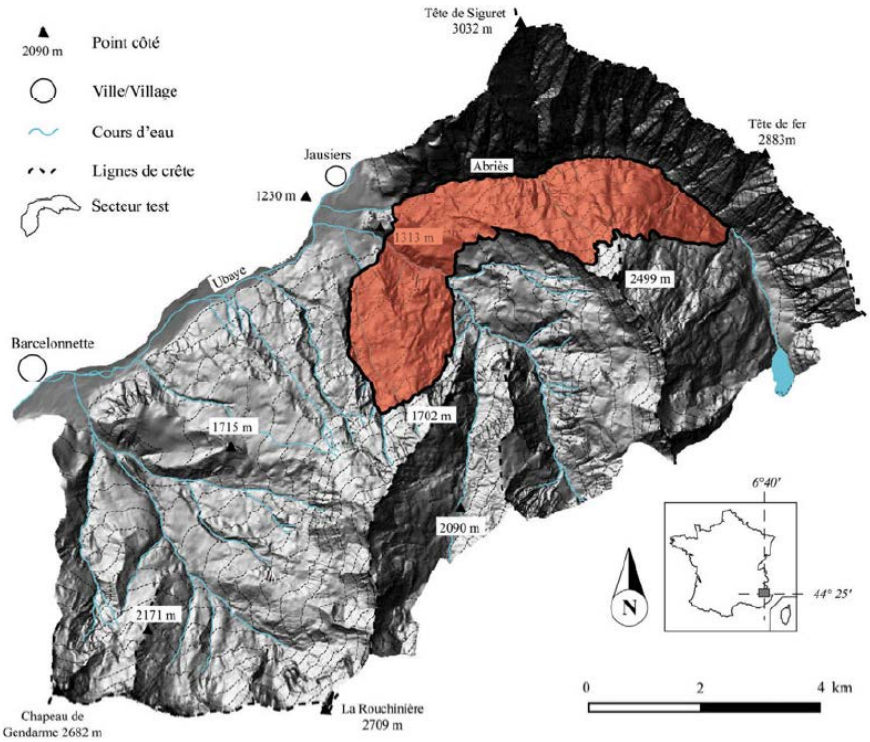


Figure 6.2.4 Extent of the calibration area used to evaluate the performance of the model, within the global study area. The sub-area has been selected because of its statistical representativeness of all classes of predisposing variables observed in the global study area.

c) Identification of the best association of data

In order to identify the minimum number of pixels able to represent the variability of the predisposing factors within the landslide triggering zones (*LTZs*), several simulations were performed with a maximum number of 460 pixels. The relation between the number of pixels of *LTZs* introduced in the model and the relative error is shown on Figure 6.2.5 for each landslide type. A threshold around 50% of the maximum number of pixels (e.g. around 230 pixels) is identified for the area.

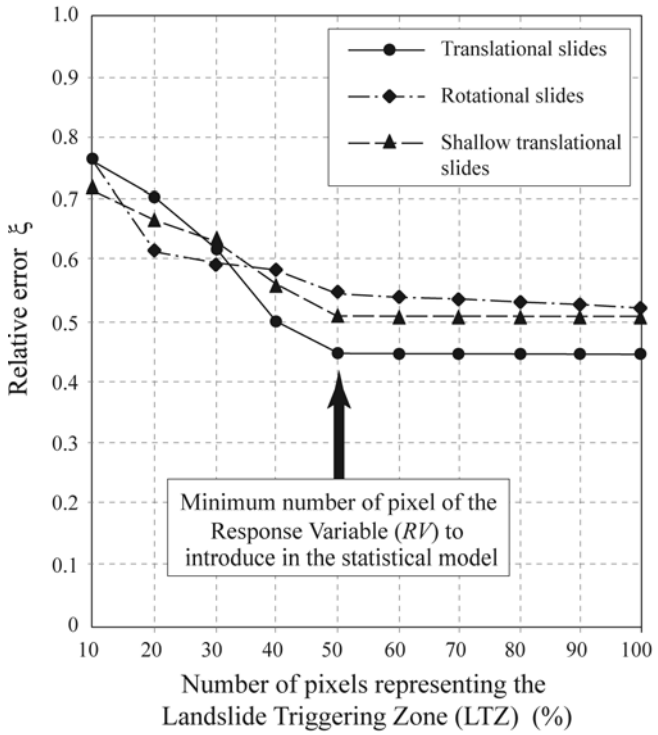


Figure 6.2.5 Relative error ξ associated to the number of RV pixels introduced in the statistical model.

In order to identify the best spatial location of pixels to represent the variability of the predisposing factors within the *LTZs*, several simulations were performed. The best results are obtained with the use of the pixels representing the most frequent combination of *PVs* observed in each *LTZs* (*RV-7*) with a relative error around ξ=0.40 to ξ=0.45.

The selection of the best *PV* dataset is performed by expert judgement taking into account only the predisposing factors and classes associated to each landslide type. For example, the factor 'hydrology' (variable *HYD*) is not introduced for the translational slides in the statistical model because these landslides do not occur nearby streams or torrents; at the opposite, the variable *HYD* is used for the rotational and shallow translational slides through a buffer of 100 m around the streams, because most of these landslides are located nearby streams.

The *PVs* are then chosen according to the statistical tests (χ^2-test, ϕ_c coefficient) and a Multiple Correspondence Analysis (*MCA*) in order to identify statistically the classes of the *PVs* which have to be merged in a neo-Predictive Variable (*nPV)*. Table 6.2.2 summarizes the results for three landslide types. The four axes F1 to F4 explain respectively 46%, 49% and 45% of the total variance associated respectively to the translational slides, the rotational slides and the shallow translational slides. Despite the low contribution of each axis (less than20%) on the cumulated variance, some information emerges from the *MCA.* For example, the axes F1, F2 and F3 associated to the translational slides confirm the relation between slope gradient (*SLO*) and the superficial formations (*SF* and *TSF*). An overview of the *MCA* results indicates that:

- For translational slides, a first group of variables (axis F1) associates the high slope gradients (35°-45°), the colluviums and/or weathered marls and the lowest soil depths. A second group of

variable (axis F2) associates the moderate slope gradients (25°-35°), the moraine deposits and soils of moderate thicknesses. A third group of variables (axis F3) associates low to moderate slope gradients (10°-15°), moraine deposits and the highest soil depths.

- For rotational slides, a first group of variables (axis F1) associates the moderate and high slope gradients (20°-35°) and the weathered marls. A second group of variables (axis F2) associates the high slope gradients (30°-35°), the lowest thicknesses of superficial deposits and the slope morphology (slope convexity).
- For shallow translational slides, a first group of variables (axis F1) associates the moderate and high slope gradients (15°-45°), the colluviums and moraine deposits, and the lowest soil depths. A second group of variables (axis F2) associates the lithological outcrops and the bedding (especially the class representing the down dip layers).

The *MCA* proposes some indications on the possible combination of classes for each variable, and allows to justify the definition of a *nPV* with a geomorphological meaning and low redundant information. The proposed *nPV* used in the statistical model are:

- In the statistical model of the translational slides the *nPV* combines (i) moraine deposits with several slope gradient classes [10-15°], [15-25°], [25-35°] and [35-45°] and (ii) colluviums or weathered marls with the slope gradient classes [25-35°] and [35-45°];
- In the statistical model of the rotational slides, the *nPV* combines the slope gradient classes [10-20°], [20-30°] and [30-40°] to all classes of superficial formations;
- In the statistical model of the shallow translational slides, the *nPV* combines the slope gradient classes [15°-25°], [25-35°], [35-45°] and [45-55°] to the superficial formations colluviums, screes and moraine deposits.

Table 6.2.2 Contribution of each variable for axes F1, F2, F3 and F4 for the three landslide types. The results for each class give the structure of each axis. The most contributive variables for each axis are indicated in grey, and are the basis of the definition of the nPVs

	SLO	LIT	SF	TSF	LAND	CUR	BED	Cumulative explained variance (%)
Translational slide								
Axis F1	39.1	0.3	27.0	20.5	13.0	-	-	12.9
Axis F2	40.6	4.2	10.8	43.7	4.6	-	-	25.3
Axis F3	36.3	0.9	16.9	32.2	13.7	-	-	36.1
Axis F4	48.1	0.9	7.9	4.5	38.6	-	-	46.0
Rotational slide								
Axis F1	33.9	14.8	26.8	2.6	19.8	2.0	-	16.4
Axis F2	26.6	0.8	5.8	35.2	7.8	24.3	-	28.4
Axis F3	24.2	29.3	2.2	40.2	3.9	0.1	-	39.8
Axis F4	35.8	0.9	1.7	47.7	0.1	13.6	-	49.3
Shallow translational slide								
Axis F1	18.7	11.9	23.5	17.4	10.7	4.2	13.5	14.8
Axis F2	19.9	20.9	8.7	11.8	9.9	6.3	22.4	26.6
Axis F3	26.7	0.0	16.5	13.8	32.7	1.0	9.2	36.1
Axis F4	40.3	0.0	2.1	17.2	24.3	6.3	9.7	44.5

The best association *nPV-PV* is tested by introducing iteratively each variable in the statistical model, for each landslide type, and analyzed on the basis of the cumulative curves associated to each WoFe simulations (Figure 6.2.6) and divided in four susceptibility classes. If the relative error does not decrease with the *nPV*, the simulation is rejected; at the opposite, if the relative error decreases, the simulation is accepted and compared to the following.

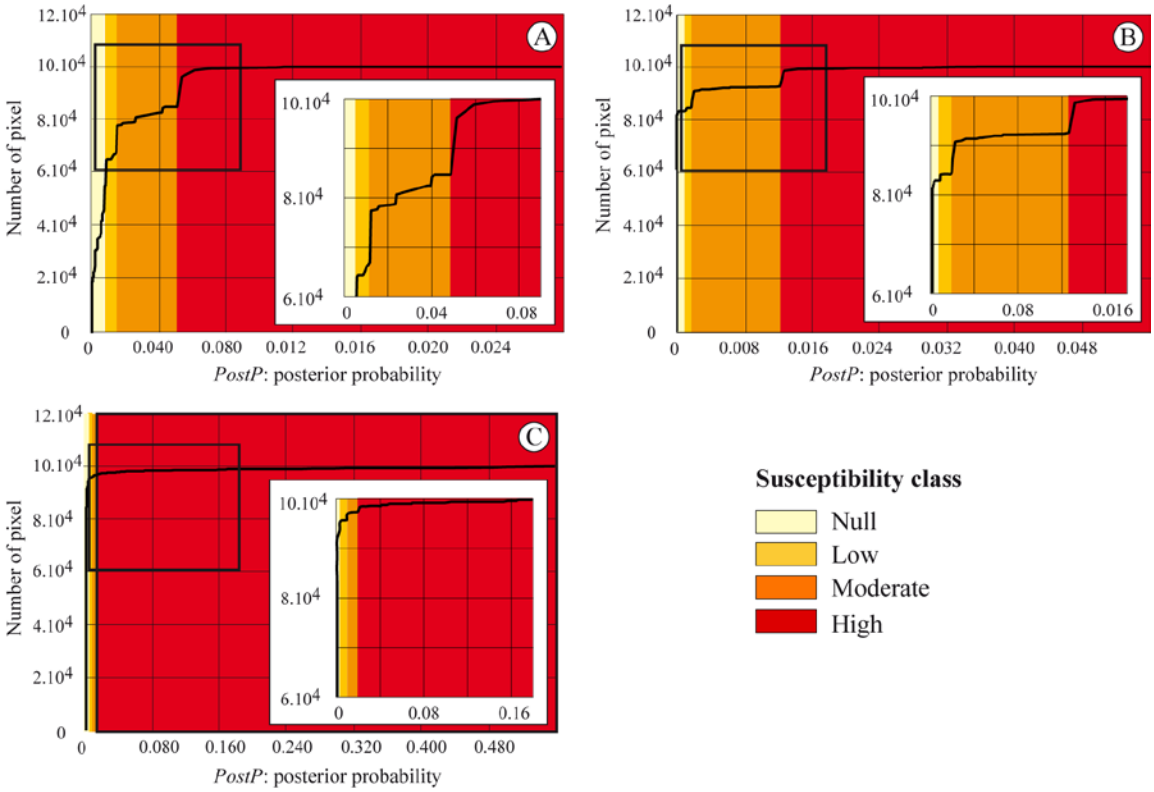


Figure 6.2.6 Cumulative curves of the best simulation for landslide susceptibility assessment in the calibration area. A: Translational slides; B: Rotational slide; C: Shallow translational slides

Figure 6.2.7 presents the susceptibility maps obtained for the translational slides with and without the introduction of the *nPV* on a small catchment of the calibration area. The predictive power of the statistical model increases with the introduction of *nPVs* for each landslide type with a respective decrease of relative error ξ of 0.10, 0.16 and 0.14 for respectively the translational slides, the rotational slides and the shallow translational slides.

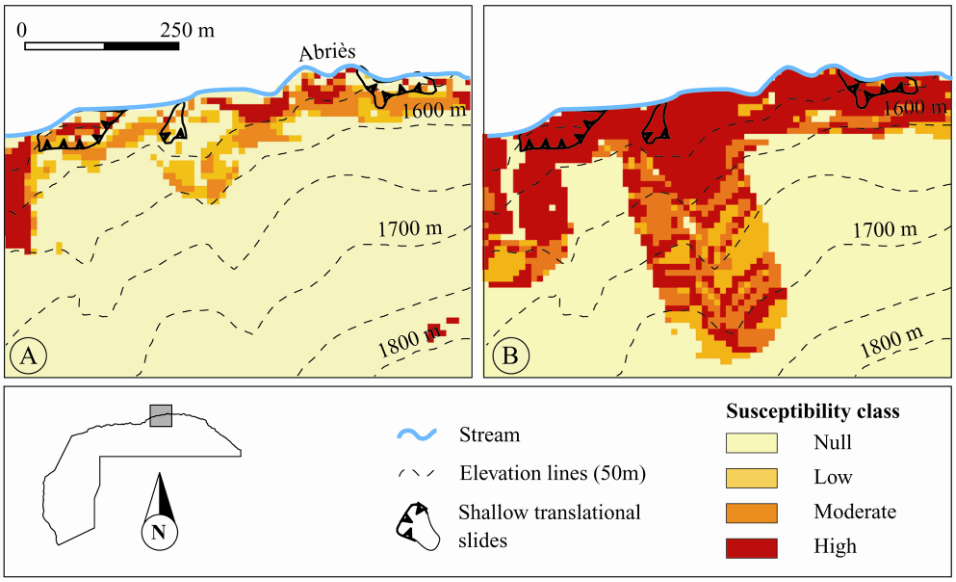


Figure 6.2.7 Example of WoFe statistical simulation performed without (A) and with (B) the introduction of a nPV. A: Statistical simulations with the PVs SLO+SF+LITH+LAND. B: Statistical simulations with the PVs nPV+LAND+CURV+BED

d) Validation of the statistical model

To test the certainty of the simulations, additional χ^2-test and ϕ_c coefficient, and a *NOT*-test have been computed. The statistical confidence of the high susceptibility class is then evaluated with a Student test, for the best statistical model with or without a *nPV*. For a confidence threshold of 0.95, the simulations indicate a percentage of presence of the high susceptibility class in the confidence zone of respectively 87.5%, 88.7% and 70.8% for respectively the translational slides, the rotational slides, and the shallow translational slides. Consequently, from a statistical viewpoint, the high susceptibility classes identified by the models using a *nPV* are certain. These tests are only relative indicators and the simulations have to be evaluated against geomorphological evidences and expert knowledge.

The surfaces of respectively high, moderate and low landslide susceptibility attain respectively 2.0 km^2, 1.2 km^2 and 0.8 km^2 for the translational slides, 0.9 km^2, 0.3 km^2, and 0.15 km^2 for the rotational slides, 0.9 km^2, 0.35 km^2 and 0.25 km^2 for the shallow translational slides in the calibration area.

Then, the weights identified on the calibration area have been applied to the North-facing hillslope of the Barcelonnette Basin for each class of *PVs* and *nPVs* (Figure 6.2.8). The statistical models recognize most of the landslide triggering zones (*LTZs*) with relative errors ξ of respectively 0.23, 0.25 and 0.22 for respectively the translational slides, rotational slides, and shallow translational slides (Table 6.2.3).

Table 6.2.3 Relative error ξ associated to the best simulations for the calibration area and the total North-facing hillslope of the Barcelonnette Basin for both the landslide triggering zone (LTZ) and the total area of landslides (L). [a] Simulations with all pixels of RV-7, [b] Simulations with 50% of pixels of RV-7

	TS *(nPV+LITH+LAND)*				RS *(nPV +LAND+CUR)*				STS *(nPV +LAND+CUR+BED)*			
	LTZ^a	LTZ^b	L^a	L^b	LTZ^a	LTZ^b	L^a	L^b	LTZ^a	LTZ^b	L^a	L^b
ξ (calibration area)	0.18	0.18	0.41	0.41	0.16	0.18	0.34	0.36	0.14	0.17	0.09	0.12
ξ (N-facing slope)	0.23	0.23	0.47	0.47	0.21	0.25	0.33	0.37	0.22	0.25	0.26	0.28

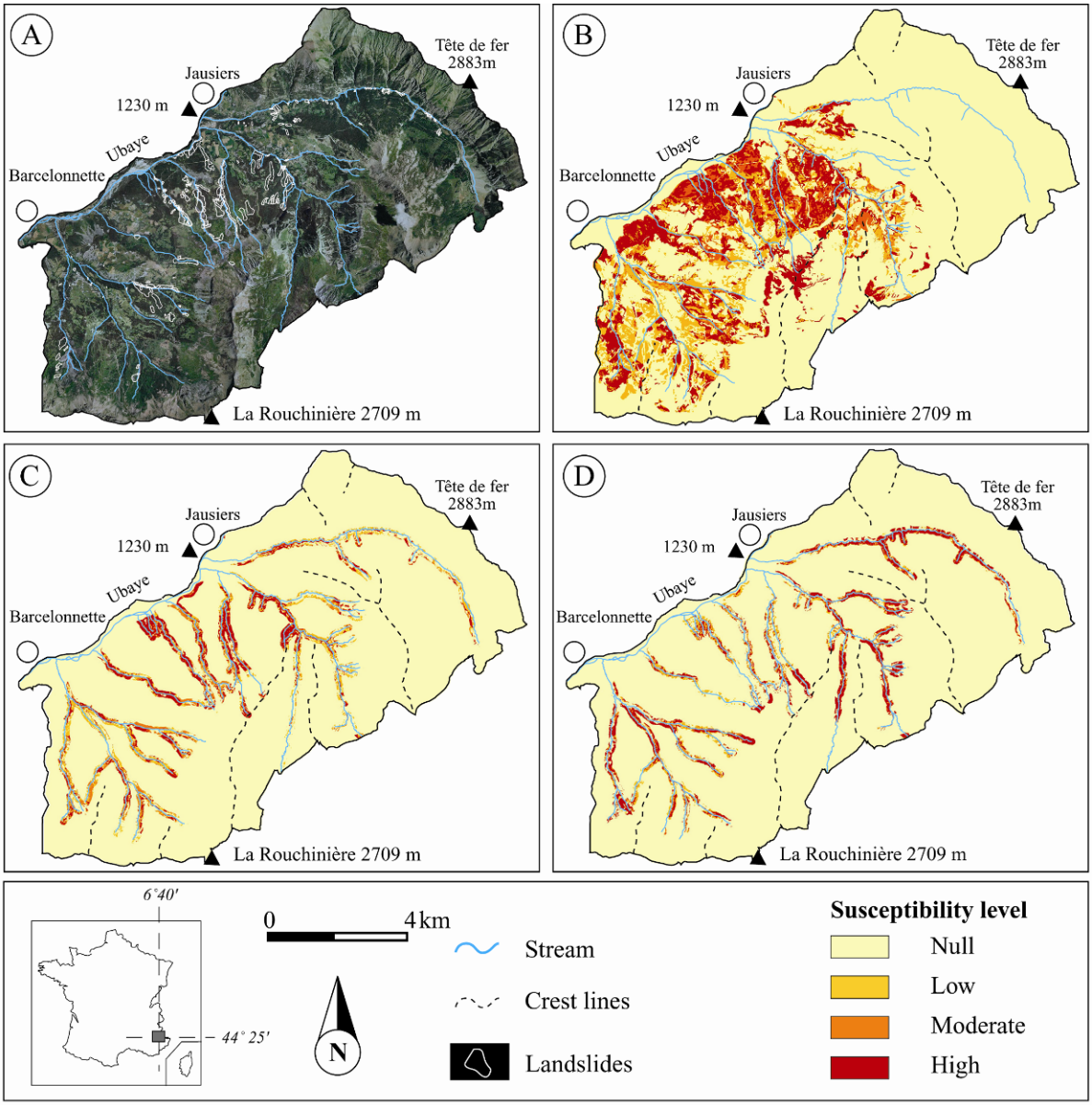


Figure 6.2.8 Susceptibility map simulated for the different landslides for the North-facing hillslope of the Barcelonnette Basin. A. Orthophoto-map. B. Translational slides. C. Rotational slides. D. Shallow translational slides

A statistical test based on a random sampling of *RV-7* has also been performed (Chung and Fabbri, 1999; Remondo *et al*., 2003): 50% of *RV-7* are introduced in the model and the relative error ξ is computed with all *LTZs* and all the total landslide areas (*L*). The relative errors ξ for *LTZs* and *L* are similar for all simulations stressing the robustness of the strategy to select the best *RV* and *PVs* dataset for landslide susceptibility assessment.

6.2.5 Comparison of two statistical susceptibility models to an expert susceptibility model

To constrain the estimation of susceptibility over the study area, two other modeling approaches have been tested, using exactly the same dataset of controlling variables. The second statistical model is a fuzzy logic approach (Thiery et al., in review); the third method is the mapping of susceptibility using the expert methodology *PPR* outlined in the *French Official Method of Landslide Risk Zoning* (MATE/METL, 1999).

a) Fuzzy logic: background

Application of different fuzzy memberships to the Barcelonnete case study are adapted from the results of the Weight-of-Evidence analysis (Thiery et al, 2007) outlined above. This analysis has indicated that, despite a robust methodology to infer a relevant statistical model, some weights still overestimate or underestimate the susceptibility of some sub-areas. Therefore, a sensitivity analysis on the fuzzy memberships has been carried out in order to identify the classes of predictive variables and the "*fuzzy inference network*". The results are validated by statistical tests and compared.

In classical theory, the membership of a variable is defined binary (true =1 or false =0). In the fuzzy approach, the membership is expressed on a continuous scale from 0 (full non membership) to 1 (full membership). The membership function can be expressed as:

$$\mu_A(x) : X \rightarrow [\,0,1] \quad (6.2.4)$$

where x is the variable, *A* is the fuzzy membership function and *X* is the universe of discourse in the interval [0,1].

The grade of membership reflects a kind of ordering that is not based on probability but on admitted possibility. Generally a low value (0 or near 0) is accorded for objects or classes which do not belong to the fuzzy set. Inversely, the grade of membership is large (1 or near 1) for objects or classes which fully belong to the fuzzy set. Thus, based on expert opinion, individual classes of variables can be evaluated regarding their membership in a fuzzy set. The membership always reflects a general hypothesis. In this study, the general proposition for all predictive variables is "*find favourable landslide location*". For categorical variable membership is not always linear but is function of the subjective judgement about the hypothesis. For example, for a categorical variable as lithology the fuzzy membership can be expressed using equation (6.2.5):

$$l = \{(x, \mu_l(x)) \mid x \in X\} \qquad (6.2.5)$$

where $\mu_l(x)$ defines a grade of membership of lithologic formation x in the class "*favourable lithology for landslide location*".

b) Fuzzy logic: modelling strategy

The modelling strategy is subdivided in three steps (Fig. 6.2.9):

i) A sensitivity analysis on the fuzzy membership values to define the best weights reflecting the expert knowledge to introduce in the model;

ii) The integration of the fuzzy membership values by intermediate hypotheses. The first intermediate hypothesis (H1) combine two or more PV in order to obtain a neo-predictive variable (NPV) representing a combination of predisposing factors favourable of landslides prone areas. This NPV can increase the predictive power of maps (van Westen *et al.*, 2003; Thiery *et al.*, 2007). The second intermediate hypothesis H2 represents the combination between the NPV and the other PVs. It reinforces H1. Figure 6.2.8 presents the different steps of the inference network.

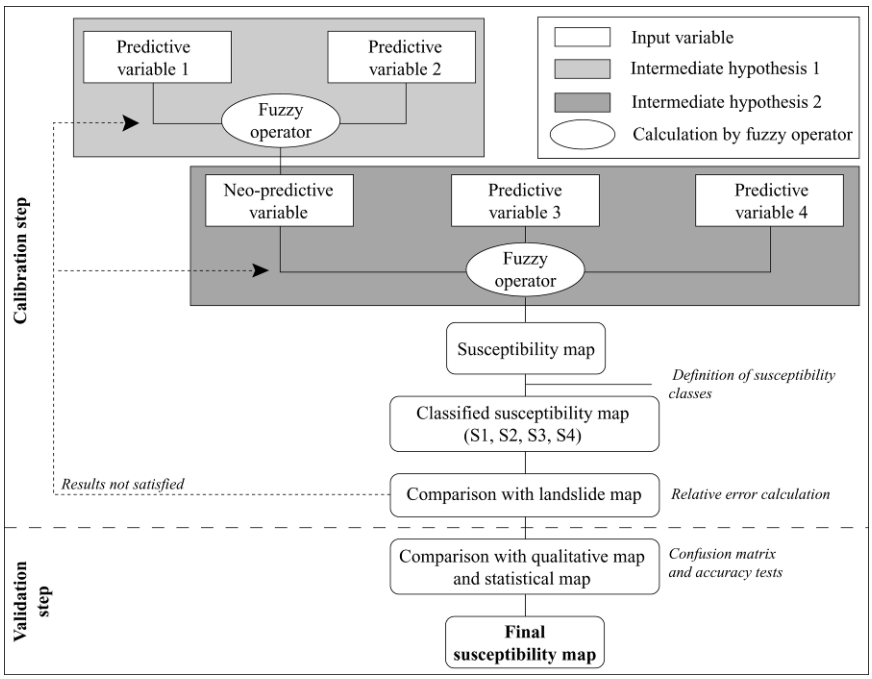


Figure 6.2.9 Schematic inference network to map landslide susceptibility in the Barcelonnette basin by fuzzy logic approach

iii) The final membership values are analysed with a cumulative curve representing the cumulative area versus the final membership values. The susceptibility simulation is divided in four classes: null (S1), low (S2), moderate (S3) and high (S4). The ranges of the

categories were derived from the histogram of the estimated susceptibility to landslides. Results are analyzed following the calculation of the relative error ξ (6.2.8). The calculation is performed by comparing the number of pixels of observations and the number of pixels of predicted value:

$$\xi = \frac{o_v - p_v}{o_v} \qquad (6.2.6)$$

where ξ is the relative error, o_v the observed value (number of pixels of the landslide triggering zone) and p_v the predicted value (number of pixel of the high susceptible class in the landslide triggering zones). Five accuracy statistic tests recommended by Fielding and Bell (1997) are calculated either: correct classification rate, misclassification rate, sensitivity, specificity, and the kappa *K* coefficient. The results show the correlation and the difference between the maps.

b) Expert heuristic model - PPR: background

The expert heuristic map was created using the French official method PPR (MATE/MATL, 1999). It requires a large overview of the area by identifying sectors with homogeneous environmental characteristics, and takes into account the possibility of landslide development (up- and downhill) for a 100 years period (human scale). Thus, the PPR methodology takes explicitly into account the temporal dimension by applying a buffer zone on the areas considered with a high susceptibility; in our case a buffer zone of 20 m is applied around the slope face considered with a high susceptibility. The expert rules have been coded in an ArcGIS script in order to automatically combine all the information. The expert rules used in the analysis are depicted in Table 6.2.4 defined for each landslide type. Each polygon represents a homogeneous area defined by the mean slope gradient, the slope aspect, the landuse and the land degradation. This database contains information about the (i) landslide types and activities, (ii) the main environmental factors considered as prone to one or several type of instabilities, and (iii) the considered susceptibility class. If the homogeneous area is subjected to instabilities, the susceptibility is high.

Table 6.2.4 Expert rules and associated environmental conditions used to create the direct susceptibility map

Susceptibility class	**Expert rule**	**Environmental conditions**		
		Deep translational slide	Shallow translational slide	Shallow rotational slide
S0: no susceptibility	Environmental conditions favourable to slope stability. No possibility of landslide developments for the next one hundred years.	SLO: 0-10° SF: moraine deposits, colluviums LAD: landuse not deteriorated, arable land, permanent crop, forest highly maintained	SLO: 0-10° SF: colluviums, moraine deposits LAD: landuse not deteriorated, arable land, permanent crop, forest highly maintained	SLO: 0-10° SF: moraine deposits LAD: landuse not deteriorated, arable land, permanent crop, forest highly maintained IRR: straight slopes

		IRR: straight slopes	IRR: straight slopes HYD: far from torrent (> 100 m)	HYD: far from torrent (> 100 m)
S1: low susceptibility	Environmental conditions are lowly favourable to slope instability. Low possibility of landslide developments for the next one hundred years. Future human and socio-economic developments of the area are possible and subject to specific attention.	SLO: 5-15° SF: moraine deposits, colluviums LAD: landuse lowly deteriorated, pasture, grassland IRR: low presence of slope accidents	SLO: 5-15° SF: colluviums, moraine deposits LAD: landuse lowly deteriorated, pasture, grassland, forests highly maintained IRR: straight slopes, low presence of accidents HYD: bank stream, near torrent (< 100 m)	SLO: 5-15° SF: moraine deposits LAD: landuse not deteriorated, pasture, grassland, forests highly maintained IRR: low presence of slope accidents HYD: bank stream, near torrent (< 100 m)
S2: moderate susceptibility	Environmental conditions are moderately favourable to slope instability. Moderate possibilities of landslide developments for the next one hundred years. Mitigation works are essential for future human and socio-ecomonic developments of the area.	SLO: 10-30° SF: moraine deposits, colluviums LAD: landuse lowly to moderately deteriorated, pasture, grassland, forests lowly maintained IRR: presence of slope accidents, hummocky topography	SLO: 10-30° SF: colluviums, moraine deposits LAD: landuse lowly to moderately deteriorated, pasture, grassland, forests moderately to lowly maintained IRR: straight and moderate steep slopes, presence of slope accidents, hummocky topography HYD: bank stream, near torrent (< 100 m)	SLO: 10-20° SF: moraine deposits LAD: landuse lowly to moderately deteriorated, pasture, grassland, forests lowly maintained IRR: presence of slope accidents, hummocky topography HYD: bank stream, near torrent (< 100 m)
S3: high susceptibility	Environmental conditions are very favourable to slope instability. High possibility of landslide developments for the next one hundred years. Future human and socio-ecomonic developments of the area are impossible.	SLO: > 20° SF: moraine deposits, colluviums LAD: landuse moderately to highly deteriorated, bare soils, forests not maintained IRR: presence of slope accidents, very hummocky topography	SLO: > 20° SF: colluviums, moraine deposits LAD: landuse highly deteriorated, grasslands, bare soils, forests lowly to not maintained IRR: steep slopes, high presence of slope accidents HYD: bank stream, near torrent (< 100 m)	SLO: > 15° SF: moraine deposits LAD: landuse highly deteriorated, grasslands, bare soils, forests not maintained IRR: high presence of slope accidents, hummocky topography HYD: bank stream, near torrent (< 100 m)

The expert map is presented in Figure 6.2.10.

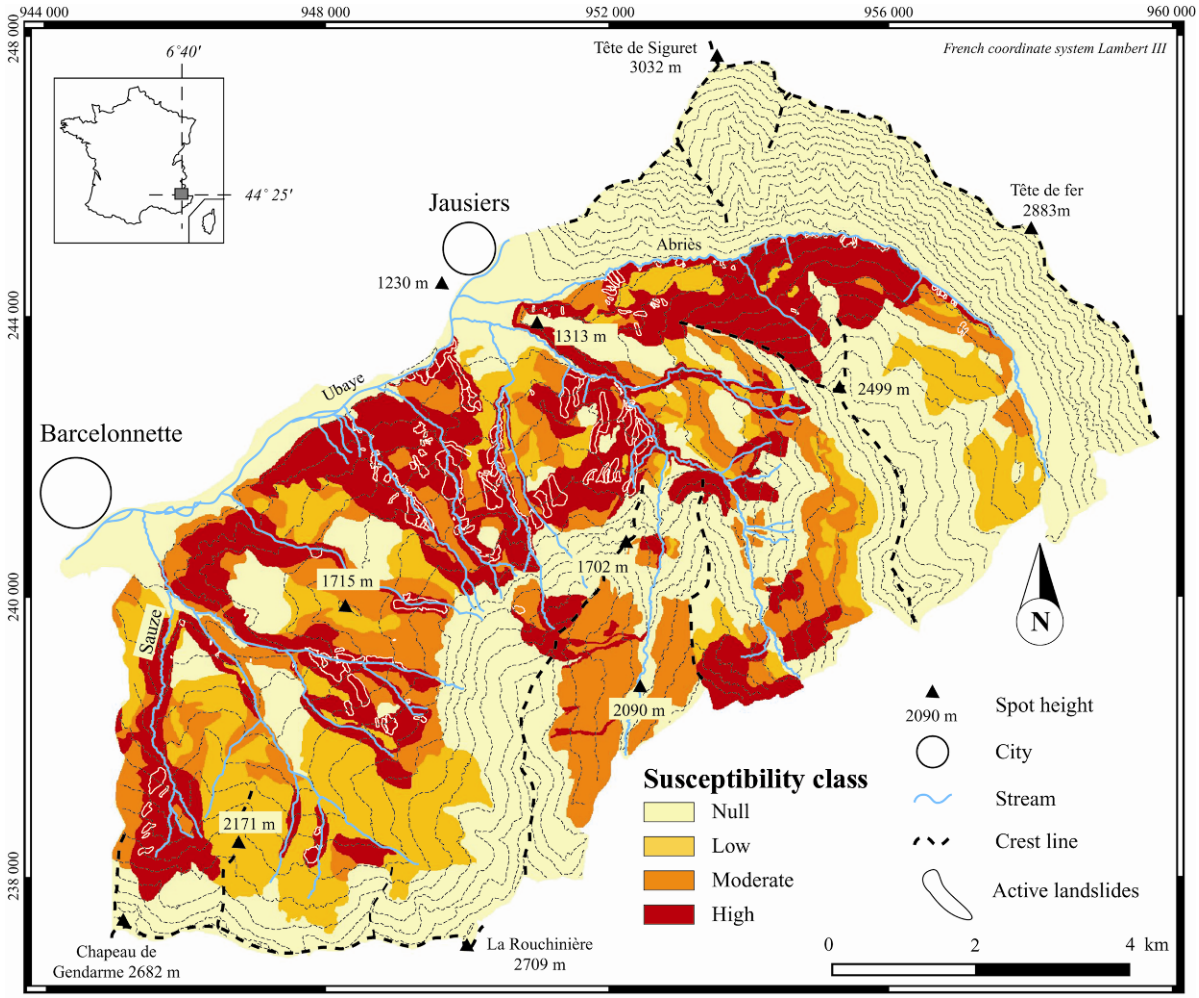


Figure 6.2.10 Expert susceptibility map obtained with the French Official Method of Landslide Risk Zoning

c) Comparison of the statistical susceptibility maps to the expert susceptibility map

The results obtained for each landslide type with the Weight-of-Evidence and Fuzzy Logic models are combined in a final global landslide susceptibility map. The final map obtained with the Weight-of-Evidence model is presented in Figure 6.2.11; this map identifies respectively 17.7 km^2, 5.8 km^2 and 6.9 km^2 for respectively the high, medium and low susceptibility classes (Figure 6.2.11).

The confusion matrix between the simulated map and the expert map (Table 6.2.5) indicates correct classification rates, especially for the high susceptibility class and a Kappa *K* coefficient of 0.41. Nevertheless the statistical map tends to underestimate (i) the surfaces with a low to moderate susceptibility, and (ii) the susceptibility associated to the portion of the areas with slopes gradients of less than 15°. The differences among the two maps for the high susceptibility class are presented on Figure 6.2.12.

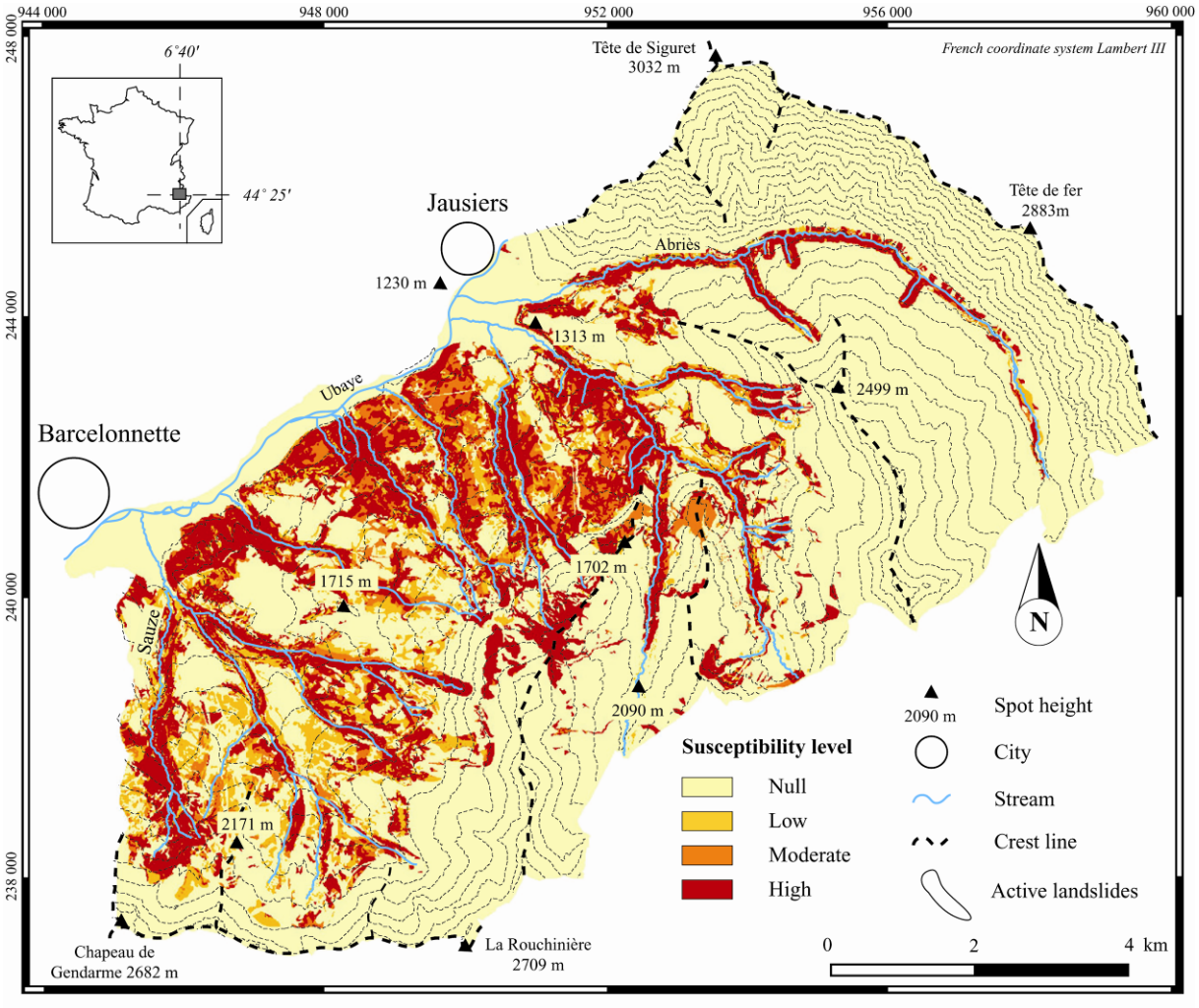


Figure 6.2.11 Final susceptibility map obtained by combination of the three individual landslide type susceptibility maps with the WofE statistical model

This discrepancy may be explained by the methodology used to select the best *RV* (*RV-7*) which mathematically increases the weights of the *PV* combination corresponding to the landslide triggering zone. Also, as only a few *LTZs* are located on slope gradients of less than 15°, the *NPVs* combining *SLO* and other *PVs* tend to underestimate *PostP* for these particular slopes.

Table 6.2.5 Statistical accuracy tests between the simulated susceptibility map and the expert susceptibility map

	Susceptibility class				
	Null	Low	Moderate	High	Global
ccr	0.73	0.81	0.85	0.91	0.61
mcr	0.27	0.19	0.15	0.09	0.45
sensitivity	0.87	0.18	0.08	0.80	0.61
specificity	0.39	0.89	0.95	0.93	0.89
Kappa *K*	0.36	0.08	0.03	0.43	0.41

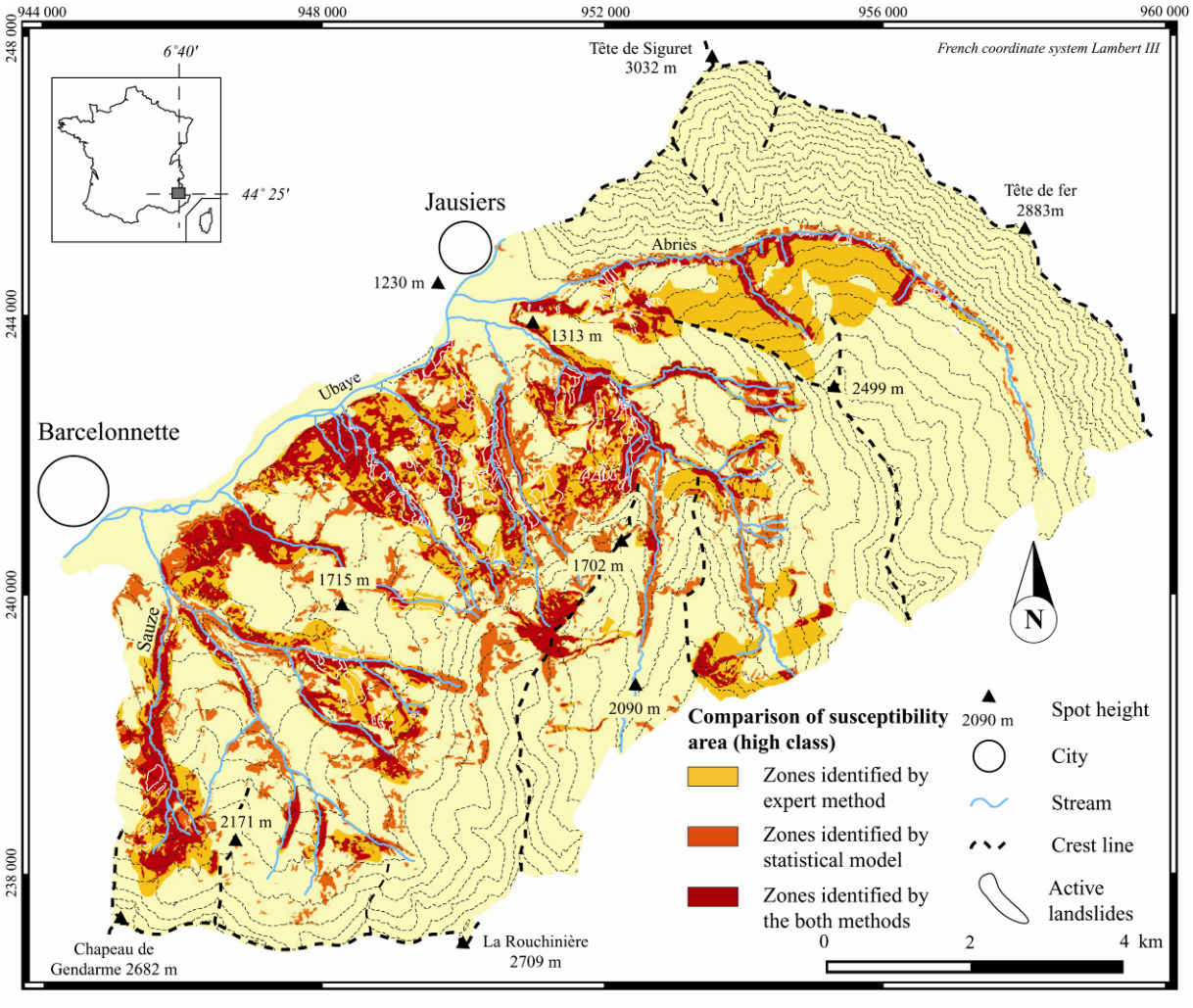


Figure 6.2.12 Difference observed between the expert susceptibility map and the total susceptibility map for the high susceptibility class.

The confusion matrix between the map obtained by the fuzzy logic method, the expert map and the Weight-of-Evidence map is presented in Table 6.2.6.

The different accuracy tests show close results between the susceptibility map obtained by fuzzy approach and bivariate technique (Table 6.2. 5) with a correct classification rate of 0.89, sensitivity of 0.78 and a kappa *K* of 0.71. Kappa *K* coefficient gives an assessment of the improvement of the model predictions over chance. A value over 0.4 indicates a good fit between the two maps.

Results between the qualitative map and the simulated fuzzy map are more heterogeneous. If for S1 and S4, the accuracy test can be considered as good (correct classification rate of 0.8 and Kappa *K* coefficient of 0.47), large differences subsist for S2 and S3. This is related to the general hypothesis of the fuzzy model: "*find favourable landslide location*". In this case, the membership values and their combination with the fuzzy operators have been performed to find principally the most susceptible landslide prone areas (Figure 6.2.13).

Table 6.2.6 Accuracy tests between the fuzzy logic model and the qualitative map and the Weight of Evidence model.. ccr: correct classification rate; mcr: misclassification rate. The definition of the index is detailed in Fielding and Bell (1997)

	Qualitative map					Weight of Evidence map				
	Susceptibility class				General	Susceptibility class				General
	S1	S2	S3	S4		S1	S2	S3	S4	
ccr	0,80	0,85	0,82	0,80	0,82	0,91	0,89	0,88	0,88	0,89
mcr	0,20	0,15	0,18	0,20	0,18	0,09	0,11	0,12	0,12	0,11
sensitivity	0,83	0,14	0,17	0,58	0,64	0,88	0,17	0,49	0,81	0,78
specificity	0,25	0,93	0,89	0,88	0,88	0,98	0,93	0,93	0,90	0,93
Kappa K	0,58	0,08	0,05	0,47	0,52	0,82	0,08	0,55	0,65	0,71

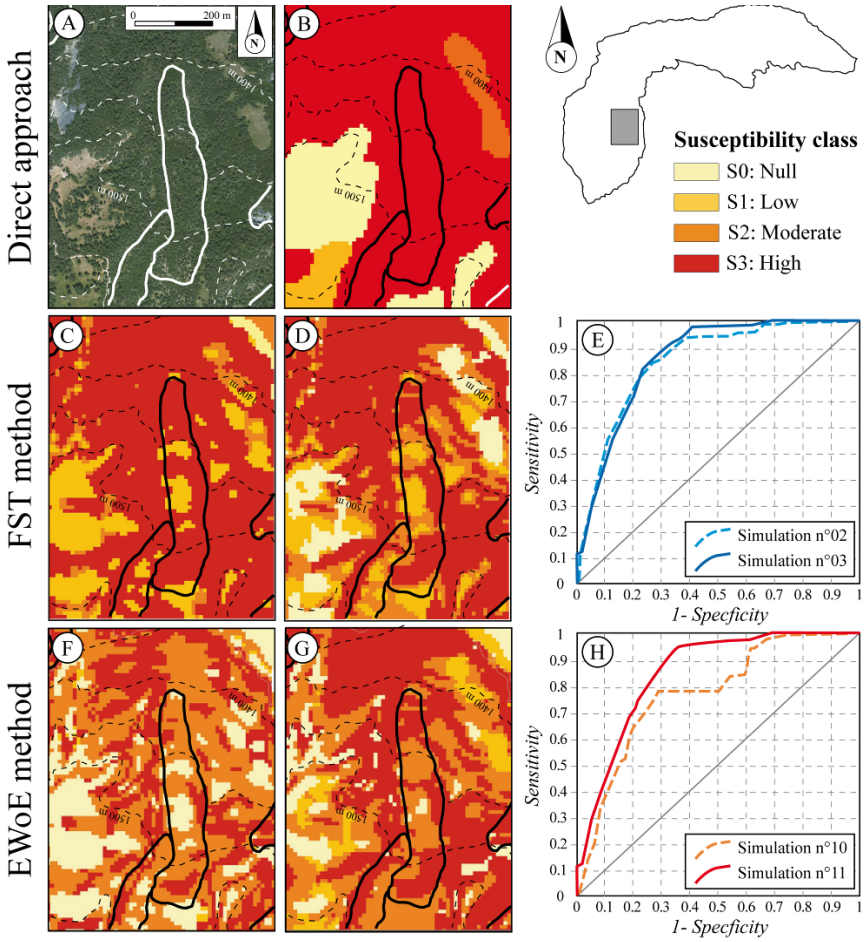


Figure 6.2.13 Examples of susceptibility maps for deep translational slides performed with the direct expert model, the Fuzzy Set theory model and the WofE model. (A) Orthophotograph of the area. (B) LS map obtained with the direct approach. LS map obtained with the FST method: (C) Simulation calculated with the PVs SLO+ FS + LIT + LAND and the operator SUM; (D) Simulation calculated with the PVs NPV-1 + LIT + LAND and the operator γ(0.975); (E) ROC curve for simulation (C) and (D). LS map obtained with the EWoE method: (F) Simulation calculated with the PVs SLO + FS + LIT + LAND; (G) Simulation calculated with the PVs NPV-1 + LIT + LAND; (H) ROC curves for simulation (F) and (G)

Table 6.2.7 Relative error ξ, ROC-AUC and surface of S3 for best combinations of PVs

Fuzzy Set model

Simulation	Landslide	Predictive variables	Fuzzy operator	Relative error ξ	ROC-AUC	S3 (% of area)
N°1	Deep translational slide	SLO + FS + LIT + LAND	PRODUCT	0.54	0.73	10 %
N°2		NPV-1 + LIT + LAND	SUM	0.11	0.86	70 %
N°3		NPV-1 + LIT + LAND	γ (0.975)	0.26	0.85	25 %
N°4	Shallow translational slide	HYD + SLO + FS + LAND + IRR+ DIP	PRODUCT	0.61	0.59	5 %
N°5		NPV-2 + LAND + IRR+ DIP	SUM	0.16	0.92	11 %
N°6		NPV-2 + LAND + IRR+ DIP	γ (0.975)	0.26	0.89	6.5 %
N°7	Shallow rotational slide	HYD + SLO + FS + LAND + IRR	PRODUCT	0.62	0.53	6 %
N°8		NPV-3 + LAND + IRR	SUM	0.18	0.94	21 %
N°9		NPV-3 + LAND + IRR	γ (0.975)	0.25	0.92	9 %

WofE model

Simulation	Landslide	Predictive variables	Relative error ξ	ROC-AUC	S3 (% of area)
N°10	Deep translational slide	SLO + FS + LIT + LAND	0.33	0.76	21 %
N°11		NPV-1 + LIT + LAND	0.22	0.86	20 %
N°12	Shallow translational slide	HYD + SLO + FS + LAND + IRR+ DIP	0.36	0.73	9.5 %
N°13		NPV-2 + LAND + IRR + DIP	0.19	0.92	7 %
N°14	Shallow rotational slide	HYD + SLO + FS + LAND + IRR	0.37	0.56	10 %
N°15		NPV-3 + LAND + IRR	0.24	0.92	9.5 %

As can be observed in Table 6.2.7 and in Figure 6.2.14 and 6.2.15, the Fuzzy Set simulations performed with the operators SUM and γ have the best prediction capacity. For the WofE model, the introduction of a NPV improves the prediction capacity of simulated maps. Figure 6.2.14 shows the best simulated maps produced by Fuzzy Set model and the WofE model for each landslide type, and Figure 6.2.14 shows the best simulated unified landslide maps.

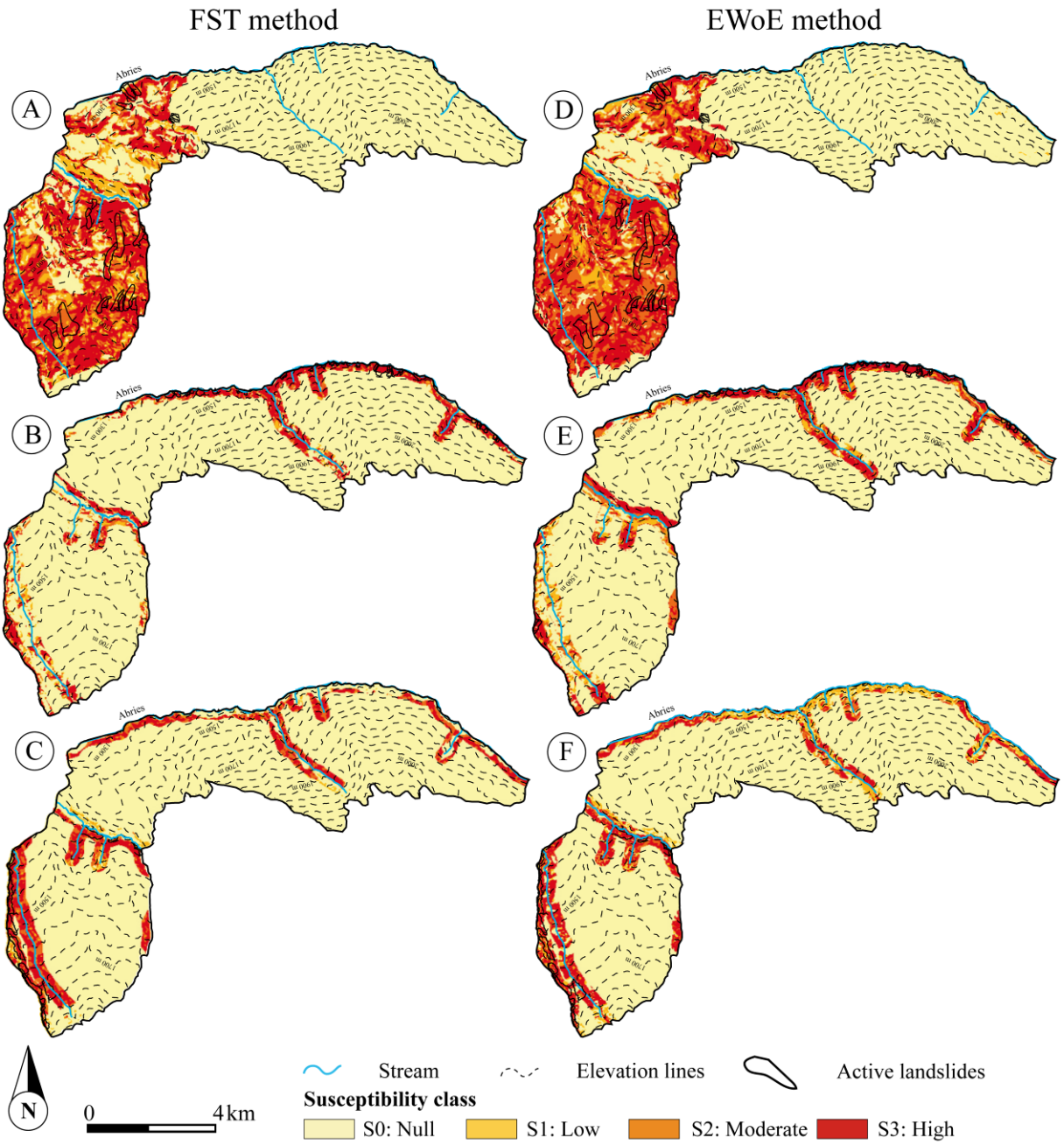


Figure 6.2.14 Landslide susceptibility maps for the landslide types observed on the test site. Susceptibility maps produced with the Fuzzy Set model: (A) Deep translational slides; (B) Shallow translational slides; (C) Shallow rotational slides. Susceptibility maps produced with the WofE model: (D) Deep translational slides; (E) Shallow translational slides; (F) Shallow rotational slides

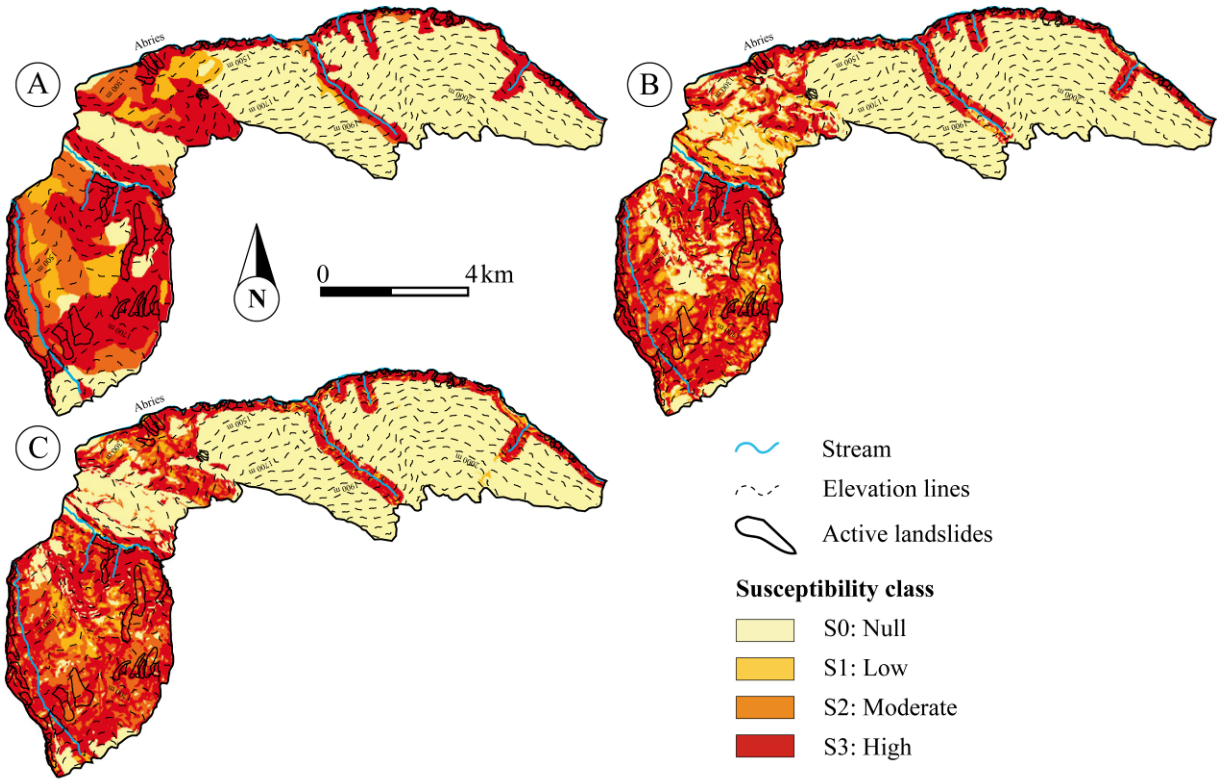


Figure 6.2.15 Final landslide susceptibility map with the highest performance; (A) Landslide susceptibility map produced with the direct expert approach; (B) Landslide susceptibility map produced with the Fuzzy Set model; (C) Landslide susceptibility map produced with the WofE model

The AUCs, calculated on the basis of the ROC curves, are above 0.85 for each simulated landslide susceptibility maps (Figure 6.2.16), representing a good predictive power for the class S3. In others words, the method to introduce the expert rules in the two indirect approaches is reliable for the high susceptibility class.

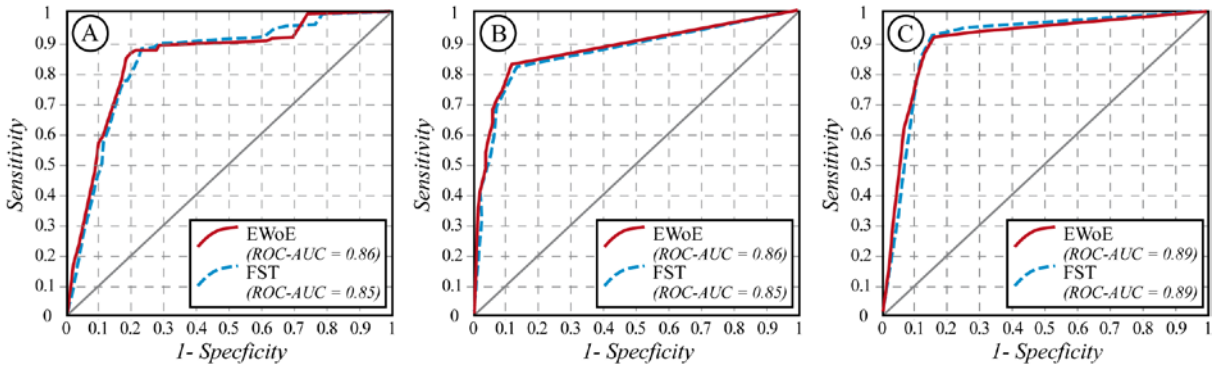


Figure 6.2.16 Tests of the high susceptibility class obtained by indirect approaches and the specific procedure to introduce expert rules by ROC curves. (A) Deep translational slides; (B) Shallow translational slides; (C) Shallow rotational slides

6.2.6 Discussion and summary of conclusions

This study has demonstrated the necessity to use specific and adapted procedures for landslide susceptibility assessment (whatever the type of statistical model used), especially when the environment is complex or if few landslide records are available. The procedure has always to combine several statistical tests and expert judgement. Several other considerations can also be pointed out from this work:

- First, the quality of the simulated susceptibility maps is highly dependent on the landslide type. As many statistical models as landslide types have to be identified because each landslide type is controlled by a specific combination of predisposing factors.
- Second, the quality of the simulated susceptibility maps depends more on the selection of relevant pixels representing the variability of the environmental factors that on the selection of all the landslide pixels.
- Third, the results of χ^2-test and ϕ_c coefficient have to be interpreted with caution, because a few pixels can severely bias the results (Dumolard *et al*., 2003). These tests are just informative and they cannot be used in rigorous terms (Pistocchi *et al.*, 2002). Therefore, instead of not incorporating the pixels posing some problems or of decreasing the total number of *RV* pixels, our procedure intends to combine some classes of the *PVs* which are conditionally dependent. An interesting procedure is then to combine a χ^2-test and a ϕ_c coefficient to a Multiple Correspondence Analysis and an expert judgement in order to select the classes of *PVs* to combine in a neo-predictive variable (*nPV*) with a geomorphological meaning. Expert judgement is critically important in the conception of the statistical model to guide the maps towards geomorphological landslide evidences.
- Fourth, the procedure of calibration/validation of the statistical models is also dependent on the prevalence and susceptibility thresholds observed on the simulated cumulative curves. Bégueria (2006) points out this aspect of the classification/validation procedure which can overestimate some areas in potentially dangerous area and underestimate other. Statistical evaluation of the models based on the ROC curve associated to other accuracy tests may therefore be useful.

References

Agterberg, F.P., Cheng, Q., 2002. Conditional independence test for weights of evidence modeling. Natural Resources Research 11(4), 249-255.

Bégueria S., 2006. Validation and evaluation of predictive models in hazard assessment and risk management. Natural Hazards 17, 315-329.

Bonham-Carter, G.F., 1994. Geographic Information System for geoscientists: modelling with GIS. Pergamon Press, Oxford, 398p.

Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1990. Statistical pattern integration for mineral exploration. In: Gaal, G., Merriam D.F. (Eds.), Computer Applications in Resource Estimation: Prediction and Assessment for Metals and Petroleum. Pergamon Press, Oxford, 1-21.

Buma, J., 2000. Finding the most suitable slope stability model for the assessment of the impact of climate change on a landslide in southeast France. Earth Surface Processes and Landforms 25(6), 565-582.

Catani, F., Casagli, N., Ermini, L., Righini, G., Menduini, G., 2005. Landslide hazard and risk mapping at catchment scale in the Arno River Basin. Landslides: Journal of International Consortium on Landslides. 2(4), 329-343.

Chung, C.F., Fabbri, A.G., 1999. Probabilistic prediction model for landslide hazard mapping. Photogrammetric Engineering and Remote Sensing 65, 1389-1399.

Davis J.C., 2002. Statistics and data analysis in geology, third edition. Wiley J. & Sons, New-York, 638p.

Dikau, R., Brunsden, D., Schrott, L., Ibsen M-L., 1996. Landslides recognition, Identification, Movement and Causes. Wiley & Sons, New-York, 251p.

Dumolard, P., Dubus, N., Charleux, L., 2003. Les statistiques en géographie. Ed. Belin, 240p.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/ absence models. Environmental Conservation 24, 38-49.

Flageollet, J-C., Maquaire, O., Martin, B., Weber, D., 1999. Landslides and climatic conditions in the Barcelonnette and Vars basins Southern French Alps, France Geomorphology 30, 65-78.

Malet, J-P., van Asch, Th.W.J., van Beek, R., Maquaire O., 2005. Forecasting the behaviour of complex landslides with a spatially distributed hydrological model. Natural Hazards and Earth System Sciences 5: 71 – 85.

MATE/MATL, 1999. Plan de Prévention des Risques (PPR) : Risques de Mouvements de terrain, Ministère de l'Aménagement du Territoire et de l'Environnement (MATE), Ministère de l'Equipement des Transports et du Logement (METL). La Documentation Française, Paris, 72p.

Pistocchi A., Luzi, L., Napolitano, P., 2002. The use of predictive modelling techniques for optimal exploitation of spatial databases: a case study in landslide hazard mapping with expert system-like methods. Environmental geology 41, 765-775.

Remondo, J., Gonzàlez-Diez, A., Dìaz de Teràn, J.R., Cendrero, A., 2003. Landslides susceptibility models utilising spatial data analysis techniques. A case study from the lower Deba Valley, Guipùzcoa (Spain). Natural Hazards 30, 267-279.

Saporta, G., 1990. Probabilités, analyse des données et statistique. Technip, Paris, 528p.

Süzen, M.L., Doyuran, V., 2004. Data driven bivariate landslide susceptibility assessment using geographical information systems: method and application to Asarsuyu catchment, Turkey. Engineering Geology 71(3-4), 303-321.

Thiart, C., Bonham-Carter, G.F., Agterberg, F.P., 2003. Conditional independence in weights of evidence: application of an improved test. IAMG, International Association for Mathematical Geology, September 7th-12th, 2003, Portsmouth, United-Kingdom.

Thiery, Y., Malet, J.-P., Maquaire, O. 2006. Test of Fuzzy Logic Rules for landslide susceptibility assessment. In Weber C. & Gancarski P. (Eds): *SAGEO 2006*, Proceedings International Conference on Spatial Analysis and Geomatics, Strasbourg, France, CD-Rom Support Proceedings, 18p.

Thiery, Y., Malet, J.-P., Maquaire, O. 2008. Application of expert rules in indirect approaches for landslide susceptibility assessment. Natural Haards and Earth System Sciences (submitted in December 2010, in review).

Thiery, Y., Malet, J.-P., Sterlacchini, S., Puissant, A., Maquaire, O. 2007. Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. *Geomorphology*, 9(1-2): 38-59.

Van Asch, T.W.J., Buma, J., 1997. Modelling groundwater fluctuations and the frequency of movement of a landslide in the 'Terres-Noires' region of Barcelonnette (France). Earth Surface Processes and Landforms 22, 131–141.

Van Westen C.J., Rengers N., Soeters R., 2003. Use of geomorphological information in indirect landslide susceptibility assessment. Natural Hazards 30, 399-419.

Van Westen, C.J., 1993. Application of geographic information systems to landslide hazard zonation. ITC Publication, vol. 15. International Institute for Aerospace and Earth Resources Survey, Enschede, The Netherlands. 245p.

6.3 CASE STUDY ON LANDSLIDE RISK QUANTIFICATION ALONG TRANSPORTATION CORRIDORS BASED ON HISTORICAL INFORMATION. NILGIRI, INDIA

By: Pankaj Jaiswal and Cees van Westen (ITC)

6.3.1 Abstract

The research was conducted along a transportation corridor with a road and a railway alignment in the Nilgiri hills in southern India. Historical records belonging to the railway and geotechnical units were used to obtain a multi-temporal landslide inventory for a 23-year period from 1987 to 2009. A substantially complete inventory was obtained for landslides initiating from cut slopes along the transportation lines. In contrast, only little information was available for landslides affecting natural slopes above the transportation lines. Most landslides were shallow translational debris slides and debris flowslides triggered by rainfall. On natural slopes most landslides occurred as first-time failures.

Various analyses were performed to quantify landslide risk along the road and the railroad and the surrounding areas. A Gumbel analysis was carried out to determine the frequency of landslides on cut slopes and on natural slopes for certain units of the transportation line. Logistic regression analysis was carried out to model the susceptible areas to landslides on natural slopes. Rainfall threshold analysis was used to estimate the temporal probability of landslides, and magnitude-frequency analysis to obtain the probability of landslide size. These data were combined in an analysis of landslide initiation hazard on cut and natural slopes. Landslide run-out analysis was carried for landslides on natural slopes. Landslide vulnerability was established for landslides with different magnitudes and for different elements at risk. Landslide hazard and risk estimation was done using landslide events that occurred between 1987 and 2007 and the results were validated using landslides that occurred in 2008 and 2009. As a final output direct risk was quantified for properties (alignments, vehicles, buildings, and plantations) and people (commuters and residents) and indirect risks due to the traffic interruption. The research also presented the perception of Nilgiri people to landslide risk and the use of the obtained hazard and risk information in reducing landslide risk to the society.

The results provided a quantitative estimate of total annual landslide losses, expressed in monetary value (US$) for properties and in annual probability of death for people. An F-N curve was used to express the societal risks. The results are of important societal value and will provide inputs for planning risk reduction strategies, for developing risk acceptance criteria and for financial analysis for possible damage in the study area. The methodology provides a cost-effective approach to estimate direct and indirect landslide risks. The methods can be applied elsewhere if a similar historical landslide data is made available.

6.3.2 Introduction

In the last decades many advances in landslide related studies have been made. Most of the studies are related to landslide identification, dynamic modelling and susceptibility analysis, but relatively few studies have been carried out on the quantitative analysis of landslide hazard and risk. The limited number of publications on landslide hazard and risk estimation is mostly due to the unavailability of a substantially complete landslide inventory and due to the lack of sufficient information on landslide damages, frequency of events and landslide sizes (van Westen et al., 2006). Since most of the above information can only be obtained from historical records the availability of such records, thus becomes crucial. In most countries there is no single agency that has the responsibility of maintaining historical landslides and damage records, rather different agencies maintain records of their area of interest. This result in incomplete and biased databases

both with respect to the area covered and to the time period investigated (Ibsen and Brunsden, 1996). The insufficient historical records on landslides, in fact, makes estimation of landslide risk difficult in most countries. More thrust is therefore needed to obtain inventory of historical landslides and associated damages so that assessment of landslide risk becomes possible.

The aim of this research was to extract landslide information from the available historical records and to convert them into a spatio-temporal landslide inventory, and to develop methods to apply this information for the quantification of landslide hazard and risk along transportation lines.

The specific objectives of the research are:

- To develop an approach to prepare a landslide inventory based on historical information,
- To develop methods of quantifying landslide hazard along transportation corridors for the purpose of risk analysis and land use planning, which include:
- a method for estimating (temporal) probability of occurrence of landslides using information on triggering events and landslide occurrence dates,
- a magnitude-frequency relationship of rainfall-induced landslides.
- To assess the vulnerability of movable and stationary elements at risk for landslide of different sizes,
- To develop methods to quantify direct and indirect risks along transportation corridors,
- To highlight the application of quantitative landslide hazard and risk information in planning risk reduction strategies

6.3.3 Study area

In this study the transportation corridor refers to a 22 km2 area encompassing a 24 km long section of a road and a 17 km long section of a railroad (Figure 6.3.1). The corridor forms a part of Coimbatore and Nilgiri districts of the western Tamilnadu region in southern India and falls within the Survey of India toposheet no 58 A/15. The road is a national highway (NH-67) and the railroad is declared by UNESCO as a “world heritage railway route”. Both form part of the main transportation lines connecting Mettupalayam to Coonoor in the state of Tamilnadu (Figure 6.3.1.).

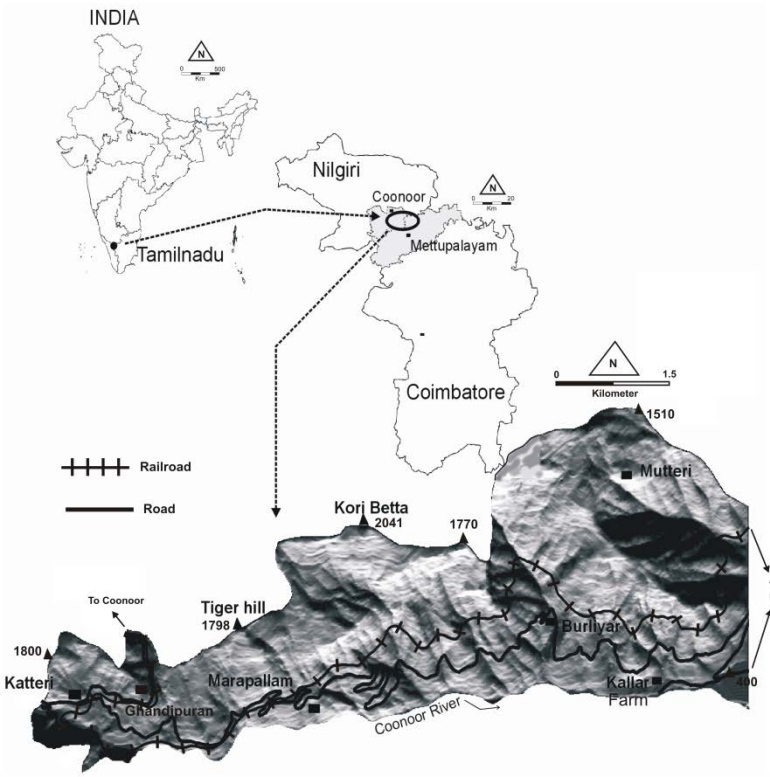


Figure 6.3.1 Location of the transportation corridor with a road and a railroad alignment. The hill shade map shows the extent of the study area

6.3.4 Landslide inventory mapping

In total 932 landslides were obtained from the historical records covering a 21-years period from 1987 to 2007 along the transportation corridor. Out of these 578 landslides (62%) were obtained from the railway slip register (from 1992 to 2007), 220 (24%) from the railroad landslide table (from 1987 to 1991) and 134 (14%) from technical reports (from 1987 to 2007). Figure 6.3.2 shows the spatial distribution of mapped landslides along the railroad in different years from 1992 to 2007. Due to the small size the location of landslide scars is represented as points. The minimum number of landslides recorded was in 1994 (5 landslides) and the maximum in 2006 (119 landslides). The figure indicates that the spatial distribution of landslides is not uniform every year except in 1992, 1996, 2001 and 2004 when landslides occurred all along the railroad. More landslides are located towards the eastern part of the study area, which is probably due to steep slopes.

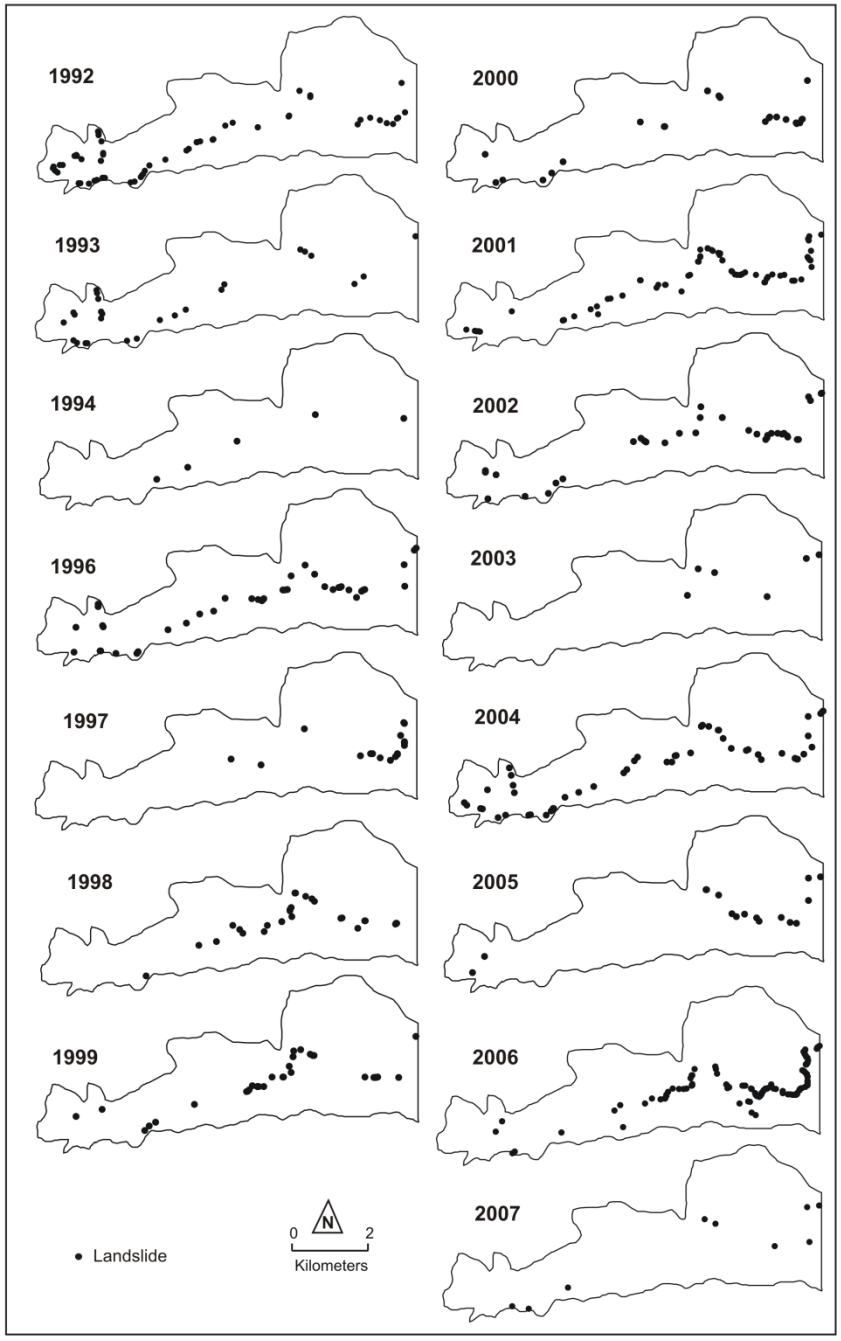


Figure 6.3.2 Spatial distribution of landslides in different years from 1992 to 2007 along the railroad. Points represent location of landslide scars on cut slopes. In 1995 no landslides occurred

For the period 1987 to 2007, a total of 111 landslide event dates were obtained from the historical records. The term “landslide event” is used hereafter for days when one or more landslides were triggered by rainfall. Out of 111 events, actual date was only known for 87 events. For 24 events the month of occurrence was available (taken from the railroad landslide table). It is assumed that in the latter case only one event had occurred each month. Table 6.3.1 provides the main characteristics of

all landslide events that have affected the area in the period from 1987 to 2007. The number of landslides per event varies from one to 166 and the number of events per year varies from one to 11.

Table 6.3.1Main landslide triggering events in the period 1987 to 2007

Event Year	Event date and number of landslides Date/month (#)	Total slide (#)	Landslide volume		
			Min. (m^3)	Max. (m^3)	Mean (m^3)
1987+	Oct (23), Nov (19); Dec (36); 14/12 (43) *	121@	50*	12,000*	706*
1988+	Sept (1); Nov (2); Dec (1)	04	--	--	--
1989+	Feb (1); March (1); Jul (1); Sept (3); Oct (1); Nov (22); Dec (2)	31	--	--	--
1990+	Jan (5); Feb (1); March (3); Oct (29); Nov (10)	48	--	--	--
1991+	Jan (6); April (1); May (3); July (1); Oct (2); Nov (46)	59	--	--	--
1992	14/11 (8); 15/11 (27); 16/11 (17); 20/11 (3); 21/11 (14)	69	6	900	77
1993	10/11 (6); 11/11 (27)	33	10	150,000	5308
1994	11/11 (5)	05	2	3600	883
1995	No event	--	--	--	--
1996	17/12 (41)	41	2	2000	120
1997	3/7 (1); 11/7 (1); 16/10 (1); 10/11 (1); 27/11 (16)	20	6	1944	133
1998	8 /1 (1); 18/8 (1); 27/9 (1); 9/10 (2); 13/10 (1); 9/11 (2); 2/12 (16); 11/12 (5); 12/12 (1)	30	3	240	54
1999	8/2 (1); 14/4 (1); 16/8 (1); 2/9 (1); 8/10 (1); 16/10 (23); 6/11 (1); 23/11 (2); 24/11 (1); 29/11 (2)	34	2	900	53
2000	25/9 (1); 14/11 (1); 15/11 (2); 22/11 (1); 24/11 (19); 1/12 (2); 31/12 (1)	27	2	600	49
2001	1/1 (1); 31/7 (1); 25/10 (1); 26/10 (3); 27/10 (22); 16/11 (4); 17/11 (26); 23/11 (3); 24/11 (1); 25/12 (13); 27/12 (7)	82	2	1200	53
2002	5/5 (1); 8/10 (19); 2/11 (2); 5/11 (12)	34	4	270	36
2003	20/3 (1); 30/4 (1); 13/5 (1); 29/5 (1); 19/8 (1); 10/11 (1)	06	2	77	17
2004	5/5 (21); 17/10 (1); 20/10 (3); 22/10 (4); 31/10 (14); 2/11 (3); 7/11 (2); 9/11 (6); 13/11 (2)	56	5	200	27
2005	7/10 (4); 22/10 (1); 7/11 (2); 13/11 (3); 25/11 (6)	16	2	100	22
2006	17/10 (4); 18/10 (9); 21/10 (2); 22/10 (5); 2/11 (8); 4/11 (2); 13/11 (1); 14/11 (166)	197	3	16,000	364
2007	8/4 (7); 28/5 (1); 27/10 (11)	19	2	420	66
	Total	932			

* data based on technical report pertaining landslide inventory along the road.

@ out of 121 landslides 78 slides are taken from the railway landslide table for which volume is not known.

+ Landslides data taken from the railway landslide table that only contains information on the number and month of occurrence of landslide (month (#)).

6.3.5 Rainfall threshold analysis

Daily rainfall data were collected from 14 rain gauges belonging to the tea estates (eight stations), the horticulture department (three stations) and the railway office (three stations).

The probability of landslide occurrences can be given by the intersection of two probabilities,

$$P[(R > R_T) \cap L] = P(R > R_T) \times P[L \mid (R > R_T)] \quad (6.3.1)$$

In which R is the daily rainfall, and R_T is the threshold rainfall causing a landslide. This means that the probability of occurrence of both ($R > R_T$) and (L) is equal to the probability of ($R > R_T$) multiplied by the probability of occurrence of (L), assuming that ($R > R_T$) has already occurred. The probability of ($R > R_T$) can be obtained by determining the exceedance probability of the rainfall threshold and the probability of $[L \mid (R > R_T)]$ relies on the frequency of occurrence of landslides after the threshold has been exceeded.

To determine the suitable number of antecedent days required for shallow debris slides and debris flowslides, 54 landslide events were selected that have occurred between 1992 and 2006 around Burliyar. These triggering events have resulted in 270 shallow landslides around this area. After analyzing the 3, 5, 15 and 30 days antecedent rainfall, according to the method suggested by Zezere et al. (2005), the 5-days antecedent rainfall was considered suitable for the analysis.

To determine R_T, a scatter plot was prepared showing daily rainfall against the corresponding 5-day antecedent rainfall, for each day with one or more triggered shallow landslides. The envelope curve was manually drawn such that it demarcates the lower end of the plotted points. The line can be represented by a linear mathematical equation.

For the calculation of the thresholds, the transportation corridor was divided into four sections (Figure 6.3.3), based on topography, land use types, and terrain gradient. Rainfall conditions at each section were determined from the nearest rain gauge.

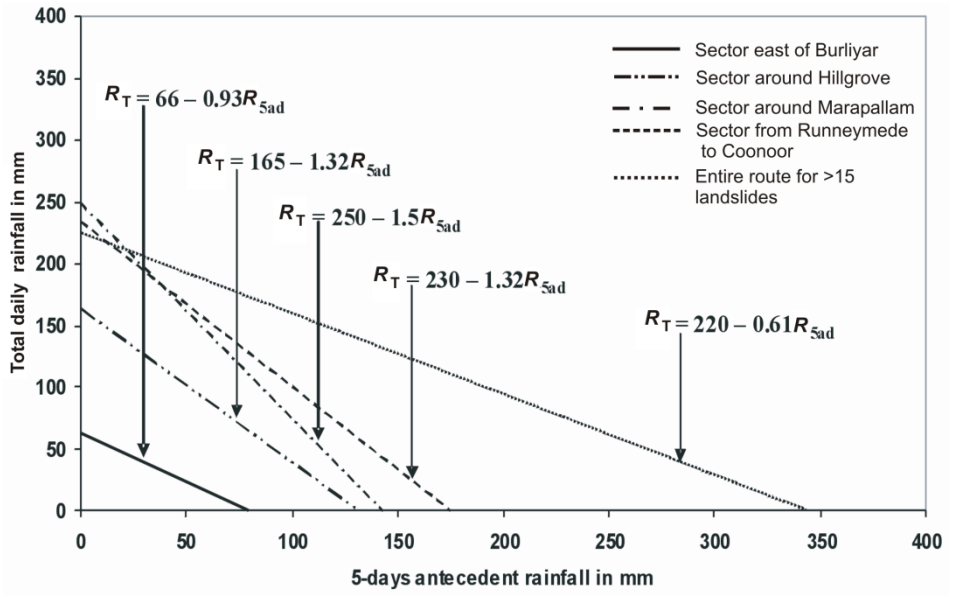


Figure 6.3.3 Envelope curves for landslides. R_T is the threshold rainfall and R_{5ad} is the 5-days antecedent rainfall

To determine *AEP* of the rainfall threshold for a particular area, R_T is calculated from the threshold equation, and the result is subtracted from *R*. Each phase of continuous positive values ($R > R_T$) is considered as the period of maximum likelihood for landslide initiation. For a given rain gauge AEP of the threshold P(R > RT) was determined using a Poisson probability model. This model has been used to determine the exceedance probability of landslides in time by, e.g., Coe et al. (2000, 2004) and Guzzetti et al. (2005). In this study, *AEP* calculation was based on the 15 years daily rainfall data from 1992 to 2006 in the months from October to December for landslide initiation along the railroad.

The next step after calculating *AEP* of the rainfall threshold is the assessment of the probability of landslide occurrence after the threshold has been exceeded. The frequency can be established from the rainfall and landslide records, for different sections of the railroad. From this frequency, the probability of (L) conditioned on ($R > R_T$), i.e., P[L| ($R > R_T$)], can be estimated. To achieve this, the transportation routes were further subdivided into eight smaller topographic units based on the variation in the land use type and the height of the cut slope (Figure 6.3.4). This was done to take account of variation in the landslide distribution in different units resulting from the unequal response of the terrain towards the threshold due to changes in local relief and land use. Other factors that play a role are differences in the height of the cut slopes and the size of the upslope area for a landslide to retrograde.

As indicated earlier, the temporal probability of landslide initiation was calculated by multiplying:

(i) *AEP* of the rainfall threshold, i.e., a probability of the threshold being exceeded in a year, by

(ii) the probability of landslide initiation given that the threshold is exceeded P[L| ($R > R_T$)].

The temporal validation of the threshold equation $R_T = 66 - 0.93\ R_{5ad}$ for the section east of Burliyar is shown in Figure 6.3.5. The validation was carried out using 2001, 2006 and 2007 rainfall, and landslide data. The 2001 and part of 2006 event data were also used in building the

model. These are included here to visualise the performance or success of the model. Figure 6.3.5 indicates that in the period from October to December the rainfall has exceeded the threshold curve several times. Between two successive positive periods (i.e., the period for which the threshold was exceeded) there may be a period with no rainfall or very low rainfall. Each rise in the threshold curve indicates that either there is a sudden increase in the magnitude of daily rainfall or there is a constant rise in five days antecedent rainfall. The width of each positive curve (or positive amplitude) denotes the period of consecutive rainy days in a given month. The crossover of the curve from negative to positive values indicates the time when the threshold is crossed and the conditions favourable for landsliding begins. One or more landslide events might be expected before the positive curve decays to the zero threshold value. Similar validations were carried out for the other threshold equations: $R_T = 165 – 1.32\ R_{5ad}$, $R_T = 250 – 1.5\ R_{5ad}$ and $R_T = 230 – 1.32\ R_{5ad}$ using the 2001, 2006 and 2007 rainfall and landslide data.

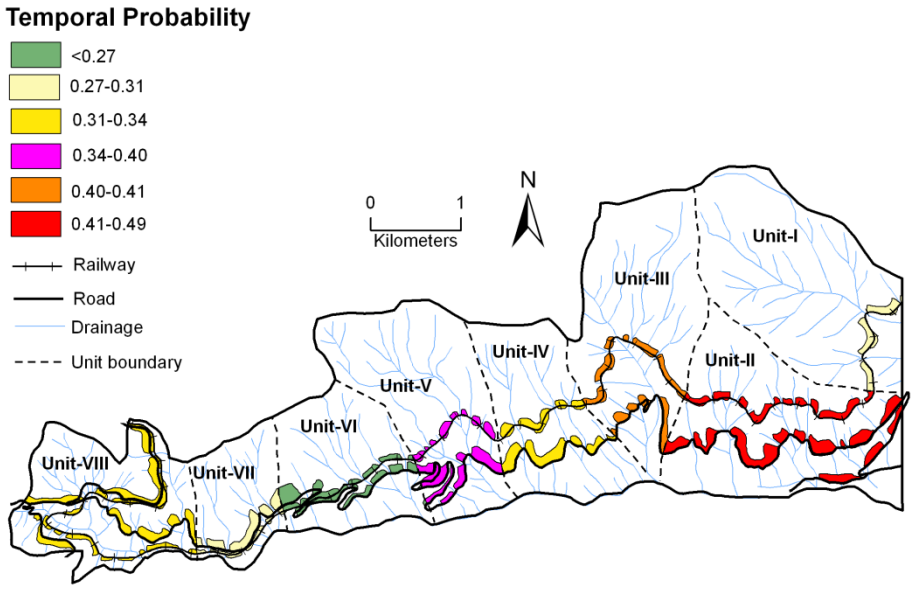


Figure 6.3.4 Annual temporal probability of landsliding on cut slopes. Temporal probability is based on the exceedance probability and frequency estimates of threshold rainfall in the different units along the railroad and the road

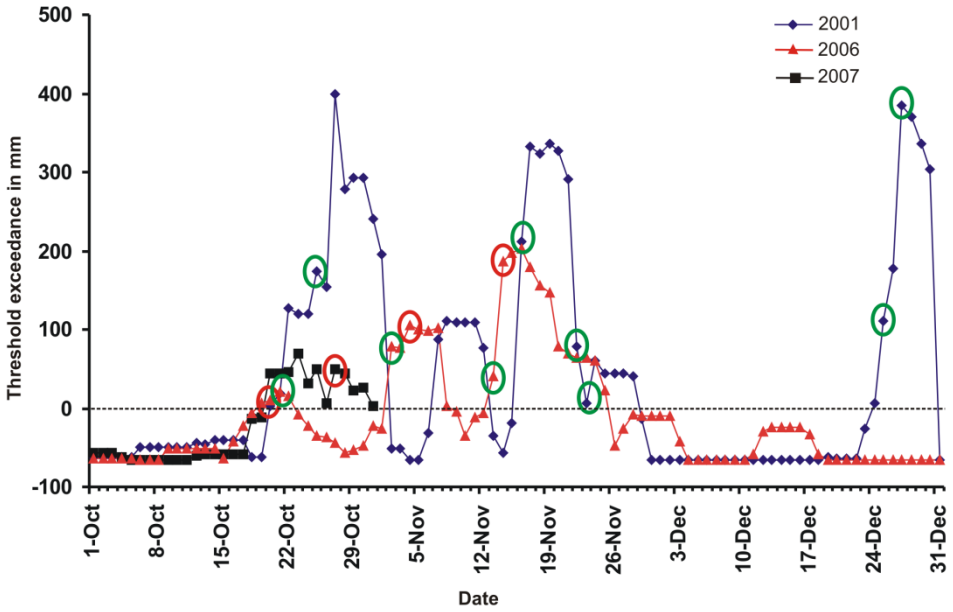


Figure 6.3.5 Validation of the threshold equation $R_T = 66 - 0.93\ R_{5ad}$ for the section east of Burliyar. Validation was done for the year 2001, 2006 and 2007. Positive values on the y-axis indicate threshold exceedance ($R > R_T$). Green circles indicate the dates of landslide events considered in the model. Red circles are the event dates that were not considered in building the threshold model

6.3.6 Landslide size analysis

To study the statistics of landslide sizes (area or volume) researchers have used the cumulative or the non-cumulative number-size distributions. Stark and Hovius (2001) suggested that a non-cumulative distribution is appropriate to observe any crossover from a non-power law to a power law scaling. In this analysis also a non-cumulative distribution was used for obtaining probability density of landslide volumes adopting the method given by Malamud et al. (2004). Non-cumulative probability density of landslide volumes was calculated for individual years for the period from 1992 to 2007. The results are shown in Figure 6.3.4.

Table 6.3.2 Probability of landslide size on cut slopes

Landslide triggering event	Probability of landslide of volume		
	$< 10^2$ m^3	$10^2 - 10^3$ m^3	$> 10^3$ m^3
<100 landslides per year	0.5 – 1 (avg = 0.85)	0.01 - 0.33 (avg = 0.13)	0 - 0.16 (avg = 0.02)
≥100 landslides per year	0.39	0.53	0.08

The frequency percentage of landslide size in different years was calculated, which was taken as the probability of occurrence of a particular landslide size on cut slopes (Table 6.3.2). The probability

was calculated for each year separately and the average value was taken for further analysis. The largest landslide recorded along the railroad is of volume ~3,600 m^3 and along the road it is ~5,250 m^3.

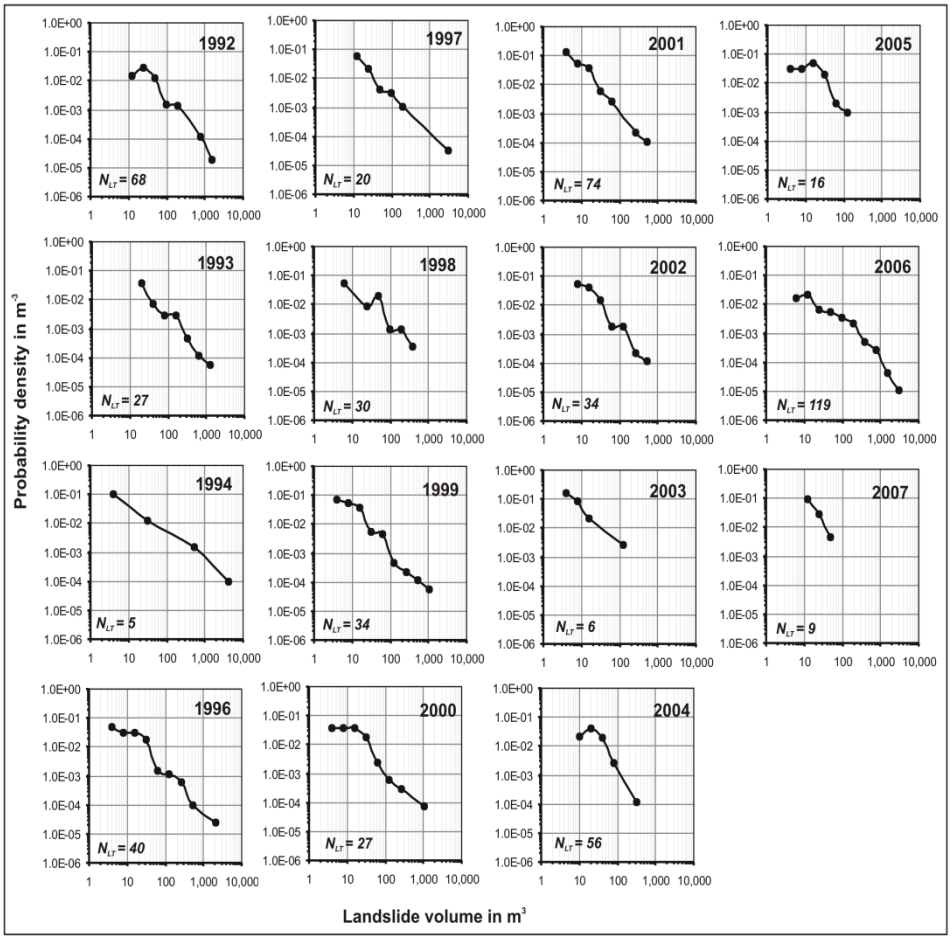


Figure 6.3.6 Probability density of landslide volumes for different years. N_{LT} is total number of landslides in the inventory

6.3.7 Landslide hazard assessment along the road

Figure 6.3.7 shows the work flow of the approach used to estimate landslide hazard along the road and railroad. For the hazard calculation, it is assumed that the probability of landslide occurrence can be calculated directly from the complete landslide inventory by analyzing the number of expected landslides per kilometer of the (rail) road for different return periods. For this the Gumbel method for frequency-magnitude analysis in which the magnitude is represented as the number of landslides per kilometer is used. The volume of expected landslides was analyzed separately using the volume-frequency analysis. Given the limitations in the available data, it is assumed that the volume estimation is independent of the number of landslides triggered by the event.

The hazard, expressed as the probability of a given number of landslides with a particular volume to occur along a specific section of the road or railroad, was obtained by multiplying the probability of having a certain number of landslides per unit length resulting from the Gumbel analysis, with the results of the volume-frequency analysis.

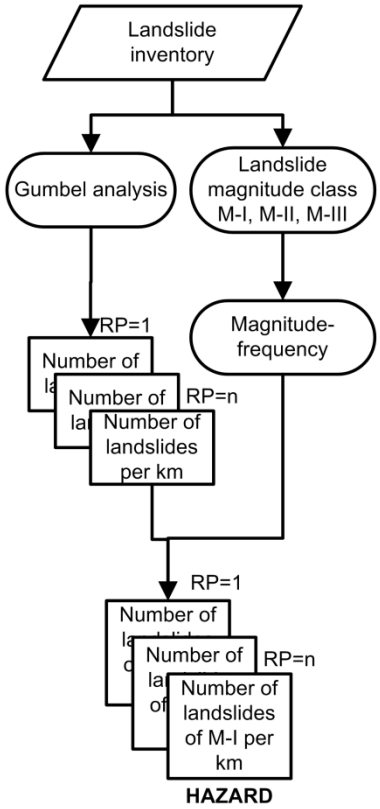


Figure 6.3.7 Flow diagram for the quantitative assessment of landslide hazard along the road and railroad

Figure 6.3.8 shows results of the Gumbel analysis for each kilometer length along the railroad represented as smooth line graph. Results for the two sections along the road are given in Table 6.3.3. Results indicate that no landslide is expected to occur along the railroad and the road on average once every year, and one or more landslides can occur on an average once in three or more years. A four kilometer stretch of the railroad (from km-10 to km-13) is relatively more prone to landslides, as is the S-I section (from km-390 to km-400) along the road. Total 56, 84, 140, 164 and 197 landslides are expected to occur along the entire railroad, and about 14, 28, 55, 66 and 82 landslides are expected along the entire road in T_3, T_5, T_{15}, T_{25} and T_{50} year return period, respectively.

The results indicate that the maximum number of landslides per kilometer is expected towards the east of Burliyar. This section of the transportation lines also has the maximum probability of experiencing rainfall events that can trigger one or more landslides (Figure 6.2.7).

Table 6.3.3 Number of landslides in different return periods (T_1 to T_{50}) along the road

Road section	Number of landslides per kilometer					
	T_1	T_3	T_5	T_{15}	T_{25}	T_{50}
S-I (total length 10 km)	0	1.01	2.04	4.08	4.93	6.12
S-II (total length 14 km)	0	0.29	0.53	1.00	1.20	1.48

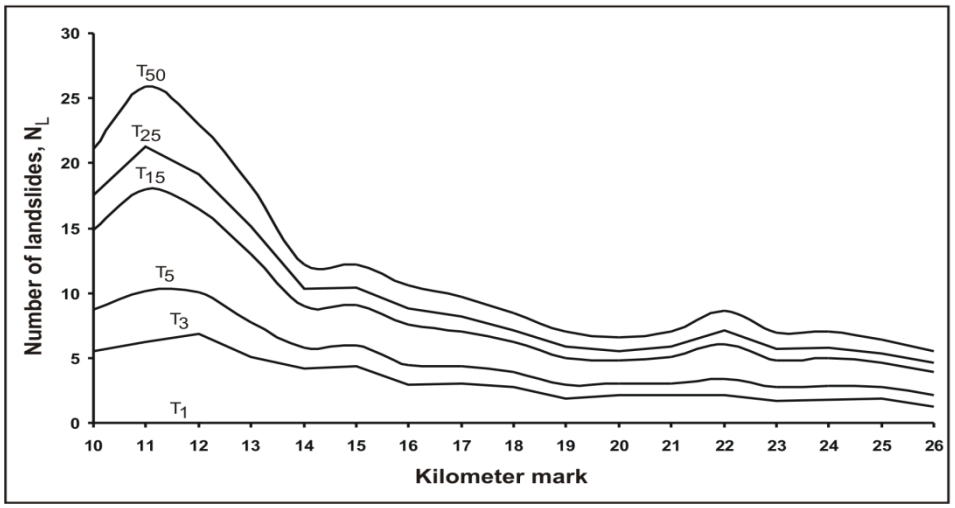


Figure 6.3.8 Total number of landslides in different return periods (T_1 to T_{50} year) along the railroad obtained using a Gumbel distribution

In total, 18 specific hazard scenarios were generated using combinations of the three volume classes (M-I, M-II and M-III) and six return periods (T_1, T_3, T_5, T_{15}, T_{25} and T_{50} year).

Two different approaches are presented for estimating landslide hazard for natural terrain above and below the transportation lines. The availability of information on the date of landslides and daily rainfall provided the opportunity to propose a probabilistic model based on a rainfall threshold to quantify hazard. The second approach is based on the statistical analysis of landslides on cut slopes to access landslide hazard for natural slopes in different return periods. The method uses a relationship between cut slope and natural slope failures to estimate the number and sizes of landslides expected on natural slopes.

Figure 6.3.9 shows the work flow of the rainfall threshold-based approach for landslide hazard estimation. It requires the estimation of three basic parameters:

(1) the magnitude probability (P_M): indicating the probability that the landslide might be of a given size,

(2) the temporal probability (P_T): indicating the annual probability of occurrence of triggering events that generate landslides, and

(3) the spatial probability (P_S): indicating the relative spatial probability of occurrence of landslides of a given type.

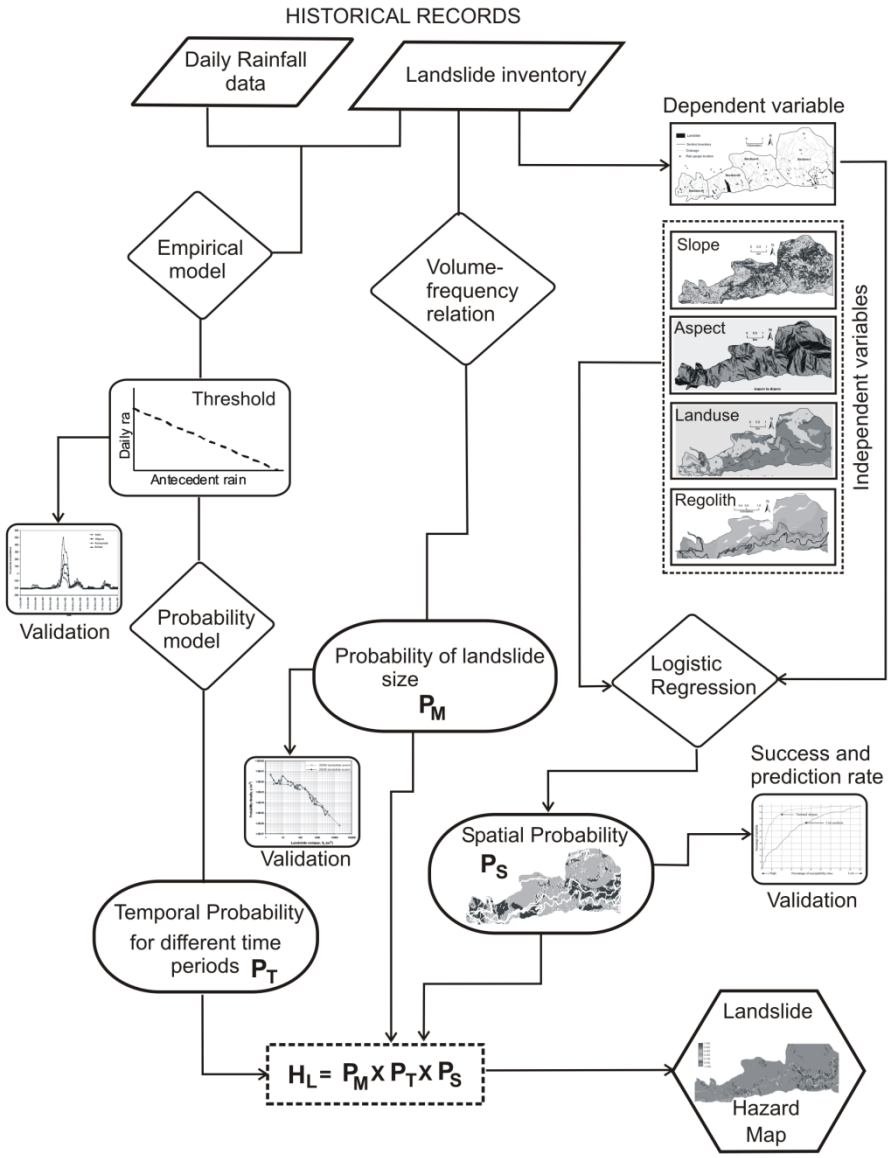


Figure 6.3.9 Parameters and process adopted for the rainfall threshold-based quantitative assessment of landslide hazard

Using the above probabilities, the landslide hazard on natural slopes can be estimated as the joint probability of the landslide size (volume), of the landslide occurrence in an established time period using rainfall threshold related to first-time slope failures and of the landslide spatial

occurrence given the local environmental setting, based on the method adapted from the earlier work presented by Guzzetti et al. (2005). It is assumed that the above three probabilities are mutually independent and the landslide hazard (H_L), i.e., the joint probability is:

$$H_L = P_M \times P_T \times P_S \quad (6.3.2)$$

Figure 6.3.10 shows the work flow of the inventory-based approach for landslide hazard estimation.

It requires information on two basic parameters:

(1) the landslide susceptible zones: indicating zones with relatively high, moderate and low susceptibility, and

(2) the number and sizes of landslides expected on natural slopes in different return periods, obtained based on a relationship between the number of landslides on cut slopes and the frequency distribution of cut and natural slope failures.

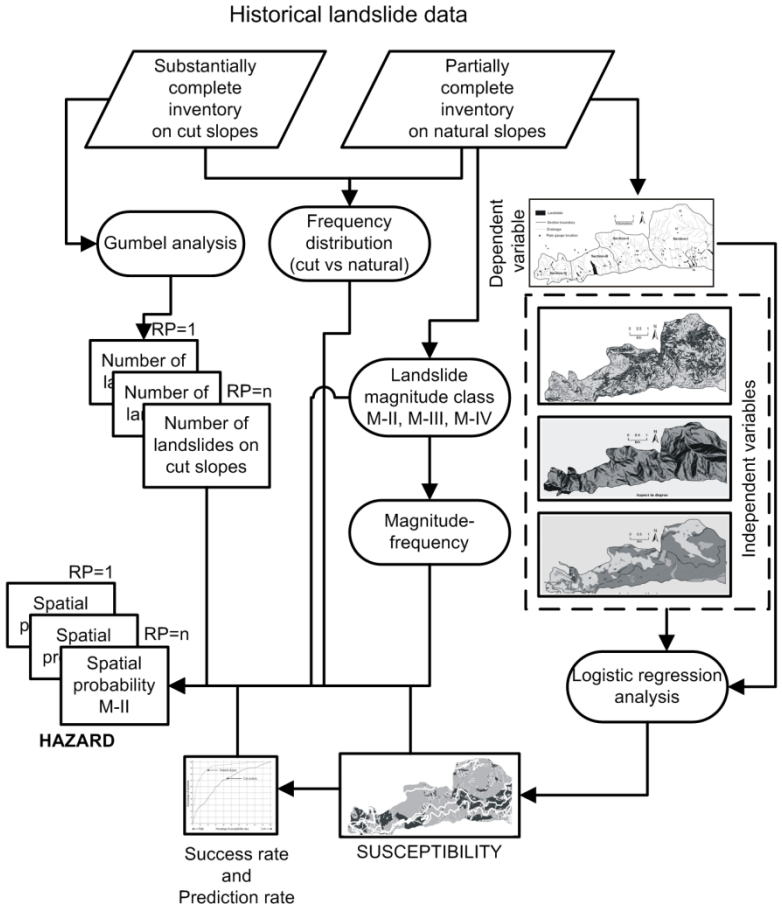


Figure 6.3.10 Parameters and process adopted for the inventory-based landslide hazard assessment

To convert the susceptibility map into a quantitative hazard map i.e., the probability of occurrence of a landslide of a given size in a given time period, the following information is required:

1) It requires information on the number of landslides on natural slopes. The inventory indicates that landslides occur on natural slopes in triggering events that cause many cut slope failures (e.g., in 2006 and 2009). Based on the data of the 2006 and 2009 events, for which a complete inventory was available for both cut slopes and natural slopes, it is estimated that during such a landslide triggering event about 90% of all landslides will occur on cut slopes and 10% on natural slopes. It is assumed that the number of landslides expected to occur on natural slopes in different return periods can be obtained indirectly if a relationship between the number of landslides on cut slopes can be established for different return periods. This relationship was established using a Gumbel model and the landslide inventory on cut slopes.

2) It requires information on the size of landslides on natural slopes. The landslides were grouped into different magnitude classes with minimum to maximum landslide volumes ranging from 100 to 1,000 m^3 (average=550 m^3) for M-II, from 1,000 to 10,000 m^3 (average=5,500 m^3) for M-III and from 10,000 to 100,000 m^3 (average=55,000 m^3) for M-IV. Based on the available inventories the percentage distribution of natural slope failures in M-II, M-III and M-IV classes was 66%, 28% and 6%, respectively. These values were used to estimate the relative frequency of landslide size in different return periods.

3) It requires information on landslides in different susceptible zones. The susceptibility map was grouped into three susceptible zones, as high (estimated probability > 0.6), moderate (probability 0.4-0.6) and low (probability ≤ 0.40). For each return period landslides were to be distributed in the three susceptible zones based on the result of the model fitting performance (success rate curve). Finally, hazard or the probability of occurrence of a landslide of a given size per pixel (spatial probability) for each return period was estimated as the ratio of the total area of landslides of a given size expected in each susceptible zone to the total area of the susceptible zone. For each magnitude class source area of a landslide was obtained by dividing the volume for the given magnitude class with the average depth of the landslide at the scarp. For the magnitude class M-II, M-III and M-IV the average depth of a landslide was taken as 2, 5 and 10 m, respectively. Thus different hazard scenarios can be obtained considering a minimum, average and maximum landslide size of each magnitude class.

6.3.8 Landslide risk assessment

The approach used to estimate landslide risk along transportation corridors is presented schematically in Figure 6.3.11. Both direct and indirect landslide risk was analyzed. Direct risk was analysed separately for elements at risk vulnerable to landslides on cut slopes and natural slopes. Indirect risk was estimated only for cut slope failures. To assess direct risk the following activities are required (AGS, 2000; Fell et al., 2008):

(1) estimation of landslide hazard for a specific return period,

(2) estimation of run-out distances for potential landslides,

(3) mapping and quantification of the elements at risk, number of people and properties (monetary value), and assessment of their temporal and spatial probability to be in an exposed position. Vulnerability values were based on expert opinion (See Tables 6.3.4, 6.3.5 and 6.3.6)

(4) assessment of probability of the landslide reaching the elements at risk,

(5) assessment of the vulnerability of the elements at risk, in terms of property damage (monetary loss) or loss of life, resulting from the specified landslide magnitudes, and

(6) estimation of specific risk for each element at risk for various landslide hazard.

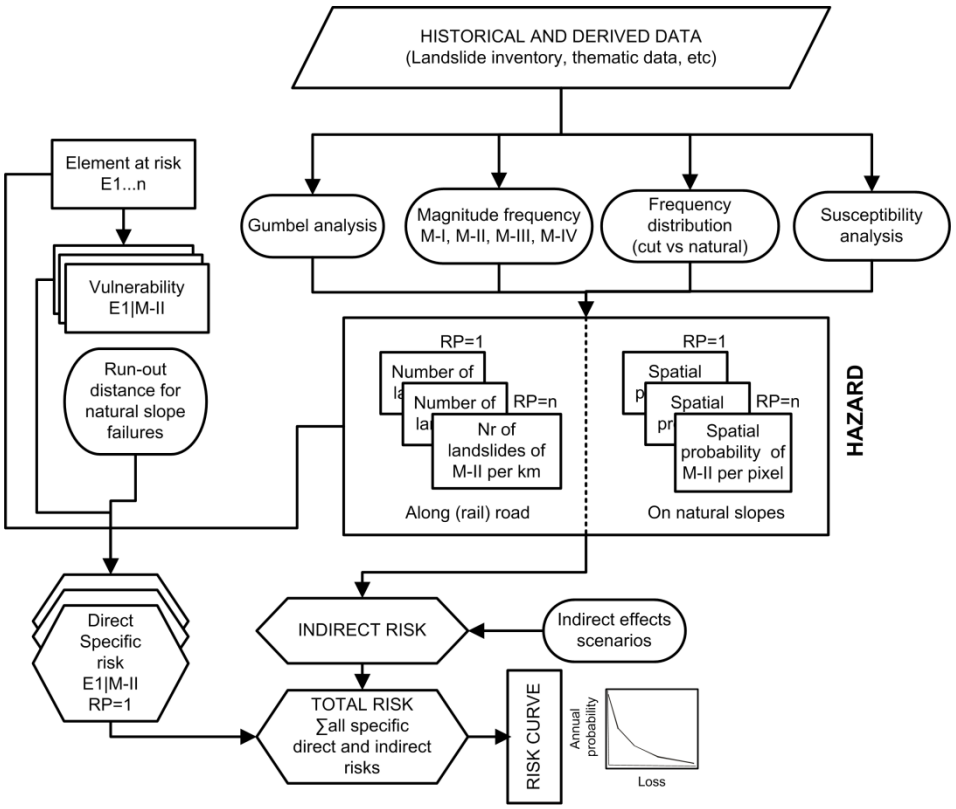


Figure 6.3.11 Flow diagram showing the process adopted for landslide risk analysis

The analysis of indirect risk depends on the socio-economic condition of the area of interest. It requires determination of the most important elements and activities in the area and how they could be affected due to the disruption (Remondo et al., 2008). The elements that are indirectly affected due to the traffic disruption include local businesses, residents, tourists, and transport and railway department.

Table 6.3.4 Estimated vulnerability of elements at risk located in landslide initiation areas

Element at Risk	Vulnerability due to a landslide of magnitude class								
	M-II			M-III			M-IV		
	min	avg	max	min	avg	max	min	avg	max
Tea/coffee plantation	0.5	1	1	1	1	1	1	1	1
Horticulture plantation	0.5	0.8	1	1	1	1	1	1	1
Railroad	0.5	0.8	1	1	1	1	1	1	1
Road	0.3	0.6	1	1	1	1	1	1	1
Building types									
Type-1	0.3	0.5	1	1	1	1	1	1	1
Type-2	0.2	0.4	0.8	0.8	1	1	1	1	1
Type-3	0.1	0.3	0.5	0.5	0.6	1	1	1	1
Type-4	0	0.1	0.2	0.2	0.8	1	1	1	1

Table 6.3.5 Estimated vulnerability of elements at risk located within run-out paths of a natural slope failure

Element at Risk	Vulnerability due to a landslide of magnitude class								
	M-II			M-III			M-IV		
	min	avg	max	min	avg	max	min	avg	max
Railroad	1	1	1	1	1	1	1	1	1
Road	0.2	0.3	0.4	0.4	0.6	0.8	0.8	1	1
Building types									
Type-1	1	1	1	1	1	1	1	1	1
Type-2	0.2	0.4	0.8	0.8	1	1	1	1	1
Type-3	0.1	0.3	0.5	0.5	0.7	0.8	0.8	1	1
Type-4	0	0.05	0.1	0.1	0.3	0.4	0.4	0.8	1

Table 6.3.6 Estimated vulnerability of elements at risk affected by a landslide from cut slopes

Element at Risk	Vulnerability due to a landslide of magnitude class		
	M-I max	M-II max	M-III max
Infrastructure			
Railroad	0.5	1	1
Road	0.2	0.4	0.8
Moving vehicle			
Bus	0.01	0.1	0.8
Lorry	0.01	0.1	0.8
Car	0.1	0.5	1
Motorbike	0.5	0.8	1
Train	1	1	1
Person in a moving vehicle (probability of death)			
Bus	0.001	0.1	0.8
Lorry	0.001	0.1	0.8
Car	0.01	0.1	1
Motorbike	0.5	1	1
Train	0.5	0.5	0.5

Direct risk to properties (i.e., infrastructure components) for landslides from cut slopes was estimated for a given return period using the expression adapted from Fell et al. (2005):

$$RD_{p:cut} = \sum_{m=1}^{m=n} (H_{cut:m} \times P_{Lm:p} \times P_{T:p} \times V_{p:Lm} \times A_p) \quad (6.3.3)$$

where, $RD_{p:cut}$ is the direct risk to properties (US$), $H_{cut:m}$ is the hazard due to landslides of magnitude class '*m*' (nr per km), $P_{Lm:p}$ is the probability of a landslide with magnitude '*m*' reaching the infrastructure (0-1), $P_{T:p}$ is the temporal probability of the infrastructural components to be exposed to a landslide of magnitude '*m*' (0-1), $V_{p:Lm}$ is the vulnerability of the property i.e., infrastructural components caused due to the occurrence of a landslide of magnitude '*m*' (0-1), and A_p is the quantification (monetary value) of the property (US$). The specific risk is calculated per standard length of the road or railroad (e.g., per kilometer). The specific risk for different landslide magnitudes for each return period is integrated to generate the combined specific risk for a particular infrastructure element.

Direct risk to properties located in the initiation areas of potential landslides on natural slopes was estimated for a given return period using the expression adapted from Fell et al. (2005):

$$RD_{p:Li} = \sum_{m=1}^{m=n} (H_{nat:m} \times V_{p:Lm} \times A_p) \quad (6.3.4)$$

where, $RD_{p:Li}$ is the direct risk to the property located in a landslide initiation area (US$), $H_{nat:m}$ is the probability of occurrence of a landslide of size '*m*' (0-1), $V_{p:Lm}$ and A_p are explained in Eq. (7.3.2). For each return period the specific risk was estimated considering the minimum, average and maximum landslide volumes of each magnitude class and their corresponding minimum, average and maximum vulnerability values. The specific risk for different landslide magnitudes was added for each return period to generate the combined specific risk resulting in minimum, average and maximum landslide risk estimated for four time periods.

Direct risk from landslides on natural slopes located above properties for a given return period was estimated using the expression adapted from AGS (2000):

$$RD_{p:Lr} = \sum_{m=1}^{m=n} (H_{nat:m} \times P_{Lm:p} \times V_{p:Lm} \times A_p) \quad (6.3.5)$$

where, $RD_{p:Lr}$ is the direct risk to the property located within the run-out path of a landslide (US$), $P_{Lm:p}$ is the probability of a landslide with size '*m*' reaching the element at risk from the upslope areas (0-1), and $V_{p:Lm}$ is the vulnerability of the element at risk due to a landslide run-out caused by a landslide of size '*m*' (0-1). $H_{nat:m}$, and A_p are the components used in Eq. 6.3.2). The specific risk for different landslide magnitudes is added for each return period to generate the combined specific risk for the particular element.

Direct risk to a moving vehicle, i.e., a vehicle being hit by a landslide, depends on the probability ($P_{T:veh}$) of the vehicle being at the location of a landslide when it occurs. This probability ($P_{T:veh}$) was used to calculate the specific risk to a moving vehicle for a given return period using the following three expressions (adapted from AGS, 2000):

$$RD_{sv} = P(V_m) \times V_{veh:m} \times A_{veh} \quad (6.3.6)$$

$$P(V_m) = 1 - (1 - P_{T:veh})^{Nr} \quad (6.3.7)$$

$$P_{T:veh} = (ADT \times L) / (24 \times 1000 \times S_{veh}) \quad (6.3.8)$$

where, RD_{sv} is the direct specific risk to vehicles (US$), $P(V_m)$ is the probability of one or more vehicles being hit by a landslide initiating from a cut slope with a magnitude '*m*' (0-1), $V_{veh:m}$ is the vulnerability of the vehicle for a landslide of magnitude '*m*' (0-1), A_{veh} is the cost of the vehicle (US$), $P_{T:veh}$ is the probability of a vehicle occupying the portion of the road onto which the landslide hits with a given magnitude 'm' (0-1), N_r is the number of landslides of magnitude '*m*'

initiating from cut slopes, *ADT* is the average daily traffic (vehicles per day), *L* is the average length of the vehicle (m) and S_{veh} is the speed of the vehicle (km/hr).

The risk of life or the annual probability of a person losing his/her life while travelling in a vehicle depends on the probability of the vehicle being hit by a landslide $P(V_m)$ and the probability of death of the person (vulnerability) given the landslide impact on the vehicle.

The specific risk to commuters for a given return period was estimated using the following expression (adapted from AGS, 2000):

$$RD_c = P(V_m) \times V_{c:m} \quad (6.3.9)$$

where, RD_c is the annual probability of death (0-1), $V_{c:m}$ is the vulnerability of the individual commuter (probability of death) given the landslide with magnitude '*m*' from a cut slope impact on the vehicle (0-1). The parameter $P(V_m)$ is estimated using Eqs. (6.3.6 – 6.3.7).

Direct specific risk to persons occupying the building affected by a landslide from an upslope area was estimated using the expression adapted from AGS (2000):

$$RD_{pop} = H_{nat:m} \times P_{Lm:bt} \times P_{pop:bt} \times V_{pop:Lmbt} \quad (6.3.10)$$

where, RD_{pop} is the annual probability that a person will be killed (0-1), $H_{nat:m}$ is the annual probability of occurrence of a landslide of size '*m*' (0-1), $P_{Lm:bt}$ is the probability of a landslide with size '*m*' reaching the building of type '*t*' from the upslope areas (0-1), $P_{pop:bt}$ is the probability that the person is present in the building of type '*t*' affected by the hazard at the time of its occurrence (0-1), $V_{pop:Lmbt}$ is the vulnerability of the person given a landslide of size '*m*' impacting the building of type '*t*' (0-1).

The indirect risk estimation requires two basic parameters: the hazard scenario that defines the blockage time of the transportation lines, and a socio-economic analysis of the study area to determine the most important activities in the area and their consequences to the society if disrupted.

Indirect combined risk for additional fuel consumption for a given return period was calculated using the following expression:

$$RI_{FC} = \sum_{m=1}^{m=n} \left[(ARL \times ADT \times FC \times TBT_m) / M_V \right] \quad (6.3.11)$$

where, RI_{FC} is the indirect combined risk (monetary loss) due to additional fuel consumption by vehicles (US$), *ARL* is the alternate road length (km), *ADT* is the average daily traffic (vehicles per

day), *FC* is the fuel cost (US$/l), TBT_m is the traffic blockage time due to landslides with magnitude '*m*' (day) and M_V is the mileage of the vehicle (km/l).

Indirect combined risk for additional travel cost for a given return period was calculated using the following expression:

$$RI_{TC} = \sum_{m=1}^{m=n} ADC \times CT \times TBT_m \quad (6.3.12)$$

where, RI_{TC} is the indirect combined risk (monetary loss) due to additional travel cost (US$), *ADC* is the average commuters per day, *CT* is the cost of ticket (US$) and TBT_m is the traffic blockage time due to landslides with magnitude '*m*' (days).

Indirect combined risk for business for a given return period was calculated using the following expression:

$$RI_B = \sum_{m=1}^{m=n} NBT \times ADI \times P_{Loss} \times TBT_m \quad (6.3.13)$$

where, RI_B is the indirect combined risk (monetary loss) to business (US$), *NBT* is the number of businesses, *ADI* is the average daily income from the business (US$/day), P_{Loss} is the probability of loss in income (0-1) and TBT_m is the traffic blockage time due to landslides with magnitude '*m*' (days).

Indirect combined risk to the railway department in a given return period was calculated using the following expression:

$$RI_R = \sum_{m=1}^{m=n} DIL \times TBT_m \quad (6.3.14)$$

where, RI_R is the indirect combined risk (monetary loss) to the railway department (US$), *DIL* is the daily income loss (US$/day) and TBT_m is the traffic blockage time due to landslides with magnitude '*m*' (days).

Figure 6.3.12 shows the estimated risks to different properties located in the initiation areas of potential landslides, as obtained using Equation (6.3.2).

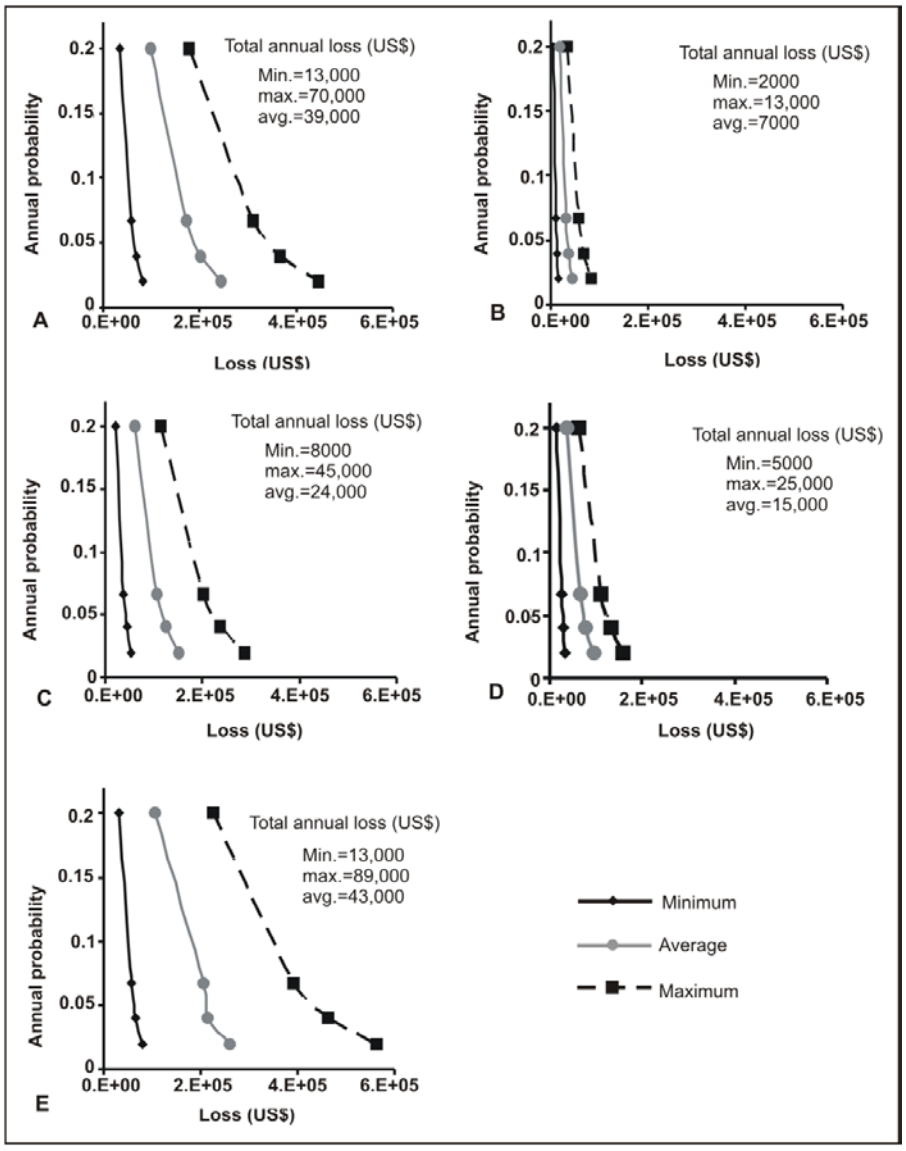


Figure 6.3.12 Specific risk to landslide initiation displayed as risk curves. A: Tea/coffee plantations; B: Horticulture plantations; C: Railroad; D: Road and E: Buildings

Figure 6.3.13 shows an example of the total risk within a 10-year period. Figure 6.3.14 displays the risk curve of the total maximum landslide risk, which is the plot of total maximum losses (US$) versus annual probabilities of the occurrence of triggering events. The total annualized maximum total loss in the study area, which is area under the curve, is estimated as US$ 432,000.

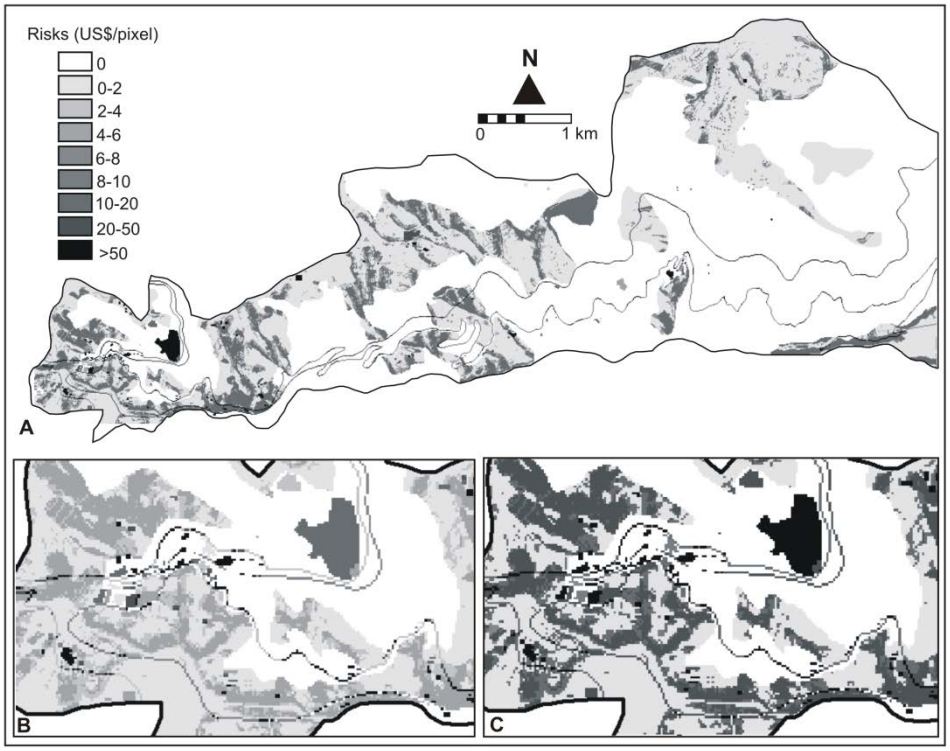


Figure 6.3.13 Specific risk to landslide initiation displayed as risk curves. A: Tea/coffee plantations; B: Horticulture plantations; C: Railroad; D: Road and E: Buildings. Expected monetary losses due to landslide initiations in US$ per pixel over a 10 year period. A: shows distribution of the average loss in the study area. B and C: show minimum and maximum loss, respectively around Katteri

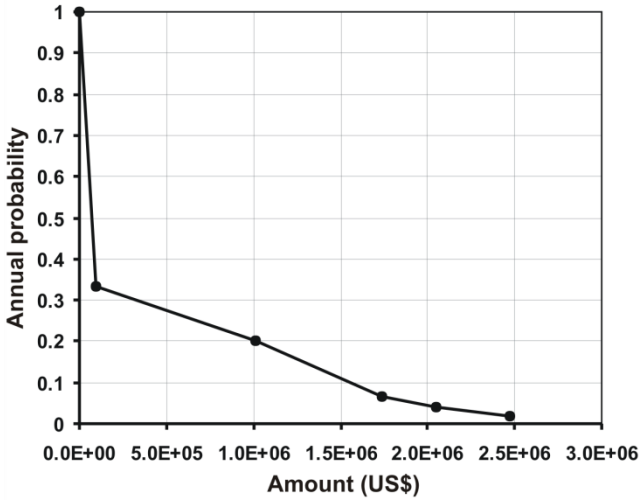


Figure 6.3.14 Risk curve for total landslide losses in the study area (maximum risk) expressed in monetary value (US$)

6.3.9 Validation of hazard and risk models

According to the guidelines of the Joint Technical Committee on landslides and Engineered Slopes (Fell et al., 2008), all models used in the hazard and risk analysis need to be validated for their performance in forecasting landslides. Validation of results of a predictive modelling is absolutely essential in order to make the model applicable for practical purposes.

Validation can be performed using a variety of datasets, but for practical purposes, validation can best be performed using landslide events independent from the ones used in modelling landslide hazard and risk. What we require is the spatio-temporal information on landslides and resulted consequences that have affected the area after the period considered in constructing the models.

Researchers commonly follow four different methods to obtain an independent landslide dataset for validation purposes: (i) they split the landslide inventory into two sets, one set used as calibration dataset for modelling and other set as validation dataset; (ii) modelling is carried out in one part of the study area and the model is extrapolated to the other part of the area where it can be validated using landslides belonging to the validation area; (iii) modelling is carried out using landslides that occurred in a certain time period and validation is performed using landslides of a different period; and (iv) validation is carried out using landslides that occurred after the period considered in the modelling. The latter is most adequate to test the validity of the "prediction" made, but as a disadvantage it requires a landslide event to occur.

In this section a temporal validation (of type iv, see above) is made of the models presented in the previous sections. This is based on a number of landslide events from November 2009. In the period between January 2008 and December 2009, a total of nine landslide events have affected the Nilgiri district, out of which the landslide event that occurred on 8, 9 and 10 November 2009 is the severest in terms of the number of landslides triggered and damage caused. On these three days continuous heavy rainfall triggered numerous landslides within the study area around Hillgrove and Katteri, and adjacent to the study area around Wellington, Ketty and Lovedale. The three days cumulative rainfall was recorded as 202 mm at Kallar farm, 387 mm at Burliyar, 560 mm at Hillgrove, 865 mm at Katteri farm, 937 mm at Coonoor and 1357 mm at Ketty rain gauge. The amount of rainfall shows a dramatic decrease from West (Coonoor) to East (Kallar farm) of the study area. The heavy precipitation in a short period resulted in flash floods and torrent streams, which triggered landslides along the channels and the accumulated debris flowed down slope causing damages to properties located close to stream courses.

For these events a landslide inventory was prepared from the railway slip register and landslide technical reports. Field mapping was carried out in February 2010. A total of 147 landslides were identified and spatially mapped on a 10,000 scale topomap. Figure 6.3.15 shows the spatial distribution of landslides triggered in November 2009. Out of the 147 slides, 71 occurred on cut slopes along the railroad and 65 along the road, and 11 on the natural slopes around Katteri and Hillgrove area. Field evidences indicate that most landslides on natural slopes were caused due to erosion (over bank or head ward) by torrent streams. A total of 80 landslides on cut slopes were reactivated old slides whereas 67 landslides, including all landslides on natural slopes, occurred in locations without clear evidence of previous landslides. Prior to November 2009, seven more landslides have occurred along the railroad, which makes a total of 154 landslides in 2009.

Within the study area the total physical damage caused by the 2009 landslide events include five houses, two shops, two cottages of a resort, one workshop, one tourist restroom, four parked cars, two parked trucks, one swimming pool. The events caused seven human casualties. Besides, the NH-67 road was closed for 57 days, and the railroad from Coonoor to Kallar was closed for 157 days, thus resulting in both direct and indirect losses.

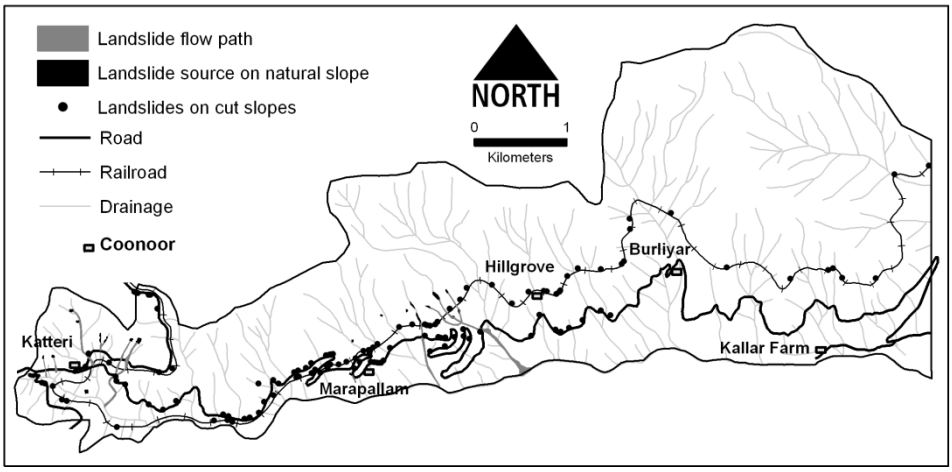


Figure 6.3.15 Spatial distributions of landslides caused by the 2009 events

6.3.9.1 Validation of rainfall threshold models

In the previous sections a method was presented to generate landslide risk information for a transportation corridor. We have used landslide events that occurred after 2007 for validating the model performance for future forecasting of landslides. For the validation of the rainfall thresholds we used a method that evaluates how often the rainfall threshold was crossed and whether this was associated with the occurrence of a landslide event. Out of the nine landslide events recorded in 2008 and 2009, five have occurred in the period between October and December 2009 due to the retreating monsoon. These five landslide events were used to validate the threshold models on cut slopes (Figure 6.1.3) and the model for natural slopes (R_T = 210– 0.54 R_{5ad}). The results are shown in Table 6.3.8. The probability of occurrence of landslides given the exceedance of the thresholds that can trigger less than 15 landslides on cut slopes within units I to VIII, P[L| ($R > R_T$)], varies from 0.20 to 1. In terms of percentage, the forecasting accuracy for most models for small landslide events (< 15 slides) remains less than or equal to 50%. The performance of the model is better for rainfall events that can trigger landslides on natural slopes and 15 to 90 failures on cut slopes (threshold for the entire route). The threshold model (R_T = 220 – 0.61 R_{5ad}) for the entire route and model (R_T = 210 – 0.54 R_{5ad}) for natural slopes give probability of P[L| ($R > R_T$)] equal to 1, which indicates that the threshold model is capable of accurately forecasting landslides of large events. The low performance of the models in units I to V is because of small threshold limits which are attributed to even a single slope failure on cut slopes.

Figure 6.3.16 shows the performance of the threshold model for natural slopes during October to December 2009. In Burliyar area, landslides occurred on 10 November when the rainfall crosses the threshold value, whereas in Hillgrove and Runneymede landslides also occurred on 8 and 9

November when the threshold was high. In the Kallar area, rainfall did not cross the threshold value and also no landslide occurred in this part of the study area. Similarly in 2008, the threshold was not exceeded in any of the rain gauges and therefore no landslide occurred in this year.

The probability of occurrence of one or more landslide events on natural slopes in a 3-year time period was estimated high (~0.65) in sections I and III and relatively low (~0.33) in sections II and IV. However, in 2009 rainfall triggered landslides in all sections, except in section I. An amount of rainfall similar to the 2009 events has never been recorded during the period of analysis (1987 to 2007). In the recent past, a very high rainfall occurred in 1979 that resulted in floods and landslides around Coonoor area.

Table 6.3.7 Validation of exceedance of rainfall threshold and occurrence of landslide events during 2008 and 2009 (RR = Railroad, NS = Natural slopes)

Units*	Threshold equation	Number of times the threshold exceeded	Landslide frequency in units	P[(L|$R>R_T$)]
RR I	$R_T = 66 - 0.93 R_{5ad}$	5	1	0.20
RR II		5	2	0.40
RR III		5	1	0.20
RR IV	$R_T = 165 - 1.32 R_{5ad}$	4	1	0.25
RR V		4	1	0.25
RR VI	$R_T = 230 - 1.32 R_{5ad}$	2	1	0.50
RR VII	$R_T = 250 - 1.5 R_{5ad}$	2	1	0.50
RR VIII		2	2	1
Entire railroad	$R_T = 220 - 0.61 R_{5ad}$	1	1	1
NS Kallar	$R_T = 210 - 0.54 R_{5ad}$	0	0	0
NS West of Burliyar		1	1	1

* for units along the railroad refer Figure 6.3.3.

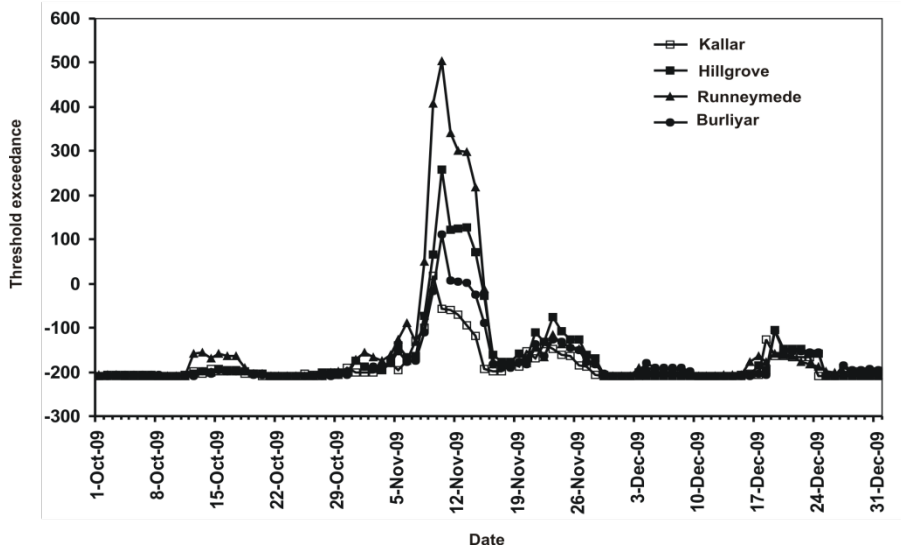


Figure 6.3.16 Validation of the threshold equation $R_T = 210 - 0.54\ R_{5ad}$ for natural slopes. Positive values on the y-axis indicate threshold exceedance ($R > R_T$)

6.3.9.2 Validation of landslide size model

In this section, the results of the probability of landslide size are compared with the result of 2009 events. Figure 6.3.17 shows the non-cumulative probability density distribution of landslide volumes of the 2006 and 2009 events. In 2009, the probability density distribution shows a distinct change in slope of the curve for failure volumes less than 200 m^3 whereas in 2006 a roll-over in form of flattening of curve was observed for failure volumes less than 200 m^3. However, in both distributions the linear portion of the curve (volume > 200 m^3) are comparable and shows a negative power law fit with $\beta = 1.7$ in 2009 and $\beta = 1.9$ in 2006.

The probability of a landslide volume exceeding 1,000 m^3 is estimated as 0.04 in 2009, which is smaller than that observed in 2006 (0.07). In fact in 2009, small landslides (volume < 100 m^3) occurred in relatively large number. This is because in 2009 most landslides occurred within section III and IV, which have gentle slopes covered by tea plantation. A similar case was observed in 1993 where the gentler terrain of section IV triggered more small landslides and resulted in a negative power law distribution of landslide volumes.

For cut slopes, the probability of occurrence of landslides in 2009 was found to be 0.61 for < 100 m^3, 0.35 for 100 to 1,000 m^3 and 0.04 for > 1,000 m^3, which is significantly different from the values used in the hazard model on cut slopes for years with more than 100 landslides i.e., 0.39 for < 100 m^3, 0.53 for 100 to 1,000 m^3 and 0.08 for > 1,000 m^3. As explained above the greater percentage of small slope failures is because of the gentle terrain condition. The percentage of landslides of larger volume would have been more if the same amount of rainfall that triggered landslides between the Marapallam and Katteri section would have affected the area east of Burliyar where slopes are steeper and the height of cut slopes is more.

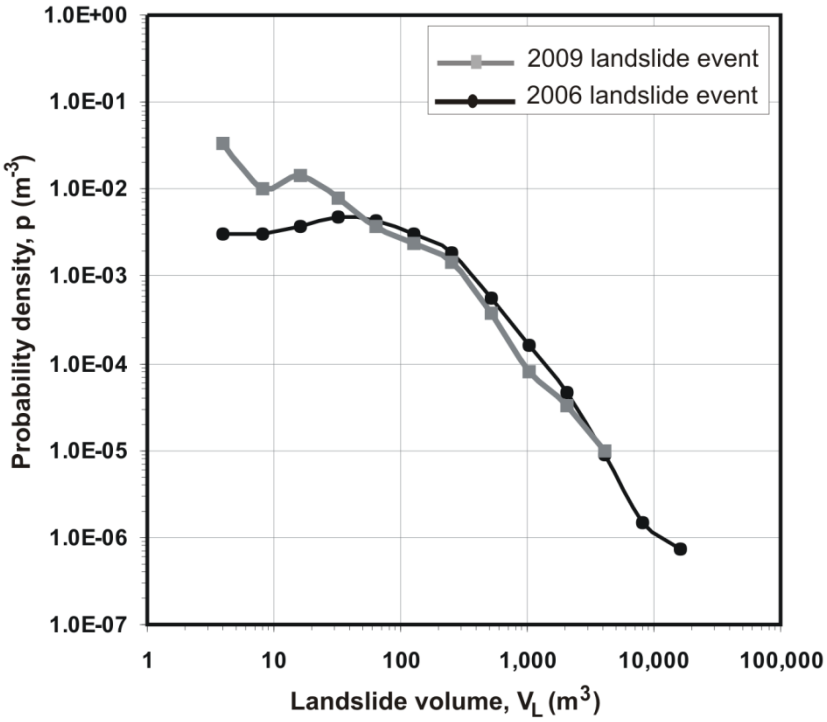


Figure 6.3.17 Probability density distributions for the landslides that occurred in 2006 and 2009

6.3.9.3 Validation of landslide susceptibility model for natural slopes

The most common approach used to validate a susceptibility model is the use of a "prediction rate curve" (Chung and Fabbri, 2003). In this section, we determined the ability of the susceptibility model, obtained using logistic regression analysis, to predict future landslides in the study area. In order to analyze the prediction skill of the susceptibility model, we used source area of landslides on natural slopes triggered by 2009 landslide events and the model proposed by Chung and Fabbri (2003). The model involves computing the proportion of the event landslide area in each susceptibility class, and showing the results using cumulative statistics.

Figure 6.3.18 shows the percentage of the study area, ranked from most to least susceptible (x-axis), against the cumulative percentage of the area of the triggered landslides in each susceptibility class (y-axis), as a dashed black line. The prediction rate curve shows that the most susceptible 20% of the study area contains 65% of the landslide source areas. This 20% of the susceptible areas also contain the high susceptible slopes that are located east of the study area. The model is able to predict about 67% of the landslide areas as unstable group (spatial probability > 0.6). Further, the most susceptible 33% of the study area contains 77% of the landslide areas whilst 80% of the landslide areas are predicted by most susceptible 40% of the study area. Most of the landslides shown in the inventory map are in areas classified as susceptible by the model, although still 12% of the slope failures are in areas with an estimated probability ≤ 0.20. The prediction of the model is better than the success rate in the high susceptible areas. It is reasonable to accept that if a rainfall event triggers landslides equally in the eastern part of the study area then the prediction rate could

be much better. In 2009, even some of the low susceptible areas covering tea plantation also experienced landslides. Very high rainfall on 9 and 10 November 2009 triggered slope failures even in gentle slopes.

The result shown in Figure 6.3.18 provides a quantitative estimate of the model prediction skill. Since the aim of susceptibility modelling is to forecast the spatial location of future landslides, therefore, it is obvious that the model prediction rate should always be better or equivalent than its success rate. In this validation test the prediction of the susceptibility model is better than the model fitting performance shown in Figure 6.3.18 at least for the high susceptible areas. Contrary to this, others researchers (e.g., Chung and Fabbri, 2003; Guzzetti et al., 2005) have observed that the prediction rate of a susceptibility model is often lower than the success rate of the model. They argued that since the success rate measures a "goodness of fit" assuming that the model is "correct" therefore the success rate is always better than the prediction rate for any study area.

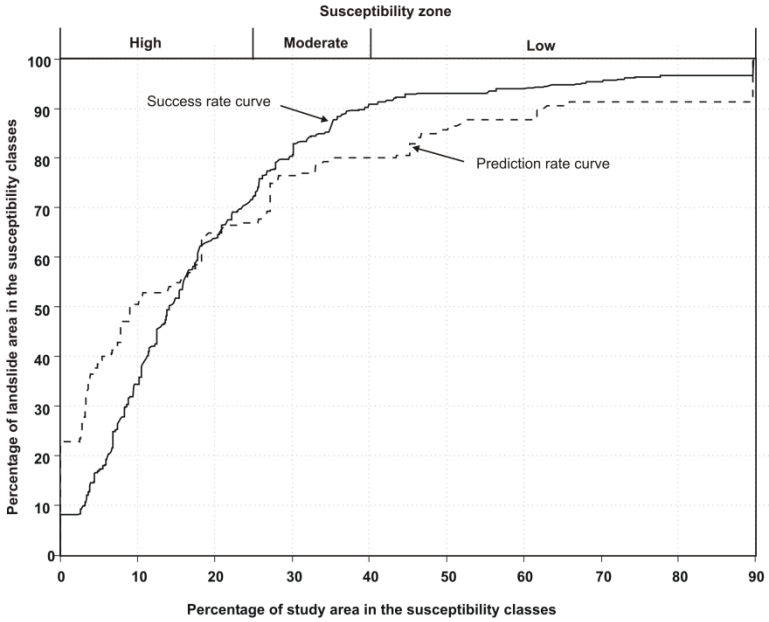


Figure 6.3.18 Graph showing prediction rate (dashed line) of the landslide susceptibility model. Continuous thin line shows success rate

6.3.9.4 Validation of landslide hazard and risk models

The validation of a hazard and risk model for different temporal scenarios is difficult because it requires information on landslides of different time periods that have affected the area after the period considered in the analysis. In a short time period it is highly uncertain to have information on landslide events of different return periods. However, the availability of information on landslides that occurred in 2009 provided an opportunity to validate the hazard and risk models, at least for the available small period of time.

Using the landslide events of 2009, we attempted to validate the results of the Gumbel analysis used to quantify hazard along the transportation lines i.e., the number of landslides per kilometer of cut slopes in different return periods. Table 6.3.8 compares the results of the Gumbel analysis with the occurrence of landslides per kilometer of cut slopes along the railroad in 2009. The table indicates that for km-17, km-19 and km-20, the number of landslides that occurred in 2009 corresponds to a return period of 25 or more years. The amount of rainfall that triggered landslides in this section of the railroad (i.e., 865 mm around Katteri farm) has never occurred in the period considered in the modelling (i.e., 1987 to 2007) but was known to have occurred in 1979 (Seshagiri and Badrinarayan, 1982). The recurrence of a landslide event after 30 years thus validated the results of the Gumbel analysis between km-17 and km-20 along the railroad.

Table 6.3.8 Validation of the landslide hazard along the railroad

Kilometer	Number of landslides per return period (year)					nr of slides in 2009
	3	5	15	25	50	
10	6	9	15	18	21	7
11	6	10	18	21	26	0
12	7	10	16	19	23	9
13	5	8	13	15	18	4
14	4	6	9	10	12	6
15	4	6	9	10	12	2
16	3	4	8	9	11	6
17	3	4	7	8	10	8
18	3	4	6	7	9	5
19	2	3	5	6	7	10
20	2	3	5	6	7	9
21	2	3	5	6	7	3
22	2	3	6	7	9	2
23	2	3	5	6	7	3
24	2	3	5	6	7	1
25	2	3	5	5	6	3
26	1	2	4	5	6	0
Total	56	84	141	164	198	78

In most of the other railroad sections between km-10 and km-17 (i.e., around Burliyar) the number of landslides that occurred in 2009 corresponds to a much lower return period (5 years or less, Table

6.3.8). Although these sections are more prone to landslides but the amount of rainfall that triggered landslides in these sections in 2009 was comparatively less than those recorded around km-19. In fact the rainfall decreased sharply from Coonoor towards the East. The amount of rainfall recorded in Burliyar on 10 November 2009 was 192 mm. Rainfall events with this amount have actually occurred three times in the past 21-year period (1987 to 2007), which corresponds to a recurrence interval of seven years. The seven year return time of the landslide triggering rainfall events corresponds well with the result of the Gumbel analysis for the section of the railroad located between km-10 and km-17.

A total 78 landslides occurred along the entire railroad in 2009, which is comparable to the result of the Gumbel analysis for a 5-year return period (i.e., 84 landslides). It is reasonable to accept here that if the entire area would have received the amount of rainfall that triggered landslides around Katteri farm (corresponding to a landslide event with a 25-year return period), then the total number of landslides would have been much higher (around 160) as predicted by the Gumbel analysis for a scenario with a 25-year return period.

Along the road, the occurrence of 65 landslides is comparable with the model output for a 25-year return period (66 landslides). The occurrence of more landslides between km-390 and km-414 (S-II) triggered by extremely high rainfall corresponds to an event similar to the one from 1979.

Examples in literature that show the validation of a hazard map using landslide events that have happened after the production of the map (the so-called “wait-and-see” method) are very few. In most landslide hazard publications, researchers have shown only the validation of a susceptibility map. Since the hazard models discussed earlier are dependent on the three probabilities i.e., the size, the temporal and the spatial, therefore validation of the three probabilities as discussed before reasonably accounts the predictive capability of the hazard model.

Validation was not possible for the risk models due to the unavailability of data on the losses that occurred in 2009. According to the model, the total direct loss for a 5-year return period along the railroad is estimated as US$ 202,000 and disruption of the rail traffic for about 200 days. In fact due to the 2009 events the railroad was closed for 157 days. The higher value for the estimated route blockage is resulted because of low coefficient of correlation (0.65) of the relation between the total volumes of debris (in m^3) on the railroad and total blockage time (days) used in the analysis.

In case of risk due to slope failures on natural slopes, the obtained risk can not be validated because there are no available data on landslide losses available for the entire area. However, the losses to buildings caused by the 2009 landslide event (see Section 6.3.4), if quantified, may be comparable to the minimum expected losses of a 25-year return period because landslides triggered in 2009 are individually small in size (< 5,000 m^3).

6.3.10 Discussion and conclusions

Validation of hazard and risk models is often difficult due to the lack of sufficient validation datasets. Though the availability of information on 2009 landslide events has given the opportunity to validate the models but still we have to “wait and watch” for more events to occur in order to validate the hazard map for different time scenarios.

For the validation of the complete study area, it is essential to have a validation dataset that is spatially distributed throughout the study area. It is also essential that the landslide triggering condition used for validation should be similar to the one used in the model, which is the basic assumption of the hazard modelling. The landslide events of 2009 occurred due to an extremely high rainfall and affected mostly the western part of the study area. The lack of data in the eastern part makes it practically difficult to validate the susceptibility and the hazard of areas located east of Hillgrove. Difficulties also arise due to a large spatial variability of rainfall over the area. As observed in November 2009, one part of the study area (East) experienced landslide events with a 5-year return period, and at the same time another part (West) experienced events with a 25-year return period. This makes the hazard and loss estimation for the entire area very difficult.

The validation of the prediction skills of all models used in the hazard analysis appears to be reasonable for the given limited time period. The result of the success rate of the susceptibility model (i.e., 73% of the landslide areas occurring within class having spatial probability > 0.6) is comparable to the prediction made by the model (i.e., 67% of the landslide areas of 2009 occurring within class having spatial probability > 0.6) and therefore the distribution of landslides shown by the success rate provide a good approximation of the distribution of future landslides in different susceptible classes within the study area.

This research is published in a number of papers by Jaiswal et al. (2008, 2009, 2010, 2011).

References

AGS Australian Geomechanics Society and Sub-committee on landslide risk management, 2000. Landslide risk management concepts and guidelines, Aust. Geomech. 35(1), 49–92.

Chung C.J., Fabbri A.G., 2003. Validation of spatial prediction models for landslide hazard mapping. Natural Hazards 30, 451-472.

Coe J.A., Michael J.A., Crovelli R.A., Savage W.Z., 2000. Preliminary map showing landslide densities, mean recurrence intervals, and exceedance probabilities as determined from historic records, Seattle, Washington. USGS Open-File report 00-0303.

http://pubs.usgs.gov/of/2000/ofr-00-0303, cited on 15 July 2008.

Coe J.A., Michael J.A., Crovelli R.A., Savage W.Z., Laprade W.T., Nashem W.D., 2004. Probabilistic assessment of precipitation-triggered landslides using historical records of landslide occurrence, Seatle, Washington. Environmental and Engineering Geoscience X(2), 103-122.

Fell R., Corominas J., Bonnard C., Cascini L., Leroi E., Savage W.Z., 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology 102, 85-98.

Guzzetti F., Reichenbach P., Cardinali M., Galli M., Ardizzonem F., 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72, 272– 299.

Ibsen M., Brunsden D., 1996. The nature, use and problems of historical archives for the temporal occurrence of landslides, with specific reference to the south coast of Britain, Ventnor, Isle of Wight. Geomorphology 15, 241–258.

Jaiswal P., van Westen C.J., Jetten V., 2011. Use of landslide hazard and risk information in risk reduction planning. Natural Hazard, under preparation.

Jaiswal P., van Westen C.J., 2011. Frequency-size relation of shallow debris slides on cut slopes along a railroad corridor. Natural Hazard, under preparation.

Jaiswal P., van Westen C.J., Jetten V., 2011. Quantitative estimation of landslide risk from rapid debris slides on natural slopes in the Nilgiri hills, India. Nat. Hazards Earth Syst. Sci., under review.

Jaiswal P., van Westen C.J., Jetten V., 2011. Quantitative assessment of landslide hazard along transportation lines using historical records. Landslides, DOI 10.1007/s10346-011-0252-1.

Jaiswal P., van Westen C.J., Jetten V., 2010. Quantitative landslide hazard assessment along a transportation corridor in southern India. Engineering Geology 116, 236-250.

Jaiswal P., van Westen C.J., Jetten V., 2010. Quantitative assessment of direct and indirect landslide risk along transportation lines in southern India. Nat. Hazards Earth Syst. Sci. 10, 1253-1267.

Jaiswal P., van Westen C.J., 2009. Estimating temporal probability for landslide initiation along transportation routes based on rainfall threshold. Geomorphology 112, 96-105.

Jaiswal P., van Westen C.J., 2009. Rainfall - based temporal probability for landslide initiation along transportation routes in Southern India. in: Malet J.P., Remaitre A., Bogaard T., (eds.), Landslide Processes: from geomorphologic mapping to dynamic modelling, proceedings of the Landslide Processes Conference, European Centre on Geomorphological Hazards (CERG), Strasbourg, France, 6-7 February, 2009. ISBN 2-95183317-1-4, pp. 139-143.

Jaiswal P., van Westen C.J., 2009. Probabilistic landslide initiation hazard assessment along a transportation corridor in the Nilgiri area, India. Geophysical Research Abstracts, Vol. 11, EGU General Assembly 2009.

Jaiswal P., van Westen C.J., Jetten V.G., 2008. Use of historic landslide inventories for hazard assessment along communication routes._ Geophysical Research Abstracts, Vol. 10, EGU2008-A-02948, 2008, SRef-ID: 1607-7962/gra/EGU2008-A-02948.

Malamud BD., Turcotte DL., Guzzetti F., Reichenbach P., 2004. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29(6), 687– 711.

Remondo J., Bonachea J., Cendrero A., 2008. Quantitative landslide risk assessment and mapping on the basis of recent occurrences. Geomorphology 94, 496-507.

Stark C.P., Hovius N., 2001. The characterization of landslide size distributions. Geophysics Research Letters 28 (6), 1091– 1094.

van Westen C.J., Asch T.W.J., Soeters R., 2006. Landslide hazard and risk zonation-why is it still so difficult? Bull. Eng. Geol. Env. 65, 67-184.

Zezere J.L., Oliveira S.C., Garcia R.A.C., Reis E., 2007. Landslide risk analysis in the area North of Lisbon (Portugal): evaluation of direct and indirect costs resulting from a motorway disruption by slope movements. Landslides 4, 123–136.