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 D1.5

Statistical and empirical models for prediction of precipitation-induced landslides

Work Package 1.3 – Statistical studies of thresholds for precipitation-induced landslides

Deliverable/Work Package Leader: ICG Revision: 2

April, 2012

Rev.	Deliverable Responsible	Controlled by	Date
0	ICG		2010-04-27
1	ICG	EPFL	2010-12-20
2	ICG	AMRA	2012-04-20

SUMMARY

This deliverable presents statistical and empirical models for predicting critical meteorological elements and their thresholds for triggering of landslides at local and regional scales. The models are evaluated using rainfall data and landslide observations from five datasets from France, Switzerland, Italy and Norway. The general conclusions of the evaluations are: (i) the analysis of hourly data, when available, indicates that the duration of triggering storms for initiating debris flows is less than 12 hours; (ii) in soil slides, the antecedent moisture conditions play an important role in the triggering conditions, and therefore should be accounted for in the threshold models; these antecedent conditions can be characterised by the soil moisture or other equivalent proxies, such as long-term precipitation and water supply; (iii) the contribution of snow melt to the of the water supply conditions on a slope should be evaluated, as it can be significant for characterising both antecedent and triggering conditions; (iv) freeze-thaw effects should be incorporated in the assessment of thresholds for rock falls and rock slides in order to reduce the misclassification errors for this type of landslides; (v) the variations of thresholds in different locations are mainly due to differences in lithological conditions, and in the case of soils the variations in permeability and depth of sediments; the soil moisture conditions (and the corresponding proxies in the threshold models) in areas with deep deposits of fine-grained soils (low permeability) are affected by long term precipitation or water supply; in areas with shallower deposits of coarse-grained soils (high permeability), the response of the ground water conditions to precipitation and water supply is faster, and therefore the thresholds are controlled almost solely by short-term precipitation or water supply.

The application of models at a local scale to the La Frasse dataset (Switzerland) demonstrates the potential for integrating field measurements of landslide displacements and rainfall observations for effectively performing forward predictions. The Barcelonnette dataset (France) was assessed for thresholds for debris flows and soil slides, which indicated triggering durations in the range of 1-9 hours, and 3-17 hours, respectively. An intensity-duration threshold was sufficient for debris flows, but for soil slides, an intensity-antecedent precipitation-duration threshold was necessary to achieve improved performance. The critical antecedent precipitation corresponded to 23 days. At a regional scale, antecedent precipitation models were also applied to datasets from two locations in Norway using daily rainfall observations. The results from this model indicate that 1-day and 7-day antecedent precipitation are critical for debris flows and soil slides in a case study in Western Norway, while 46-day antecedent precipitation is critical for earth slides in a selected case study in South-Eastern Norway. The assessment in Western Norway indicated that the occurrence of rock slides and rock falls is weakly associated with rainfall parameters, suggesting the necessity of incorporating other relevant effects, such as freeze-thaw conditions. Modelling of the occurrence of debris flows requires the use of rainfall observations with a high sampling rate (e.g., hourly). The FLaIR model was calibrated successfully for predicting triggering conditions for soil slides in the Barcelonnette basin, and in three case studies in Italy. The appendices present experiences of parallel studies that clearly indicate that water supply (precipitation combined with snow melt) can be a critical parameter for threshold evaluation.

Note about contributors

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

Lead partner responsible for the deliverable:
ICG

Deliverable prepared by:
José Cepeda, Simone Colonnelli, Nele Kristin Meyer, Kalle Kronholm

Partner responsible for quality control:
EPFL, AMRA

Other contributors:

EPFL:	Alessio Ferrari, Lyesse Laloui, John Eichenberger and A. Marvuglia
AMRA:	Pasquale Versace, Carolina Maria Gualtieri, Davide De Luca
CNRS:	Alexandre Remaître and Jean-Philippe Malet
ETHZ:	Harikrishna Narasimhan and Michael Faber

CONTENTS

1 OVERVIEW OF THE WORKPACKAGE

The following presents an overview of the WP.

1.1 OBJECTIVES

Evaluate the reliability of statistical and empirical models for predicting critical meteorological elements and their thresholds for triggering of landslides at the local and regional scale. This WP, integrated with WP1.2, provides a complete set of tools for prediction of landslides induced by weather, with particular reference to climatic changes, furnishes input for early warning systems and identifies the key parameters that should be focused on in the climate change scenarios.

1.2 DESCRIPTION OF WORK

Selected statistical and empirical models will be used to identify the most important critical meteorological elements and the thresholds of the elements responsible for the triggering of landslides at the local and regional scales. The models will be tested on rich datasets from several sites to enable a direct comparison of the model results (benchmarking). This dataset will be essentially provided partners through monitoring of the established test sites, but can include also data from the literature and from other sites. In particular, the model results will be compared using a set of in terms of the goodness and uncertainty of their predictions under the assumption that the models are implemented in an early warning system. The WP will comprise the following tasks.

1.2.1 Task 1: Identification of models most suitable for the testing

The models most suitable for testing will be identified from the literature. The selected models must a) handle either of the two scales under consideration – i.e. the local and the regional scales – or both scales, and b) must be able to identify both the most critical meteorological triggering elements as well as the element threshold. The selected models will include artificial neural networks, regression analysis, classification trees and logistic regression. They will be carefully tested through data provided by literature and monitoring, in order to assess their reliability in different geomorphological and environmental situations, even very different from those for which they have been conceived and developed.

1.2.2 Task 2: Defining a set of evaluation criteria for model comparison

The evaluation criteria to compare the models by must assume that these are a part of an early warning system. This means that the models should identify various levels of alarm stages (e.g. alert, warning, and evacuation) with a minimum number of false alarms and missed alarms and a maximum number of correctly predicted alarm stages. With the optimised evaluation criteria, a set of critical meteorological elements and their thresholds are obtained. Available data, concerning for instance the Sarno site in Italy, where an early warning system is active, or other sites, can be used for testing of the reliability of the procedures associated with the different systems. The comparison of different methods might enable to assess what could have been happened during weather crises, if one system or another would have been used.

1.2.3 Task 3: Model evaluation

The responsible institutes will use their models to produce a set of the most important critical meteorological elements, the threshold values and the resulting optimised evaluation criteria as

defined in Task 2. Based on these results the models will be compared over the different scales and test sites.

1.3 DELIVERABLES

D1.5 – Statistical and empirical models for prediction of precipitation-induced landslides.

2 MODELS SUITABLE FOR TESTING

This chapter gives an overview of the models that were deemed suitable for testing, and describes the work carried out in Task 1.

2.1 GENERAL REQUIREMENTS

2.1.1 Scale

The models tested must handle prediction of events at two scales: a) local scale and b) regional scale. That is, the same model setup must be able to handle both scales. However, the input data and resulting meteorological elements and associated thresholds can of course be different at the two scales.

2.1.2 Triggering elements

The models must be able to handle data of the most usual meteorological elements associated with landslide events. The following list includes some of these elements:

- Precipitation (hourly, daily, summed values).
- Temperature (hourly, daily, min, max).
- Snowmelt.
- Wind speed and wind direction.

In addition, the models must be able to handle the combination of at least two of these triggering elements, for example a long term precipitation threshold combined with a short term precipitation threshold.

2.1.3 Alarm stages

If the tested models are coupled with meteorological forecasts, they can be part of early-warning systems for managing landslide risk (Keefer et al. 1987). In order to achieve this, it is proposed that the tested models should be able to predict four alarm stages as indicated in the table below.

Table 1. Alarm stages for model results.

	Organisational consequence	Probability of triggering
Green	None	<5%
Yellow	Alert	5-20%
Orange	Warning	20-50%
Red	Evacuation	>50%

2.1.4 Penalty weights for falsely predicted alarm stages

The models must be able to incorporate penalty weights for wrong prediction of alarm stages. The weights for falsely predicted alarm stages are given in chapter 3.

Since not all available datasets have these alarm stages, weights differ between the datasets. Weights for each dataset are described in chapter 4.

2.1.5 Input to climate change models

In order to predict possible future changes in landslide frequency, meteorological elements and associated triggers should be able to be used in conjunction with climate change models.

According to communication with colleagues in SafeLand working with climate change models, the models to be used in SafeLand are able to produce data with the following characteristics:

- Temporal resolution: Hourly values are stored.
- Spatial resolution: 10 km or 25 km. This can be further aggregated or be reanalyzed using spatial statistics.
- Thresholds in absolute values can be used as input.
- Thresholds from multiple meteorological elements can be used with the models.

Hence, there appears to be no major restrictions in the type of thresholds that are the outcome of the tested models.

2.2 CONCEPTUAL DEFINITION OF A THRESHOLD

For the purposes of this report, the following definitions are established:

- **Meteorological element (*ME*)**: a value or statistic that can be directly obtained from:
  - readings at a meteorological station,
  - weather forecasts,
  - climate model predictions,
  - remote sensing devices (radar, satellites).

  The value should be in "raw" conditions, i.e. it should not be the result of the application of empirically calibrated functions to a set of meteorological elements. Examples of meteorological elements: peak 3-hour precipitation, observed daily precipitation, storm duration, 30-day accumulated precipitation, 24-hour forecasted precipitation, minimum temperature, wind speed, wind direction, differences in snow depth, NASA's TMPA data, etc. An example of a value that, in this context, is not considered as a meteorological element is snowmelt estimated using both empirically calibrated functions and observations of temperature and wind speed.
- **Meteorological event**: a continuous occurrence of meteorological elements that is bracketed in time based on conventional criteria. For example, it can be defined as a continuous period during which the hourly precipitation has been greater than 1 mm.
- **Time element (*TE*)**: a categorical or continuous variable that describes the time or season of the year when a meteorological event takes place. Possible choices of time elements are: the Julian date (*JD*), the Julian day number (*JDN*), the season of the year, month, etc.
- **Meteorological-Time set of elements (*MTS*)**: is the combination of a selection of meteorological and time elements that characterise a meteorological event. A very common choice of *MTS* is: date, time, duration (*D*) and mean precipitation intensity (*I*).
- **Landslide incidents (*LI*)**: is a continuous or categorical variable that describes the occurrence, extent, type and size of landslide events associated to a meteorological event. For example, this variable can be associated to the categories or classes: "no landslides", "debris flows", "rock falls", "rock falls and debris flows", etc. The categories can also include linguistic descriptions of the extent of occurrence ("single", "few", "many", "abundant") and size ("small", "moderate", "large"). These categories can be associated to numerical ranges of frequency and volume. In its simplest form, this variable can be associated to a binary result: "no landslides" or "landslides".

- ***MTSLI* time series**: a collection of a Meteorological-Time set of elements (*MTS*) and Landslide Incidents (*LI*) in a region during a period of observations.
- ***ME* space:** is an *m*-dimensional space, where *m* is the number of meteorological elements (*ME*) in a *MTSLI* time series. The most usual *ME* space is a 2D space (*m* = 2) for representing mean precipitation intensity (*I*) and duration (*D*).
- **Threshold (*Th*):** is a boundary separating regions in an *ME* space based on the classification of an *MTSLI* time series using the Landslide incidents (*LI*) variable. In most of the cases, only two meteorological elements *ME* are used, so the *ME* space is a coordinate plane and the threshold is a curve. In particular, if only the non-occurrence and occurrence cases (*LI* = "no landslides" or *LI* = "landslides") are considered, a threshold is a curve separating meteorological events that triggered (one or more) landslides from those that did not trigger any.

Depending on the available classes of meteorological events, thresholds can be estimated as:

- Only occurrences of landslides: the threshold is a lower bound to the meteorological events. See Figure 1a. Examples: Caine (1980), Crosta (1998) and Guzzetti et al. (2007; 2008).
- Only non-occurrences of landslides: the threshold is an upper bound to the meteorological events. See Figure 1b. Examples: Cannon and Ellen (1988).
- Both, occurrence and non-occurrences of landslides: in this case, the threshold is a boundary separating the two classes. See Figure 1c. Examples: Wieczorek and Sarmiento (1988) and Jakob (2009). Some authors have estimated thresholds by fitting curves to the occurrence class (Rodolfo and Arguden 1991; Van Westen and Daag 2005).

From Figure 1, it can be noted that a threshold may not classify correctly all the meteorological events. Events with observed landslides ("+" in Figure 1) and located below the threshold are missed events, while events with no landslides ("-" in Figure 1) and above the threshold are false alarms. One of the main challenges in the estimation of thresholds is to achieve an optimum performance such that both missed events and false alarms are minimised as much as practical operational conditions allow (i.e., in the framework of an early-warning system). The assessment of the performance of thresholds is further discussed in section 2.3.7.

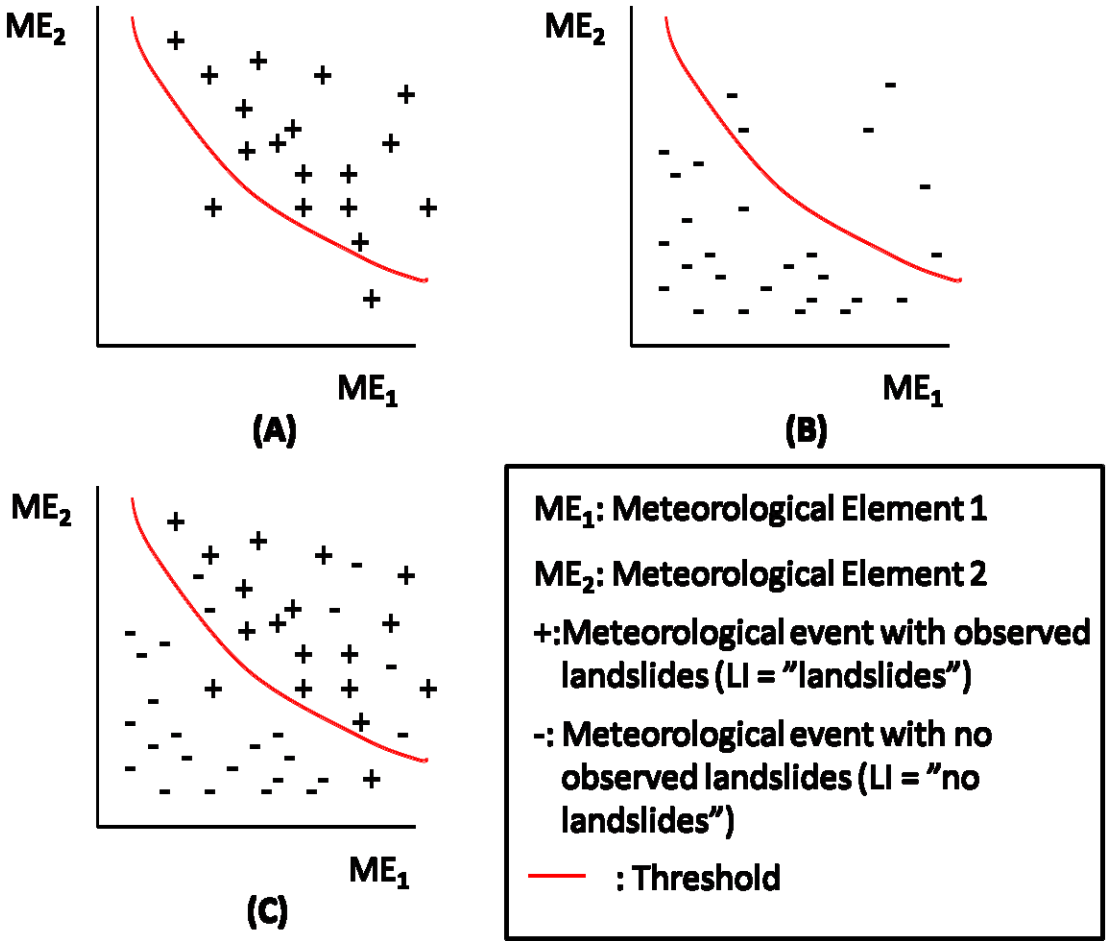


Figure 1. **Types of thresholds based on the available classes of meteorological events: (a) Only landslide occurrences available; (b) only non-occurrences available; (c) both occurrences and non-occurrences are available. For simplicity, a threshold based only on two meteorological elements (*ME*) is shown. The proportion of "landslide" events in relation to "no landslides" events is usually very small and does not correspond to the proportion shown in (c), which is presented for illustration purposes only. Adapted from Cepeda and Devoli (2008).**

If more than two classes are defined for the variable *LI*, then more than one threshold can be estimated. As an example, Figure 2 shows the set of thresholds proposed by Lumb (1975). The classes "Isolated", "Minor", "Severe" and "Disaster" were established based on the number of landslide observations and using limiting values of 1, 10 and 50 events. Note that, in this case, the meteorological elements (*ME*) are 1-day rainfall and 15-day rainfall.

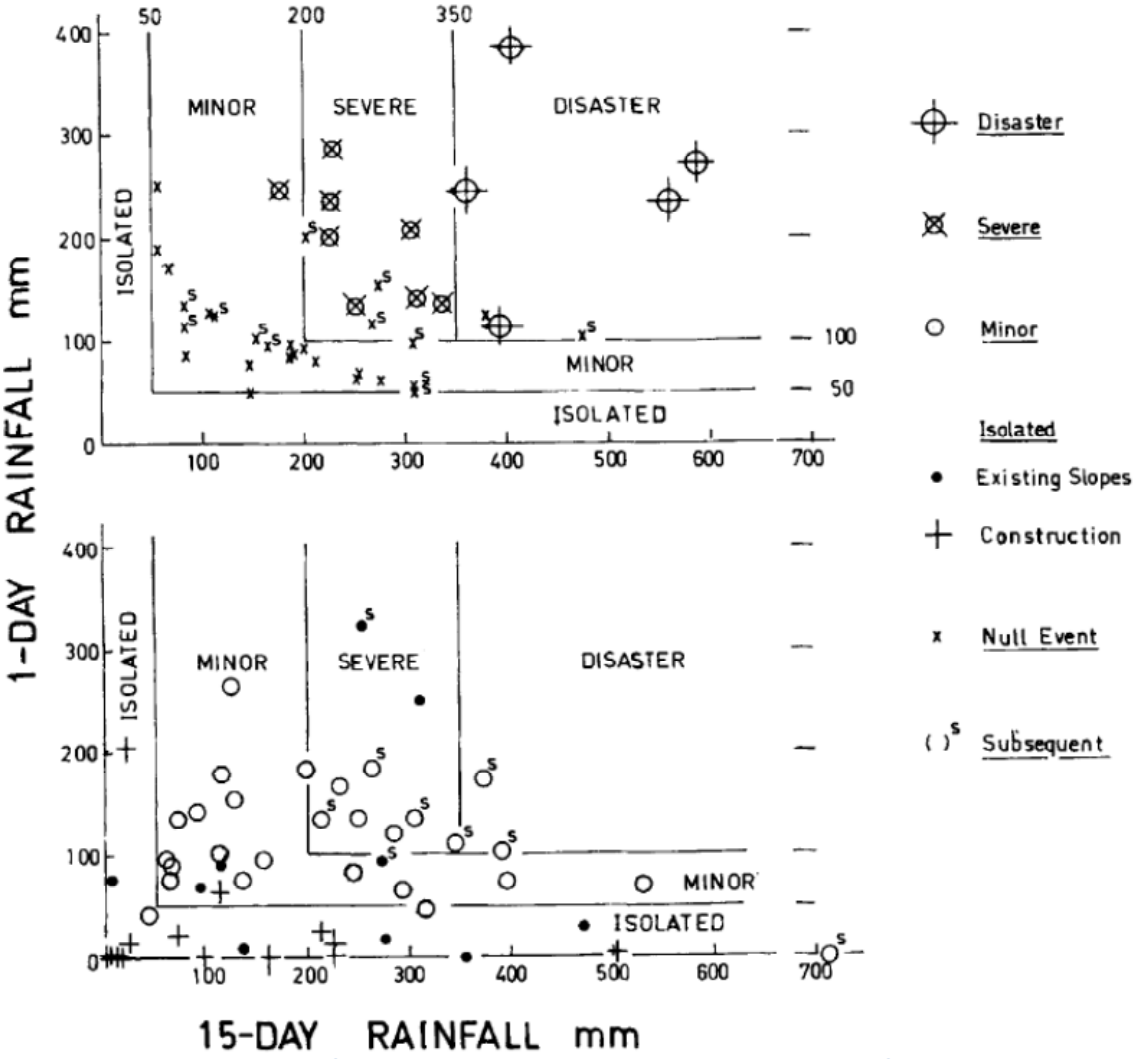


Figure 2. Thresholds for Hong Kong based on the combination of 1-day and 15-day rainfall (Lumb 1975)

Data mining and pattern recognition techniques can be useful for prediction of extreme precipitation events (Dhanya and Kumar 2009). A number of useful references on pattern recognition and data mining can be identified in the literature (Jain et al. 2000; Chen et al. 1996; Cortes and Vapnik 1995; Burges 1998; Jain et al. 1999; Fayyad et al. 1996)

In the estimation of thresholds, antecedent precipitation has a predominant importance since it controls initial groundwater conditions before the onset of a landslide-triggering rainfall event. The first author to note effects of antecedent precipitation was Campbell (1975). The time interval that is critical for the calculation of antecedent precipitation is a matter of debate (Rahardjo et al. 2001; Wieczorek and Glade 2005).

2.3 CLASSIFICATION TECHNIQUES

2.3.1 Visual

The earliest estimations of thresholds classified events based on visual inspection of the data and using constant thresholds or power-law forms. Some key references are Lumb (1975), Caine (1980), Larsen and Simon (1993), among others.

2.3.2 Linear regression

Estimations of thresholds for landslides in active volcanic settings have usually fitted least-square curves to landslide-triggering rainfall events. Many of these landslides are not initiated as a rupture in the terrain but as a gradual sediment-bulking of erosion runoff. Important references are Rodolfo and Arguden (1991) and van Westen and Daag (2005). These authors have estimated thresholds for active volcanoes in the Philippines.

2.3.3 Logistic regression

The logistic regression model predicts the probability of occurrence of a landslide-triggering rainfall by fitting the precipitation observations to a logistic curve. Frattini et al. (2009) have employed this model for estimations of thresholds in northern Italy.

2.3.4 Discriminant analysis

The aim of a discriminant analysis is to find a linear function that combines several parameters (e.g., rainfall parameters) and enables the separation of two or more classes of events (e.g., non-landslide triggering and landslide triggering). This model has been used for estimation of thresholds in British Columbia, Canada (Jakob and Weatherly 2003; Jakob 2009).

In the implementation of a discriminant analysis, it is possible to assign values of prior probabilities to the classes of observations. The sum of the prior probabilities of the classes is 1. In practice, these prior probabilities operate as weights of the classes. For example, a prior probability of 0.9 to landslide-triggering events (and thus, 0.1 to non-triggering events) will produce a conservative classification (a threshold) that gives preference to reducing the number of missed events, but lowers the priority to the reduction of false alarms. In the opposite case, a prior of 0.9 to non-triggering events (and thus, 0.1 to landslide-triggering events) will produce a non-conservative classification (a threshold) that gives preference to reducing the number of false alarms, but lowers the priority to the reduction of missed events. The default choice is to assign the same prior probability to both triggering and non-triggering events (i.e., both are 0.5). Varying the prior probabilities in the evaluation of a model allows estimating various sets of model parameters for obtaining the corresponding Receiver Operating Characteristic (ROC) curve (see sections 2.5 and 3.2). Furthermore, based on the conceptual framework proposed by Cepeda et al. (2010), the prior probabilities can serve as a proxy to landslide susceptibility levels.

2.3.5 Bayesian inference

A method of statistical inference is the Bayesian inference, in which observations are fed to a model in order to update or produce new inferences regarding some parameters or hypotheses. This method has been employed for assessment of global and European thresholds (Guzzetti et al. 2007; Guzzetti et al. 2008).

2.3.6 Classification trees

Classification trees are binary classifiers that are constructed by repeated splits of subsets of a sample into two descendant subsets, beginning with the whole sample itself. Each split occurs at a node, usually defined by a relational expression involving a predictor (rainfall parameter) and a constant (a threshold). Each terminal subset is a leaf of the tree. A classic background reference on this method is Breiman et al. (1984). Classification trees have been used for estimation of thresholds for triggering of landslides (Rességuier 2006).

2.3.7 Comparison of classification techniques

Patel (2003) has suggested some guidelines for the comparative assessment of several classifications techniques, in the context of data mining. These guidelines are presented in Table 2.

Table 2. Comparison of data mining techniques as a guideline for large data sets (Patel 2003).

H: high, M:medium, L:low.

	Multiple Linear Regression	Logistic Regression	Discriminant Analysis	Naïve Bayes	Neural Nets	Trees	k-Nearest Neighbors
Accuracy	M	M	M	HM	H	M	HM
Intepretability	H	H	M	H	L	H	L
Speed-Training	H	H	H	HM	L	HM	H
Speed-Deployment	H	H	H	H	H	HM	L
Effort in choice and transformation of indep.Vars.	HM	HM	HM	HM	L	L	ML
Effort to tune performance parameters	L	L	L	ML	H	ML	ML
Robustness to Outliers in indep vars	ML	ML	ML	ML	HM	H	HM
Robustness to irrelevant variables	H	H	HM	H	L	ML	L
Ease of handling of missing values	M	M	M	H	M	H	ML
Natural handling both categorical and continuous variables	H	H	ML	M	H	H	L

Receiver Operating Characteristic (ROC) analysis is a useful technique for comparing the performance of classification models. A comprehensive review of ROC analysis is presented by Fawcett (2006). An application of ROC analysis to rainfall thresholds is presented in section 2.5.

2.4 I-D MODELS

Intensity-duration (I-D) thresholds are the most common type of empirical thresholds proposed in the literature for predicting landslide occurrence induced by rainfall. They are devoted to forecast first-time failures of shallow landslides and triggering of rapid mass movements. The fundamental assumption is that, in a homogeneous region, landslides occur once a precipitation threshold has been overcome. Combinations of intensity and duration for rainfall events which resulted in landslides are usually plotted in logarithmic coordinates and the thresholds are obtained by drawing lower-bound lines to the points. Most commonly, the thresholds are drawn visually, i.e., without any rigorous mathematical, statistical, or physical criterion (Guzzetti et al. 2007). The first I-D threshold was estimated by Starkel (1979).

Threshold equations

Caine (1980) was the first in attempting to define in a general way the rainfall threshold for catastrophic slope failures of the debris flow type at a global scale. The author pointed out that the rainfall threshold was neither defined by the total amount of rainfall nor by instantaneous rainfall intensity. Instead, a product of these two components was to be used. Analyzing data coming from different climatic zones and involving many geological systems at different antecedent moisture conditions in the failed material the following threshold relation was proposed between rainfall intensity (I, in mm/h) and duration (D, in h):

$$I = 14.82 \cdot D^{-0.61} \quad (1)$$

Since the work of Caine (1980), different rainfall I-D thresholds have been proposed. Guzzetti et al. (2007) summarized the relationships proposed by 52 authors worldwide (Table 3 and Figure 3). The general form for the I-D threshold is

$$I = c + aD^{b} \quad (2)$$

where I is the average intensity and a, b and c being curve parameters. Curves are usually represented in logarithmic scales. When $c = 0$ a straight line is obtained in such a representation. This would imply that for long lasting rainfall events even limited intensities could induce landslides. Introducing a c coefficient different than zero allows considering an asymptotic behaviour for long rainfall durations (Wieczorek 1987; Crosta and Frattini 2001), imposing a minimum intensity (c) for high values of D.

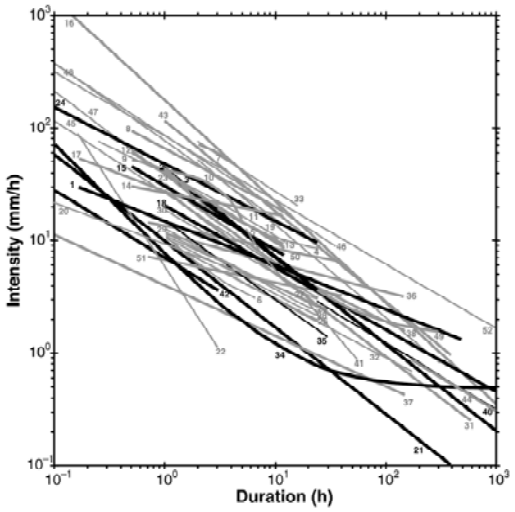


Figure 3. **Collection of rainfall I-D thresholds; numbers refer to # in Table 3; legend: very thick line, global threshold; thick line, regional threshold; thin line, local threshold (Guzzetti et al. 2007).**

Table 3 I-D thresholds for the initiation of landslides. Extent: G, global threshold; R, regional threshold; L, local threshold. Area, the area where the threshold was defined. Landslide type: A, all types; D, debris flow; S, soil slip; Sh, shallow landslide, L, lahar. Rainfall intensity in mm/h, rainfall duration in hours. Equations in italics were estimated. Source: 1, Caine (1980); 2, Moser and Hohensinn (1983); 3, Cancelli and Nova (1985); 4–5, Cannon and Ellen (1985); 6, Wieczorek (1987); 7–15, Jibson (1989); 16, Guadagno (1991); 17, Rodolfo and Arguden (1991); 18, Ceriani et al (1992); 19, Larsen and Simon (1993); 20, Arboleda and Martínez (1996); 21, Clarizia et al (1996); 22, Tuñgol and Regalado (1996); 23, Zimmermann et al. (1997); 24, Paronuzzi et al (1998); 25–30, Bolley and Olliaro (1999); 31, Calcaterra et al. (2000); 32, Montgomery et al. (2000); 33, Wieczorek et al. (2000); 34, Crosta and Frattini (2001); 35, Marchi et al. (2002); 36, Ahmad (2003); 37, Jakob and Weatherly (2003); 38, Aleotti (2004); 39, Barbero et al. (2004); 40, Floris et al. (2004); 41, Baum et al. (2005); 42, Cannon and Gartner (2005); 43, Chien-Yuan et al. (2005); 44, Corominas et al (2005); 45–48, Giannecchini (2005); 49, Hong et al. (2005); 50–51, Jan and Chen (2005); 52, Zêzere et al. (2005). From Guzzetti et al. (2007).

#	Extent	Area	Landslide type	Equation	Range
1	G	World	Sh, D	$I = 14.82 \times D^{-0.39}$	$0.167 < D < 500$
2	R	Carinthia and E Tyrol, Austria	S	$I = 41.66 \times D^{-0.77}$	$1 < D < 1000$
3	L	Valtellina, Lombardy, N Italy	S	$I = 44.668 \times D^{-0.78}$	$1 < D < 1000$
4	L	San Francisco Bay Region, California	D	$I = 6.9 + 38 \times D^{-1.00}$	$2 < D < 24$
5	L	San Francisco Bay Region, California	D	$I = 2.5 + 300 \times D^{-2.00}$	$5.5 < D < 24$
6	L	Central Santa Cruz Mountains, California	D	$I = 1.7 + 9 \times D^{-1.00}$	$1 < D < 6.5$
7	R	Indonesia	D	$I = 92.06 - 10.68 \times D^{1.00}$	$2 < D < 4$
8	R	Puerto Rico	D	$I = 66.18 \times D^{-0.52}$	$0.5 < D < 12$
9	R	Brazil	D	$I = 63.38 - 22.19 \times D^{1.00}$	$0.5 < D < 2$
10	R	China	D	$I = 49.11 - 6.81 \times D^{1.00}$	$1 < D < 5$
11	L	Hong Kong	D	$I = 41.83 \times D^{-0.58}$	$1 < D < 12$
12	R	Japan	D	$I = 39.71 \times D^{-0.62}$	$0.5 < D < 12$
13	R	California	D	$I = 35.23 \times D^{-0.54}$	$3 < D < 12$
14	R	California	D	$I = 26.51 \times D^{-0.19}$	$0.5 < D < 12$
15	G	World	D	$I = 30.53 \times D^{-0.57}$	$0.5 < D < 12$
16	R	Peri-Vesuvian area, Campania Region, S Italy	D	$I = 176.40 \times D^{-0.90}$	$0.1 < D < 1000$
17	L	Mayon, Philippines	L	$I = 27.3 \times D^{-0.38}$	$0.167 < D < 3$
18	R	Lombardy, N Italy	A	$I = 20.1 \times D^{-0.55}$	$1 < D < 1000$
19	R	Puerto Rico	A	$I = 91.46 \times D^{-0.82}$	$2 < D < 312$
20	L	Pasig-Potrero River, Philippines	L	$I = 9.23 \times D^{-0.37}$	$0.08 < D < 7.92$
21	G	World	S	$I = 10 \times D^{-0.77}$	$0.1 < D < 1000$
22	L	Sacobia River, Philippines	L	$I = 5.94 \times D^{-1.50}$	$0.167 < D < 3$
23	R	Switzerland	A	$I = 32 \times D^{-0.70}$	$1 < D < 45$
24	R	NE Alps, Italy	D	$I = 47.742 \times D^{-0.507}$	$0.1 < D < 24$
25	L	Rho Basin, Susa Valley, Piedmont, NW Italy	D	$I = 9.521 \times D^{-0.4955}$	$1 < D < 24$
26	L	Rho Basin, Susa Valley, Piedmont, NW Italy	D	$I = 11.698 \times D^{-0.4783}$	$1 < D < 24$
27	L	Perilleux Basin, Piedmont, NW Italy	D	$I = 11.00 \times D^{-0.4459}$	$1 < D < 24$
28	L	Perilleux Basin, Piedmont, NW Italy	D	$I = 10.67 \times D^{-0.5043}$	$1 < D < 24$
29	L	Champeyron Basin, Piedmont, NW Italy	D	$I = 12.649 \times D^{-0.5324}$	$1 < D < 24$
30	L	Champeyron Basin, Piedmont, NW Italy	D	$I = 18.675 \times D^{-0.565}$	$1 < D < 24$
31	R	Campania, S Italy	A	$I = 28.10 \times D^{-0.74}$	$1 < D < 600$
32	L	Mettman Ridge, Oregon	A	$I = 9.9 \times D^{-0.52}$	$1 < D < 170$
33	L	Blue Ridge, Madison County, Virginia	D	$I = 116.48 \times D^{-0.63}$	$2 < D < 16$
34	G	World	Sh	$I = 0.48 + 7.2 \times D^{-1.00}$	$0.1 < D < 1000$
35	L	Moscardo Torrent, NE Italy	A	$I = 15 \times D^{-0.70}$	$1 < D < 30$
36	R	E Jamaica	Sh	$I = 11.5 \times D^{-0.26}$	$1 < D < 150$
37	R	North Shore Mountains, Vancouver, Canada	Sh	$I = 4.0 \times D^{-0.45}$	$0.1 < D < 150$
38	R	Piedmont, NW Italy	Sh	$I = 19 \times D^{-0.50}$	$4 < D < 150$
39	L	Piedmont, NW Italy	A	$I = 44.668 \times D^{-0.78} \times N$	$1 < D < 1000$
40	L	Valzangona, N Apennines, Italy	A	$I = 18.83 \times D^{-0.59}$	$24 < D < 3360$
41	L	Seattle Area, Washington	S	$I = 82.73 \times D^{-1.13}$	$20 < D < 55$
42	G	World	D	$I = 7.00 \times D^{-0.60}$	$0.1 < D < 3$
43	R	Taiwan	A	$I = 115.47 \times D^{-0.80}$	$1 < D < 400$
44	R	Pyrenees, Spain	A	$I = 17.96 \times D^{-0.59}$	$D > 168$
45	L	Apuane Alps, Tuscany, Italy	Sh	$I = 26.871 \times D^{-0.638}$	$0.1 < D < 35$
46	L	Apuane Alps, Tuscany, Italy	Sh	$I = 85.584 \times D^{-0.781}$	$0.1 < D < 35$
47	L	Apuane Alps, Tuscany, Italy	Sh	$I = 38.363 \times D^{-0.743}$	$0.1 < D \leq 12$
48	L	Apuane Alps, Tuscany, Italy	Sh	$I = 76.199 \times D^{-0.692}$	$0.1 < D \leq 12$
49	R	Shikoku Island, Japan	A	$I = 1.35 + 55 \times D^{-1.00}$	$24 < D < 300$
50	R	Central Taiwan	D	$I = 13.5 \times D^{-0.20}$	$0.7 < D < 40$
51	R	Central Taiwan	D	$I = 6.7 \times D^{-0.20}$	$0.7 < D < 40$
52	L	N of Lisbon, Portugal	A	$I = 84.3 \times D^{-0.57}$	$0.1 < D < 2000$

Methodology to establish the thresholds

For most of the published rainfall thresholds, the applied mathematical or statistical criteria are not specified (Guzzetti et al. 2007). Several methodologies have been adopted (Guzzetti et al. 2008):

- The threshold is drawn manually on a plot using linear or logarithmic coordinates as the lower boundary for the rainfall conditions that have induced slope failures;
- The equation describing the threshold is first adopted (e.g., a power law, hyperbolae, etc.), and the parameters of the selected equation are adjusted in order to fit the empirical data;
- A function is fitted to the empirical data and then it is shifted to visually match the lower boundary of the empirical data;

Even if the described procedures may produce visually acceptable results, they are not easily or clearly reproducible since different results may be obtained working with the same set of empirical data (Guzzetti et al. 2008). To overcome this limitation, Guzzetti et al. (2007) proposed an objective methodology to define a threshold model: a curve of the form proposed by Caine (1980) is adopted and a probability approach is used to find the scale intercept (a, in Eq. 2) and the slope (b, in Eq. 2) for the power law curve representing the minimum threshold, by defining a Bernoulli probability of a data point occurring at a given value of rainfall intensity I and rainfall duration D.

Scales

I-D thresholds can be classified in global, regional or local, depending on the territorial extent to which they apply (Guzzetti et al. 2007). Global thresholds try to establish worldwide minimum level below which landslides do not happen. Examples have been proposed by Caine (1980), Innes (1983), Jibson (1989), Clarizia et al. (1996), Crosta and Frattini (2001), and Cannon and Gartner (2005). Regional thresholds are established for areas extending from a few to several thousand square kilometres with similar meteorological, climatic, and physiographic characteristics. Among others, examples have been proposed by Ceriani et al. (1992), Calcaterra et al. (2000), Corominas et al. (2005). Local thresholds are applicable to single landslides or to group of landslides in areas extending from a few to some hundreds of square kilometres. Global thresholds may be relevant where local or regional thresholds are not available, but may result in false positives, i.e., prediction of landslides that do not occur (Guzzetti et al. 2007). Different climatic regions have in fact different thresholds values due to adaptation of natural slopes to the climatic conditions which have prevail since the last glaciation (Nadim et al. 2009).

Normalized rainfall I-D thresholds

Crosta (1998) pointed out that regional and local thresholds perform reasonably well in the area where they were developed, but cannot be easily exported to neighbouring areas. Morphological, lithological, meteorological and climate variability is in fact not considered in the ID thresholds (Jakob and Weatherly 2003). To make comparable thresholds obtained for different areas or regions, the rainfall intensity values are normalized using empirical

measures of the local climatic conditions (Guzzetti et al. 2007). Two normalizations are usually performed (Guzzetti et al. 2008). The first is obtained dividing the event rainfall intensity by the mean annual precipitation (MAP) (e.g., Cannon 1988; Ceriani et al. 1992; Wieczorek et al. 2000; Aleotti et al. 2002; Bacchini and Zannoni 2003). The second normalization is performed by dividing the event rainfall intensity by the rainy-day normal (RDN), the ratio between the MAP and the average number of rainy days in a year (Wilson and Jayko 1997).

2.5 ANTECEDENT PRECIPITATION MODEL

Simple thresholds based on antecedent precipitation can be formulated according to the following equations:

$$A_n = \alpha_n \quad (3)$$

$$A_n = \alpha_n \quad \text{AND} \quad A_p = \alpha_p \quad (4)$$

$$1 + \alpha_1 A_n + \alpha_2 A_p = 0 \quad (5)$$

Where: A_n and A_p are the antecedent n-day and p-day precipitation, and α_n, α_p, α_1 and α_2 are constants of the model.

Considering that as long term antecedent precipitation increases, the short term antecedent precipitation required for triggering landslides should be reduced, the terms α in Equation 5 should have the same sign.

2.6 I-A-D MODEL

The I-A-D model has been proposed by Cepeda et al. (2009). The starting point for the formulation of the new model is the power law function first proposed by Caine (1980) as a threshold curve defining a lower bound to global observations of landslide-triggering storms:

$$I = \alpha D^{\beta} \quad (6)$$

Where:
I: mean threshold rainfall intensity (commonly expressed in mm/h)
D: duration (commonly expressed in hours)
β: slope of threshold curve in log-log space
α: numerical value of the mean rainfall intensity for D = 1 h. In order to ensure dimensional consistency for all values of D, α must have dimensions of $[L/T^{(\beta+1)}]$ or units of mm/h$^{(\beta+1)}$ for commonly used units of I and D.

Following a conventional definition of threshold, a rainfall event with a duration D, will be likely to trigger landslides if it has a mean intensity I such that $I \geq \alpha \cdot D^{\beta}$. The exponent β typically varies between -2.00 and -0.19 (Guzzetti et al. 2007). This function has been widely used by a number of authors for estimating thresholds at all geographic scales: global, regional and local. The most recent global update was presented by Guzzetti et al. (2008).

Based on analyses of numerical experiments by Rahardjo et al. (2007), Dirksen et al. (1992) and Terlien (1998) and on empirical observations by Crosta (1998), Cepeda et al. (2010)

proposed a conceptual framework that correlates regional or local empirical threshold curves with landslide susceptibility levels. This framework proposes that a threshold curve with the power law form of Eq. 6 may be associated with a set of parallel curves in the log-log Intensity-Duration space, and each member of this set defines a threshold for a particular range of landslide susceptibility levels. A brief description of the framework follows. If a set of m parallel threshold curves is defined in the log-log Intensity-Duration space (i.e, β is constant), such that $\alpha_1 < \alpha_2 < \ldots \alpha_i \ldots < \alpha_m$, then each curve corresponds to a susceptibility level ($Susc_1 > Susc_2 > \ldots Susc_i \ldots > Susc_m$) and the i-th threshold curve is defined as a lower bound to rainfall conditions that are more likely to trigger landslides on slopes having susceptibility levels higher than $Susc_i$.

Landslide susceptibility is controlled by slope geometry and the shear strength along the potential sliding surfaces. Shear strength in turn is dependent on the constitutive properties of the slope materials, their current state and the geometry and distribution of discontinuities. In soils, the current state is defined by the stress-strain history and the present void ratio and pore pressures. In fractured rock masses, water pressures in the fractures are a controlling factor since they reduce shear strength and may induce driving forces that reduce the factor of safety. Variations in pore pressures are associated with changes in groundwater conditions usually resulting from antecedent rainfall. As a consequence, a well-known connection has been established between antecedent rainfall and susceptibility levels. This connection has been addressed by many authors using physically-based approaches; see e.g. Edgers and Nadim (2004) and Nadim et al. (2009).

Based on the aforementioned conceptual framework (Cepeda et al. 2010) relating thresholds and susceptibility, the latter being controlled by antecedent precipitation, among other factors, the coefficient α in Eq. 6 can be formulated as follows:

$$\alpha = f(Susc) = g(\ldots, A_n, \ldots) \quad (7)$$

where, α and $Susc$ are defined as previously, A_n is the n-day or n-hour antecedent precipitation, generally measured in mm, and f and g are functions relating the listed variables. The ellipses that precede and follow A_n represent all the other factors controlling landslide susceptibility, which for the present purposes are not explicitly included in the formulation. It should be noted that

$$\frac{df}{dSusc} < 0 ; \frac{\partial g}{\partial A_n} < 0 \quad (8)$$

in agreement with the fact that a rise in antecedent precipitation is expected to increase pore pressures and landslide susceptibility, conversely reducing thresholds levels (i.e., reducing α).

Some empirical thresholds have incorporated antecedent rainfall using a power law form, such as the formulation by Tatizana et al. (1987):

$$I = 2603 A_{96h}^{-0.933} \quad (9)$$

or the thresholds for Nicaragua and El Salvador proposed by Heyerdahl et al. (2003):

$$I_C = 258 A_{96h}^{-0.32} \quad (10)$$

where I_C is the critical hourly rainfall in mm/h. Note that both formulations use 96-hour antecedent precipitation and that the exponent of the antecedent precipitation parameter is negative, fulfilling Eq. 8.

Considering that the left hand side in Eq. 9 and 10 refers to hourly intensity, both equations can be rewritten in the following form:

$$I = \alpha_1 A_n^{\alpha_2} (1)^{\beta} \quad 11$$

or, if attempting a generalised form for other values of duration:

$$I = \left[\alpha_1 A_n^{\alpha_2}\right] D^{\beta} \quad 12$$

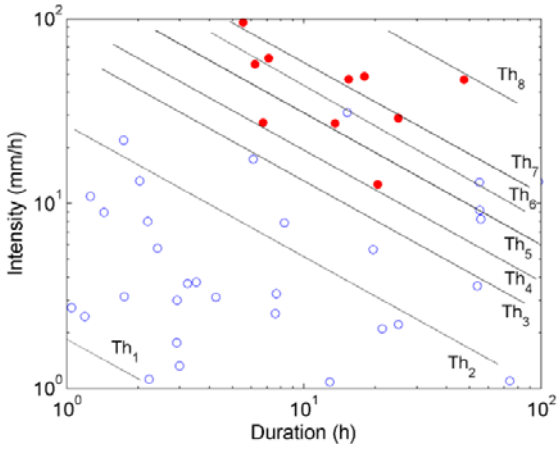
where α_1 and α_2 are constant parameters of the threshold, and the other values are defined as previously. Note that the expression in brackets corresponds to the coefficient α in Eq. 6 and to the function g in Eq. 7. Also notice that this formulation is lumping in the coefficient α_1 the contribution of all other factors controlling landslide susceptibility. Actually, Eq. 12 can be seen as a generalised form of Caine's power law (Eq. 6).

Since the duration of accumulation n for estimating the antecedent precipitation parameter A_n is not defined in Eq. 12, a procedure for selecting the optimum value of n needs to be established. It is desirable that this selection procedure may also be used for obtaining some or all the other parameters, namely α_1, α_2 and β. Traditionally, the selection of optimum threshold parameters has been done by visual inspection of available data or by performing a least-squares fit to all the observations. It is only recently that more systematic approaches have been applied such as those presented by Jakob and Weatherly (2003), Guzzetti et al. (2007; 2008) and Jakob (2009). Jakob and Weatherly (2003) and Jakob (2009) have performed discriminant function analyses (see e.g., Dillon and Goldstein 1984) using landslide-triggering and non-triggering storms. Guzzetti et al. (2007; 2008) have used a Bayesian approach (see e.g., Congdon 2006; Woodworth 2005) to estimate thresholds as lower bounds to rainfall events that triggered landslides. In this section, the procedure that will be used is the Receiver Operating Characteristics (ROC) graph, a technique for selecting classifiers based on their performance.

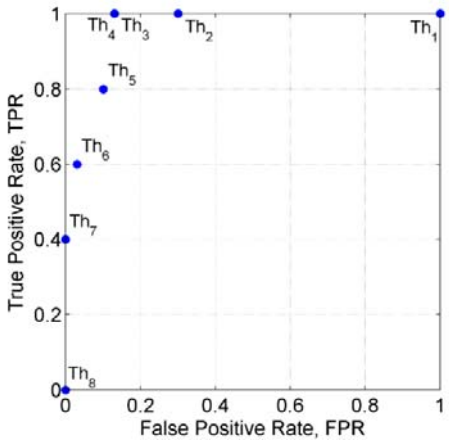
A ROC curve is a graphical representation of the performance of a *binary classifier* as a *discrimination threshold* is varied. The development and earliest usage of this technique dates back to World War II when it was used for the analysis of radar signals before being employed in signal detection theory. Fawcett (2006) presents a concise and comprehensive review of ROC analysis. In landslide hazard assessments, ROC curves have been extensively used in recent years for comparative evaluation of models of landslide susceptibility (see e.g., Frattini et al. 2004; and 2008; Chang et al. 2007). Since this technique has not yet been applied in the context of rainfall thresholds for landslide triggering, a brief description of the ROC analysis is presented in the following paragraphs employing rainfall observations to illustrate the relevant concepts. Most of the terms, variable names and symbol conventions are adopted from Fawcett (2006).

In order to keep the illustration as simple as possible, the following explanation will use a small hypothetical set of Intensity-Duration values from a rain gauge station representative of a landslide-prone area and a threshold formulation based on Eq. 6. The Intensity-Duration observations are presented in Figure 4a as circle markers. The open circles are non-triggering storms or *negative* observations in the context of the ROC analysis. The red filled circles are landslide-triggering storms or *positive* observations. The total number of negatives and positives are represented by the symbols N and P, and are 30 and 10, respectively in the present example. In this illustration, the *binary classifier* is a threshold function having the

form of Eq. 6 while β is kept constant. The coefficient α is the *discrimination threshold*. Figure 4a presents a set of 8 classifiers (black lines labeled as Th_1, Th_2, …) obtained using a constant value of β and 8 different values of the discrimination threshold α, such that $\alpha_1 < \alpha_2 < \ldots < \alpha_7 < \alpha_8$. Note that in this case, any set of discrimination thresholds using the same binary classifier will always produce a set of parallel curves.



(a)



(b)

Figure 4. Illustration of ROC analysis with rainfall observations. (a) Intensity-duration plot of non-triggering storms (open circles) and triggering storms (filled circles) and 8 different thresholds (black lines). (b) Discrete classifiers in the ROC space associated with the thresholds and rainfall observations in (a).

Some basic terms used in ROC analysis are now introduced in the context of rainfall thresholds using the curve Th_6 in Figure 4a as a classifier. The position of an observation relative to Th_6 defines a *hypothesized class* or prediction. A positive observation located at or above the threshold is called a *true positive* and below the threshold is termed a *false negative*. The corresponding terms for negative observations at or above the threshold and

below it are *false positives* and *true negatives*. The total number of true positives, false negatives, false positives and true negatives are represented by the variables *TP*, *FN*, *FP* and *TN*. Their respective values using Th_6 in Figure 4a are 6, 4, 1 and 29. These terms can be represented in a tabular form using a *confusion matrix* as shown in Table 4.

Table 4. Confusion matrix in the context of rainfall thresholds for triggering of landslides. Orange cells: landslide-triggering storms. Green cells: non-triggering storms. Colour filled cells: correct classifications. Hatched cells: incorrect classifications.

		True class (observations)	
		Positives (landslide-triggering storms)	Negatives (non-triggering storms)
Hypothesized class (predictions)	At or above threshold	True Positives ***TP***	False Positives ***FP*** (false alarms)
	Below threshold	False Negatives ***FN*** (missed events)	True Negatives ***TN***
	Σ	***P***	***N***

Two important metrics can be derived using the above variables:

$$TPR = \frac{TP}{P} \quad (13)$$

$$FPR = \frac{FP}{N} \quad (14)$$

Where *TPR* is the *true positive rate* and *FPR* is the *false positive rate*. In ROC analysis, *TPR* is also termed *hit rate*, *recall* or *sensitivity*. *FPR* is also termed *false alarm rate* and is also equal to 1 – *specificity*. *TPR* and *FPR* are 0.6 and 0.033 in Figure 4a concerning Th_6. These values can be represented in the ROC space as a *discrete classifier* plotting *FPR* on the horizontal axis and *TPR* on the vertical, see blue marker labelled Th_6 in Figure 4b. The other classifiers Th_1, Th_2, … Th_8 shown in Figure 4a are also plotted in Figure 4b.

Some important characteristics of the ROC space are illustrated using Figure 4b:

1. Most conservative classifiers. A discrete classifier that will unconditionally produce predictions at or above the threshold is plotted as (1,1) in the ROC space. This corresponds to threshold Th_1 in Figure 4a.

2. Least conservative classifiers. A discrete classifier that will unconditionally generate predictions below the threshold is plotted as (0,0) in the ROC space and corresponds to curve Th_8 in Figure 4a.

3. Perfect classification. A discrete classifier that does not produce any missed events or false alarms, i.e. $FP = FN = 0$, is plotted as (0,1) in the ROC space and produces a perfect classification. From an examination of Figure 4a, it can be seen that this set of observations and the present choice of binary classifier cannot produce a perfect classification.

4. A discrete classifier in the ROC space does not map a unique threshold. A point in the ROC space may represent several thresholds that map the same pair (*FPR*, *TPR*). See for example, thresholds Th_3 and Th_4 in Figure 4a and Figure 4b. Hence, some additional condition needs to be established to distinguish several thresholds mapping the same discrete classifier. A possible choice of criterion is using some metrics of the distance between the true positives or true negatives and the threshold. If such measure is applied to Figure 4a, Th_3 can be regarded as more conservative than Th_4, even though both thresholds map the same discrete classifier.

5. A set of monotonically changing discrimination thresholds, α, will produce a set of discrete classifiers which produces a monotonic curve termed *ROC curve*. Since rainfall thresholds are derived using a finite number of observations, the resulting ROC curve will always be a step function.

6. The area under a ROC curve, *AUC*, is a measure of the performance of a binary classifier. When a binary classifier maps a perfect discrete classifier (see item 3 above), *AUC* equals to 1, the maximum value of this measure. When two or more binary classifiers are compared based solely on their values of *AUC*, the best performance corresponds to the classifier with the largest *AUC*.

Applying the above concepts of ROC analysis to the threshold function in Eq. 12, the value of threshold $I_{n,k,i}$ for an *i*-th rainfall observation with duration, intensity and *n*-duration antecedent precipitation, D_i, I_i, A_{ni}, is given by:

$$I_{n,k,i} = \alpha_{1,k} A_{n_i}^{\alpha_{2,k}} D_i^{\beta_k} \qquad 15$$

where $\alpha_{1,k}$, $\alpha_{2,k}$, β_k and n are model parameters of the threshold. The parameters that bear the k subscript are varied while n is kept constant. Then, the total number of true positives and false positives, TP_n and FP_n, can be updated for each *i*-th rainfall observation as follows:

$$\text{If } I_i \geq I_{n,k,i}\text{, then } \begin{cases} \text{If the } i\text{-th event triggered at least one landslide, then} & TP_n = TP_n + 1 \\ \text{If the } i\text{-th event did not trigger any landslides, then} & FP_n = FP_n + 1 \end{cases} \qquad 16$$

It is possible to generalize the formulation in Eq. 16 based on the observation that short-, intermediate- as well as long-term antecedent precipitation may play a role in threshold values. This has been incorporated in other threshold formulations by relying on combinations of n (e.g., 3- and 15-day) rather than using a single value (see e.g., Chleborad et al. 2006 and Jakob 2009). Then, a more general form of Eq. 16 is:

$$\text{If } I_i \geq I_{n,k,i} \text{ and } I_i \geq I_{p,l,i} \text{ then} \begin{cases} \text{If the } i\text{-th event triggered at least one landslide, then} & TP_{np} = TP_{np} + 1 \\ \text{If the } i\text{-th event did not trigger any landslides, then} & FP_{np} = FP_{np} + 1 \end{cases} \qquad 17$$

Where $I_{p,l,i}$ is a threshold estimated for the *i*-th storm using *p*-hours or *p*-days antecedent precipitation and the *l*-set of parameters $\alpha_{1,l}$, $\alpha_{2,l}$, β_l. TP_{np} and FP_{np} are the total number of true positives and false positives resulting from this threshold definition. Note that in this extended

form, the definition of the *i*-th storm requires the antecedent *p*-hours or *p*-days precipitation A_{pi} in addition to the previous set of parameters D_i, I_i, A_{ni}.

Considering Eq. 17, a single combination of *n* and *p* will produce a ROC curve with an area AUC_{np}. With further consideration of item 6 (see list of characteristics of the ROC space), and testing several combinations of *n*- and *p*-hours (or –days) for the antecedent precipitation, then the optimum combination of the *n* and *p* parameters, *(np)opt* occurs when:

$$AUC_{(np)opt} = \max\left(AUC_{np}\right) \quad \textbf{18}$$

The choice of the *k*- and *l*-sets of parameters may be done in two steps: first, a discrete classifier is selected on the optimised *(np)opt* ROC curve; second, a norm is used to select the optimum *k*- and *l*-sets of parameters among all those parameters that map the same discrete classifier previously selected in the first step. The criterion for the first step will be a discrete classifier with no missed events (i.e., *TPR* = 1) and having the lowest amount of false alarms (i.e., minimum *FPR*):

$$mFPR_{(np)opt} = \min\left(FPR_{(np)opt,TPR=1}\right) \quad \textbf{19}$$

Where $FPR_{(np)opt,TPR=1}$ is a vector with all the *FPR* on the *(np)opt* ROC curve having *TPR* = 1. Then, the selected discrete classifier from the first step is mapped ($mFPR_{(np)opt}$, 1) on the ROC space. Now, the second step is necessary in order to resolve the non-unique mapping of thresholds on the ROC space (recall item 4 in the list of characteristics of the ROC space). Calculating a norm $r_{(np)opt,kl,j}$ for the distance between the *j*-th true positive observation and the thresholds $I_{n,k,j}$ and $I_{p,l,j}$:

$$r_{(np)opt,kl,j} = \sqrt{\left(I_{n,k,j} - I_j\right)^2 + \left(I_{p,l,j} - I_j\right)^2} \quad \textbf{20}$$

A norm $r_{(np)opt,kl}$ for all the true positives may be defined as:

$$r_{(np)opt,kl} = \sqrt{\sum_{j=1}^{TP} r_{(np)opt,kl,j}^{\ 2}} \quad \textbf{21}$$

Then, all *k*- and *l*-sets of parameters of the optimised thresholds are within a range bounded by the most conservative set *(kl)opt_cons* and the least conservative set *(kl)opt_lib*, which can be obtained by maximizing and minimizing a vector containing the norms calculated with Eq. 21:

$$r_{(np)opt,(kl)opt\_cons} = \max\left(r_{(np)opt,kl}\right) \quad \textbf{22}$$

$$r_{(np)opt,(kl)opt\_lib} = \min\left(r_{(np)opt,kl}\right) \quad \textbf{23}$$

Where $r_{(np)opt,(kl)opt\_cons}$ and $r_{(np)opt,(kl)opt\_lib}$ are the norms for the most and least conservative sets of *k* and *l* parameters.

Finally, the optimised threshold equations are:

$$I = \alpha_{1,k} A_n^{\alpha_{2,k}} D^{\beta_k} \quad \textit{and} \quad I = \alpha_{1,l} A_p^{\alpha_{2,l}} D^{\beta_l} \quad \textbf{24}$$

Where *n*, *p*, and the *k*- and *l*-sets of parameters have been obtained by the aforementioned optimisation procedure.

2.7 FLaIR MODEL

In empirical models it is possible to identify a mobility function *Y(t)*, that is a generic function of the antecedent rainfall that can be correlated to rainfall occurrence.

More precisely if $P(E_t)$ is the probability of occurrence of a landslide at time *t*, and if we assume that $P(E_t)$ only depends on *Y(t)*, we can express:

$$P[E_t] = \begin{cases} 0 & if \quad Y(t) < Y_1 \\ g[Y(t)] & if \quad Y_1 \le Y(t) \le Y_2 \\ 1 & if \quad Y(t) > Y_2 \end{cases} \qquad 25$$

where *g(.)* is a not decreasing generic function, which can assume the values between [0;1] in the interval $[Y_1; Y_2]$, Y_1 is the value of *Y(t)* below which the mobilization is impossible, and Y_2 is the value of *Y(t)* over which the mobilization is certain.

By assuming $Y_1=Y_2=Y_{cr}$, the threshold scheme occurs which identifies a threshold value Y_{cr} of a mobility function *Y(t)* which separates the conditions "impossible mobilization" and "certain mobilization", that is:

$$\begin{cases} P[E_t] = 0 \quad if \; Y(t) < Y_{cr} \\ P[E_t] = 1 \quad if \; Y(t) \ge Y_{cr} \end{cases} \qquad 26$$

The threshold approach is the most widely approach used in studies of rainfall-induced landslides because function identification and parameters calibration in Eq. 25 are very difficult owing to the lack of experimental data.

Sirangelo and Versace (1992) proposed the hydrological model FLaIR (Forecasting of Landslides Induced by Rainfall), which considers the mobility function as a convolution between the rainfall infiltration rate $I(.)$ and a filter function $\psi(.)$:

$$Y_F(t) = \int_0^t \psi(t-\tau) I(\tau) d\tau \qquad 27$$

where *F* subscript is for FLaIR.

The infiltration rate *I(u)* is assumed proportional to the rainfall intensity *P(u),* according to the following simple relationship:

$$P_*(u)=\begin{cases}P(u) & when \quad P(u)\le P_0\\ P_0 & when \quad P(u)>P_0\end{cases} \qquad 28$$

where P_0 depends on soil characteristic, and c_0 is a factor of proportionality. Because the mobility function is defined up to an arbitrary multiplicative factor, it is possible to choose $c_0c=1$ so that:

$$Y_F(t)=\int_0^t \psi(t-u)P_*(u)du \qquad 29$$

The function $\psi(.)$ is typical for each case study and plays a central role in mobility function evaluation. It can assume different expressions (Iiritano et al. 1998), like rectangular:

$$\psi(t)=\begin{cases}1/t_0 & \text{if} \quad 0<t\le t_0\\ 0 & \text{elsewhere}\end{cases} \qquad 30$$

exponential:

$$\psi(t)=\frac{1}{k}e^{-\frac{t}{k}} \qquad t\ge 0, k>0 \qquad 31$$

gamma:

$$\psi(t)=\frac{b^a}{\Gamma(a)}t^{a-1}e^{-bt} \qquad t\ge 0, a>0, b>0 \qquad 32$$

power:

$$\psi(t)=mt^{-q} \qquad 0<t<T, m>0, 0<q<1 \qquad 33$$

combination of two exponential functions:

$$\psi(t)=\omega\beta_1\exp(-\beta_1 t)+(1-\omega)(-\beta_2 t) \qquad t\ge 0, \beta_1\ge\beta_2>0, 0\le\omega\le 1 \qquad 34$$

Starting from the historical information, that is landslide dates and antecedent rainfalls, it is possible to define the transfer function, the mobility function and its threshold values (Eq. 26). Two techniques have been proposed to estimate the parameters, called *'ranking'* and *'crossing'* (Sirangelo and Versace 1996); and Versace et al. 2000).

Using the 'ranking' technique, parameters θ can be estimated by evaluating the mobility function $Y_F(t,\underline{\theta})$ on the range of variability of the parameters and adopting the values that allow $Y_F(t,\underline{\theta})$ to assume its k highest values in correspondence with the k known historical movements of the landslide.

For each admissible parameter, it is possible to detect a variation interval for the mobility function between a lower limit function value $f_L(.)$ and an upper limit function value $f_U(.)$. These represent, respectively, the highest value which did not produce any movement and the lowest value for which movement occurred.

The second technique, called '*crossing'*, has been proposed to reduce the difference between the upper and lower limit functions in the case of landslide with only one movement. The upper limit function value $f_U(.)$ is the lower value of $Y_F(t,\underline{\theta})$ that is associated to a historical mobility, while lower limit function value $f_L(.)$ is the upper value that $Y_F(t,\underline{\theta})$ assumes in the immediate proximity of landslide event. Moreover the parameters will be considered admissible if they satisfy the 'ranking' condition and produce mobility functions locally no decreasing before historical landslides.

The FLaIR model can be assumed as a general framework for the large part of empirical relationship between landslide and antecedent rainfall (Capparelli et al. 2009; Capparelli and Versace 2010).

Moreover FLaIR exhibits many advantages if compared with other empirical approaches. Firstly, it is able to consider the real pattern of rainfall input, as it gives different values of mobility function for rainfall with the same average intensity but with different hyetographs. In such cases other models, like I-D, give, instead, the same value of mobility function.

Comparison between current rainfall values and critical ones also is easier by considering FLaIR because comparison is just between two values ($Y_F(t)$ and $Y_{F,cr}$) and not between two curves.

Besides, depending on the shape of filter function $\psi(.)$, in FLaIR model, the weight of a short and intense rainfall pulse changes with time, while it remains the same if a cumulative rainfall is considered, like other empirical models do. So in FLaIR model effect of such a pulse is not anymore "persistent".

Finally, due to different shapes of the filter function $\psi(.)$, the FLaIR model is much more flexible than other models and can represent different types of slope responses with respect to rainfall inputs: from very fast responses to slower ones.

By using FLaIR model, different applications can be carried out. It can be used for individual (punctual) cases analyzing historical cases of mobilisation.

The many applications running on several selected case studies in Italy showed that the transfer function shapes appear strictly linked to the landslide dimensions (Sirangelo et al. 2003). The triggering mechanisms for the shallow landslides are usually caused by very intense rainfall accumulated over short time intervals; deep movements depend on rainfall accumulated over long time intervals.

It is possible to join the model with a rainfall forecasting model, as meteorological or stochastic generators which provide a probabilistic prediction of future rainfalls. The latter represent the new input for the FLaIR model which is so able to evaluate the probability that a mobility function exceeds, in a future time *t,* a fixed critical value $Y_F(t)$, providing useful information about the evolving conditions in advance. Capparelli and Tiranti (2010) describe an example concerning landslide thresholds in the Western Alps sector of the Piemonte region (Northern Italy) where debris flows are the predominant type of landslide. By linking the mobility function to meteorological forecasts, the model was used for developing an early warning system for real time forecasting of slope hazard. Versace et al. (2003) and Giorgi et

al. (2009) describe integration of FLaIR with two different stochastic modules: PRAISE model (Prediction of Rainfall Amount Inside Storm Events) proposed by Sirangelo et al. (2007) and a DRIP model (Disaggregated Rectangular Intensity Pulse) proposed by Heneker et al. (2001). The outputs are represented by a probability, estimated at time τ, that the forecasted value of mobility function, at time $t > \tau$, will exceed the critical value.

If suitable adequate geotechnical monitoring instrumentation (inclinometers, piezometers, extensometers, etc) are available in the landslide area, other applications of the FLaIR model are possible. The mobility function, in fact, can be correlated to field data concerning movement of landslide bodies, or change in pore pressure (Capparelli et al. 2009).

The most important references for the FLaIR model are: Sirangelo and Versace (1992), Capparelli et al. (2009) and Versace and Capparelli (2008).

2.8 NEURAL NETWORKS

Neural networks for the prediction of the behaviour of slow-moving landslides

Artificial neural networks (ANN) belong to the category of statistical methods. The method is characterized by its capacity for automatic learning, i.e. the ability to learn from their environment and to improve their performance. The first applications in the domain of soils or structures date back to the 90s (e.g., Goh 1994; Lee and Lee 1996; Aleotti and Chowdhury 1999). A detailed description of their use for the prediction of slow displacement movements of slopes can be found in Mayoraz and Vulliet (2002); the principal concepts presented by the authors are recalled in the following.

Theoretical bases

An *artificial neural network (ANN)* is a computational tool composed of a large number of very simple processors: the artificial neurons. Each neuron can perform elementary calculations, as weighted sum, on *input parameters* and provide an *output parameter*. The network set-up allows exchanging information through a network of dense connections (Figure 5).

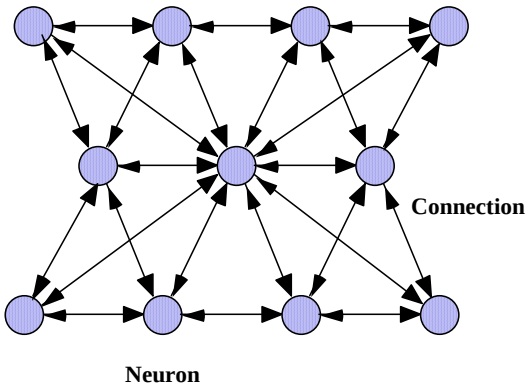


Figure 5 Artificial neural network.

The *perceptron* is the type of neural network most commonly used for tasks of classification and function approximation (Lippmann 1987). The simplest perceptron is composed of a

single artificial neuron of the McCulloch and Pitts type (McCulloch and Pitts 1943) (Figure 6), having N input signals and one output.

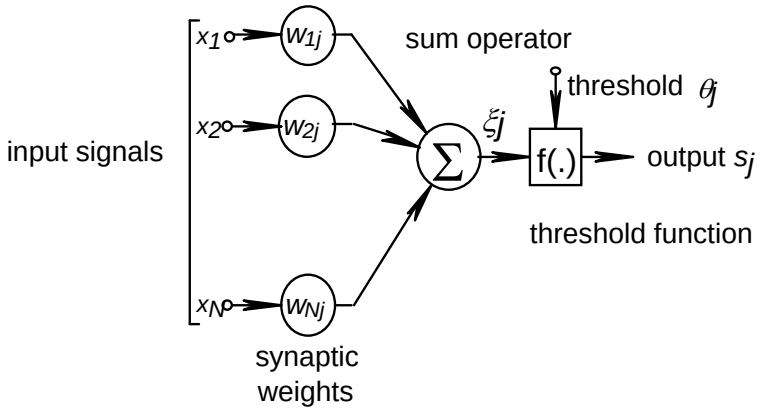


Figure 6 **Basic artificial neuron *j* of the McCulloch and Pitts type (from Mayoraz and Vulliet 2002).**

The output s_j of this kind of neuron has the form:

$$s_j = f(\xi_j - \theta_j) = f(\sum_{i=1}^{N} w_{ij} x_j - \theta_j) \qquad 35$$

where w_{1j}, w_{2j}, ...,w_{nj} are the transmission (or synaptic) weights, which are the real coefficients affected to the input signals $\{x_1, x_2, ..., x_n\}$, ξ_j is the internal state of the neuron, f is a function affected by a threshold θ_j. Various forms can be adopted for the threshold function f (Figure 7). These are continuously increasing limited function such as simple step functions, sigmoidal forms or linear piecewise functions. The maximum and the minimum of the function correspond to the upper and lower states of the automate (generally the output varies between 0 and 1, or between –1 and 1).

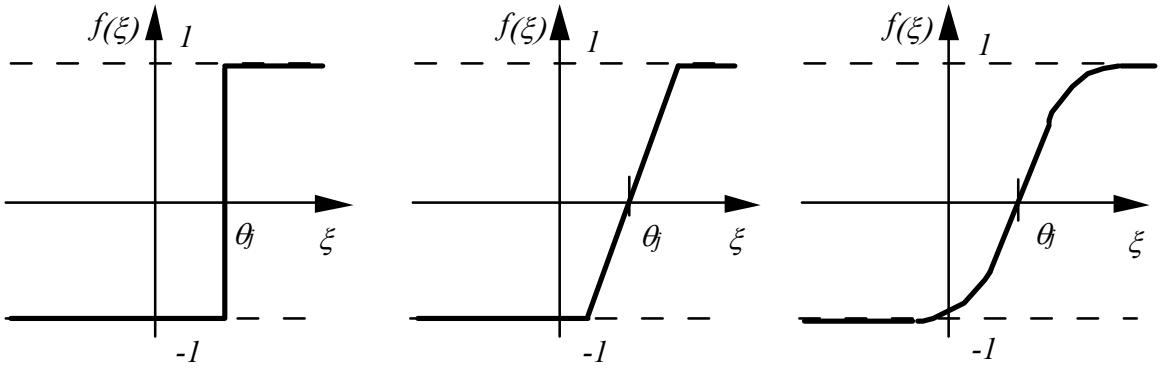


Figure 7 **Different threshold function forms : step, linear and sigmoid (from Mayoraz and Vulliet 2002).**

Equation 35 shows that such a neuron performs a linear discrimination in the vector space of its inputs. The internal weight w_{ij} of the neuron and the threshold θ_j determine the equation of

the separating hyperplane. When the problem is not linearly separable, a *multilayer perceptron* can be created combining the basic units by cascading in order to form networks of several neurons, organised in layers (Figure 8). A *transfer function* T is constructed automatically for the neural network. If the output only has one component, this function is of the following form:

$$T(\underline{X}_k) = T(x_{1(k)}, x_{2(k)}, \ldots, x_{N(k)}) = \sum_{i=1}^{M} v_i f\left(\sum_{j=1}^{N} w_{ij} x_{j(k)} - \theta_i\right) - \mu \qquad 36$$

where $x_{j(k)}$ are the components of the k^{th} input vector, N and M are the number of inputs and the number of neurons in the intermediate layer, respectively; w_{ij} are the synaptic weights between the inputs and the neurons in the intermediate layer, v_i the synaptic weights between the neurons in the hidden layer and the output neuron; θ_j are the thresholds of the neurons in the hidden layer, μ is the threshold of the output neurons and f is the sigmoid function given by:

$$f(\xi) = \frac{1}{1 + e^{-\xi}} \qquad 37$$

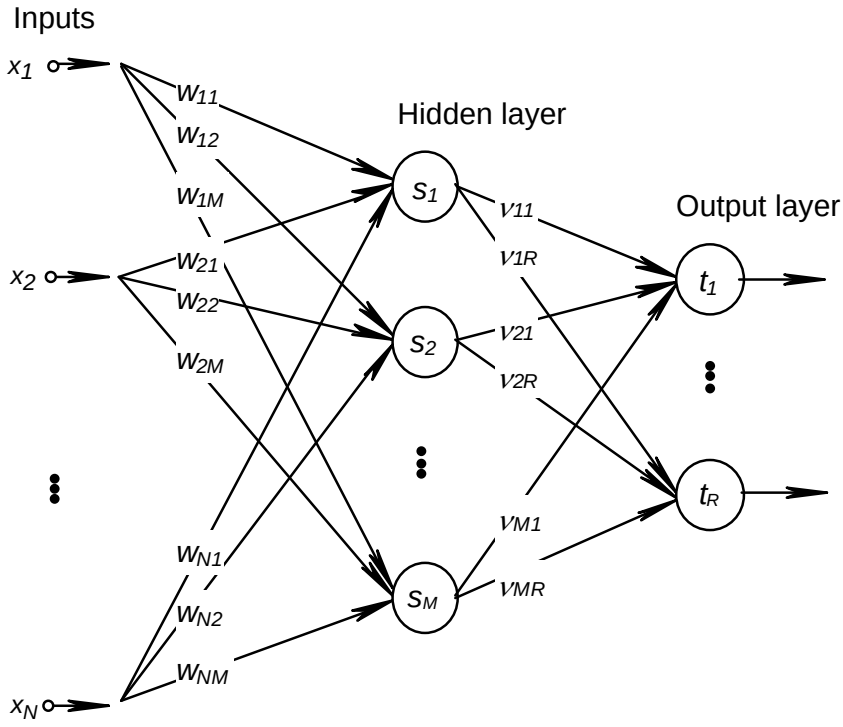


Figure 8 **Multilayer perceptron with N components inputs, 1 intermediate layer of M neurons and 1 output layer of R neuron (from Mayoraz and Vulliet 2002).**

The learning process

ANN can learn from their environment and to improve their performance. The use of such a system generally includes two distinct phases: a configuration phase, also called the *training* phase, and an application phase called the *test* phase during which the outputs, unknown to the network until that moment, are presented to it. *Supervised* learning is an *iterative* process during which the weights of the connections are progressively adjusted. During this phase, some examples of input vectors are presented to the system, which, through the use of a learning algorithm, minimise a predefined criterion, generally the average global quadratic error E at the output of the network, defined by:

$$E(\underline{W}) = \frac{1}{2P}\sum_{k=1}^{P}\sum_{i=1}^{M}(t_i(\underline{X}_k) - o_i(\underline{X}_k))^2 = \frac{1}{2P}\sum_{k=1}^{P}(T(\underline{X}_k) - o(\underline{X}_k))^2 \qquad 38$$

$o(\underline{X}_k)$ is the desired output of the network and $T(\underline{X}_k)$, the output calculated by the transfer function. t_i is the output of the i^{th} neuron of the ouput layer and P is the number of input vectors. Learning in a multilayer perceptron may be seen as a process of regression of the non-polynomial function T.

A generalisation of the learning algorithm for simple perceptrons (one layer) is used for multilayer perceptrons. This is based on the principle referred to as *backpropagation* (LeCun 1985). Let $\{\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_P\}$ be a set of P points in an N dimensional space and the set $\{o(\underline{X}_k)\}, k = \{1,2,\ldots,P\}$ its corresponding outputs. The partial derivatives of error with respect to the weight of the last layer are calculated directly. The technique of back-propagation is a process which permits the *back* calculation of the partial derivatives of error e_j due to the neuron j with respect to the weight of the neurons connected to j of the preceding layers and permits the progressive adjustment of these weights according to the general relation:

$$\Delta w_{ij} = -\frac{\eta}{P}\frac{\partial E}{\partial w_{ij}} = -\frac{\eta}{P}\sum_{k=1}^{P} J_{(k)}{}^{T}(w_{ij})\, e_{j(k)}(w_{ij})$$

where

$$e_{j(k)}(w_{ij}) = s_{j(k)}(w_{ij}) - o_{j(k)}(w_{ij}) \qquad 39$$

$$J_{(k)}^{T} = \frac{\partial e_{j(k)}}{\partial w_{ij}}$$

η is the learning parameter, s_j is the transfer function of the neuron j (Eq. 35), and J is the Jacobian error matrix with respect to the weights. The index *(k)* indicates that these values concern the k^{th} input vector. In the case where the threshold function is that of Eq. 37, it may be noticed that its derivation respects the following property:

$$\frac{\partial f}{\partial \xi} = f(\xi)(1 - f(\xi)) \qquad 40$$

When the number of layers and the number of examples are large, this method, which is a gradient descent method, can reach a local minimum in weight space that is rather far away from the desired absolute minimum. It is then necessary to restart the training from other initial weights, randomly chosen. The use of complex perceptrons is therefore limited by calculation difficulties. Thus, pre-processing of data is often necessary in real applications to extract the pertinent attributes rather than presenting the raw data to the neural network.

Prediction of slope behaviour

The presented method has been used to predict the behaviour of slow-moving landslides (Mayoraz and Vulliet 2002). Possible applications are: (i) prediction of the pore water pressure as a function of the rainfall; (ii) prediction of the velocity as a function of the variation in pore water pressure; (iii) prediction of displacement rates from rainfall data.

2.9 EMPRICAL DYNAMIC MODELS FOR THE PREDICTION OF LANDSLIDES

This section describes the use of empirical dynamic models for the analysis of the behavior of slow-moving landslides. General information on building dynamic models is available in specific books; in the following, the main concepts presented by Ljung (2002) are briefly recalled and adapted to the case of landslide analysis.

Dynamic models aim to describe relationship between measured quantities. Input and output are distinguished. The output is partly determined by the inputs. With reference to the behavior of slow-moving landslides the movement rate is usually the model output; in this sense a monovariate time series is obtained, that is a time series with a single output variable. A single-output system with input u and output y is called linear if it is possible to describe it by a model that takes the form:

$$y(t) = G\left(q^{-1}\right)u(t) + H\left(q^{-1}\right)e(t) \qquad 41$$

where G and H are transfer functions in the time delay operator, q^{-1}; $e(t)$ is a white noise signal that is independent of past inputs; $u(t)$ is an exogenous signal. For the present work, rainfall evolution is assumed as the input signal. Input can generally include several other measurements evolving with time such as pore water pressure, temperature, relative humidity. All quantities that are not explicitly considered in the model and that affect the landslide behavior (i.e. the model output) are indicated as disturbance.

The delay operator works on a signal in the following way:

$$q^{-d}x(t) = x(t-d) \qquad 42$$

where d is a multiple of the sampling period of the data.

All the common models used can be summarized by the general equation:

$$A\left(q^{-1}\right)y(t) = \frac{B\left(q^{-1}\right)}{F\left(q^{-1}\right)}u(t) + \frac{C\left(q^{-1}\right)}{D\left(q^{-1}\right)}e(t) \qquad 43$$

where

$$A\left(q^{-1}\right) = 1 + a_1 q^{-1} + \ldots a_n q^{-n}$$
$$B\left(q^{-1}\right) = b_0 + b_1 q^{-1} + \ldots b_m q^{-m}$$
$$C\left(q^{-1}\right) = 1 + c_1 q^{-1} + \ldots c_k q^{-k}$$

$$D\left(q^{-1}\right)=1+d_1q^{-1}+\ldots d_lq^{-l}$$
$$F\left(q^{-1}\right)=1+f_1q^{-1}+\ldots f_rq^{-r}$$

According to the different values of the polynomials in Eq. 43, the following models are defined: Finite Impulse Response (FIR) model (A=F=C=D=1); *Box-Jenkis* (BJ) model (A=1); AutoRegressive Moving Average with eXogenous inputs (ARMAX) model (F=D=1); *Output-Error* (OE) model (A=C=D=1) and AutoRegressive with eXogenous inputs (ARX) model (F=C=D=1).

The predictor associated with Eq. 43 can be expressed in "pseudo-linear" regression form as (Ljung and Söderström 1983):

$$\hat{y}\left(t\middle|\theta\right)=\theta^T\varphi\left(t,\theta\right) \quad (44)$$

where θ is the *parameter vector* and φ is the *regression vector*.
For example, in the ARX model the *regression* and *parameter vectors* are defined respectively by:

$$\varphi\left(t\right)=\left[y\left(t-1\right)\ldots y\left(t-n_a\right),u\left(t-n_k\right)\ldots u\left(t-n_k-n_b\right)\right]^T \quad (45)$$

$$\theta=\left[-a_1,\ldots-a_{n_a},b_0,\ldots b_{n_b}\right]^T \quad (46)$$

In an explicit form, the predicted value $\hat{y}\left(t\right)$ of the output variable y can also be written as:

$$\hat{y}\left(t\right)=a_1y\left(t-1\right)+\ldots+a_{n_a}y\left(t-n_a\right)+b_0u\left(t-n_k\right)+\ldots+b_{n_b}u\left(t-n_k-n_b\right) \quad (47)$$

or

$$\hat{y}\left(t\right)+a_1y\left(t-1\right)+\ldots+a_{n_a}y\left(t-n_a\right)=b_0u\left(t-n_k\right)+\ldots+b_{n_b}u\left(t-n_k-n_b\right) \quad (48)$$

Which is the same as Eq. 47, within the signs of the a coefficients. The coefficients n_a and n_b in these expressions are also called *embedding parameters*, the coefficient n_k is the *delay parameter*.

The equation used for an ARMAX model would instead be:

$$\hat{y}\left(t\right)+a_1y\left(t-1\right)+\ldots+a_{n_a}y\left(t-n_a\right)=b_0u\left(t-n_k\right)+\ldots+b_{n_b}u\left(t-n_k-n_b\right)+c_1e\left(t-1\right)+\ldots+c_{n_c}e\left(t-n_c\right) \quad (49)$$

In the case of multiple input time series, a value of n_b and n_k has to be set for each of input channels. Generally speaking, they can be different for each input channel, even though the most common case is to have the same parameters for each of the input channels, especially if they all represent the same kind of variable.

A second approach can be attempted based on the nonlinear versions of the time series forecasting models mentioned above.

Nonlinear counterparts to the linear time series forecasting model structures are obtained by equations of the type:

$$y(t) = g\left[\varphi(t,\theta),\theta\right] + e(t) \qquad 50$$

or on *predictor* form:

$$\hat{y}(t|\theta) = g\left[\varphi(t,\theta),\theta\right] \qquad 51$$

where $\varphi(t,\theta)$ is again the *regression* vector, while θ is the vector containing the adjustable parameters in the neural network known as *weights*. g is the function realized by the neural network and it is assumed to have a feed-forward structure. Depending on the choice of the regression vector, different nonlinear model structures can be obtained. If the regression vector is selected as for ARX models, the model structure is called NNARX (the acronym of Neural Network ARX). Likewise, NNFIR, NNARMAX, NNOE structures also exist.

The structures of NNARX and NNARMAX models are depicted in Figure 9.

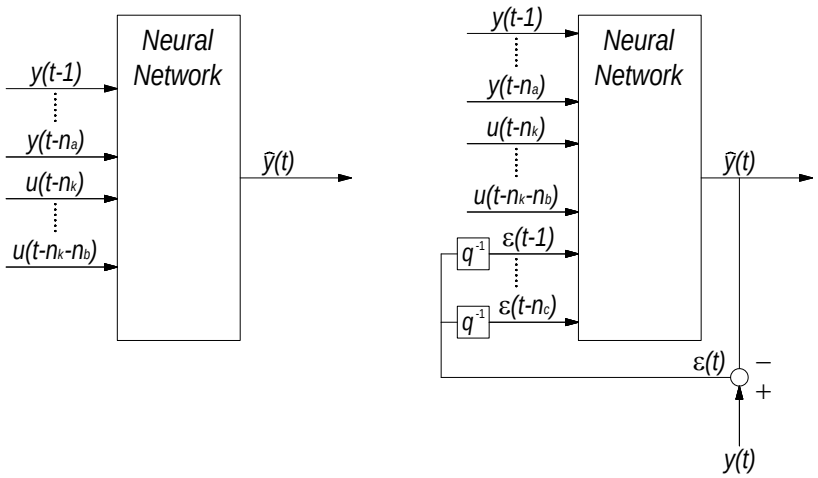


Figure 9. The NNARX (left hand side) and NNARMAX (right hand side) model structures.

2.10 CONSIDERATION OF UNCERTAINTY IN DATE AND TIME OF OCCURRENCE OF LANDSLIDE EVENTS FOR ESTIMATION OF THRESHOLDS.

Landslide inventories suffer frequently of lack of information or uncertainty in the location and timing (date and time) of events. For example, the Norwegian database of landslide includes more than one thousand events with an uncertainty in time of occurrence ranging between a few hours to 64 days. The traditional approach to address this uncertainty has been to discard all events above a specified uncertainty, and use the remaining events without accounting for their time uncertainty. In the present report, the following procedure is proposed for incorporating uncertainty in time (or date) of occurrence in the estimation of thresholds:

a. For each landslide event in the inventory, produce a symmetrical probability density function (PDF) distributed in time and bracketed by the value of uncertainty given in the inventory. The PDF can be uniform, triangular, or other form, but its area must be equal to 1.

b. Assuming that the landslide events are not mutually exclusive, apply the addition rule for estimating the union of the PDFs of all events.

c. Classify the resulting distribution according to the level of total time-uncertainty (e.g., low, medium, and high).

d. Estimate thresholds for the different levels of time-uncertainty.

The above procedure allows to incorporate all events in an inventory, regardless of their uncertainty in the time (or date) of occurrence. The procedure has significance when it is assumed that the time uncertainty is inversely proportional to the magnitude of the landslide.

3 EVALUATION CRITERIA

In order to test model performance, a set of evaluation criteria have been defined against which the models can be tuned.

3.1 PENALTY WEIGHTS FOR FALSELY PREDICTED ALARM STAGES

False predictions of an alarm stage by the model will result in a penalty. The size of the penalty increases with increasing potential severity of the false alarm. For example, a predicted alarm stage Yellow while the real stage was Orange is potentially more severe to a community than a predicted alarm stage Orange while the real stage was Yellow. An example of penalty weights given to falsely predicted alarm stages is given in Table 5.

Table 5. Example of weigths due to differences between observed and predicted alarm stage.

		Observed			
		Green	Yellow	Orange	Red
Predicted	Green	0	2	4	6
	Yellow	1	0	2	4
	Orange	2	1	0	2
	Red	3	2	1	0

Since not all datasets in the database available for testing have these alarm stages, penalty weights differ between the datasets.

3.2 PERFORMANCE OF CLASSIFICATION

The relative performance of the classification of several threshold models can be evaluated using one or more of the following criteria:

a. Confusion matrix. See Table 4.
b. ROC curve. The ROC curve is obtained by evaluating the threshold model with several sets of parameters. See section 2.5.
c. Area under the ROC curve (AUC). The maximum AUC is 1 for a model with a perfect classification. When comparing several models, the ones with highest AUC have the best performance.
d. Minimum distance from the ROC curve to the point of perfect classification. When comparing the ROC curves for several models, the minimum distance to the point of perfect classification (FPR = 0, TPR = 1) provides an indication of the model performance. The lowest distances correspond to models with best performance.
e. Misclassification error rate. This parameter is obtained as the ratio between the misclassified observations (missed events plus false alarms) divided by the total number of observations.

3.3 FLaIR MODEL

The FLaIR model can be evaluated by using ROC Analysis (Receiver Operating Characteristics, see e.g., Fawcett 2006; Frattini et al. 2004; and 2008; Chang et al. 2007), for detecting the success/failure to predict landslide and non-landslide events.

If the model outputs and the time series of the observed landslides are known, there are four possible outcomes:

a. If a landslide occurs and it is predicted by the model, a *true positive TP*, also named *correct alarm CA*, is counted;
b. If a landslide occurs and it is not predicted by the model, a *false negative FN*, also named *missed alarm MA*, is counted;
c. If a landslide does not occur, but it is predicted by the model, a *false positive FP*, also named *false alarm FA*, is counted;
d. If a landslide does not occur, and it is not predicted by the model, a *true negative TN* is counted.

Starting from these definitions, a two-by-two contingency matrix can be constructed (Table 6), representing the counting of the outcomes. In details, the numbers along the major diagonal represent the correct decisions, while the other diagonal provides the errors between the various classes.

Table 6. Contigency matrix in the context of landslides prediction.

		True class (**Observation**)	
		Landslide occurrence	**No landslide occurrence**
Predicted class (**Prediction**)	**Predicted landslide**	Correct Alarms *CA*	False Alarms *FA*
	No predicted landslide	Missed Alarms *MA*	True Negatives *TN*

This matrix forms the basis for many common metrics, useful for model performance evaluation; it is possible to define:

1. The *Hit Rate HR*, estimated as:

$$HR = \frac{Correct\ Alarms}{Correct\ Alarms + Missed\ Alarms} = \frac{CA}{CA + MA} \quad (52)$$

2. The *False Alarm Rate FAR*, estimated as:

$$FAR = \frac{False\ Alarms}{True\ Negatives + False\ Alarms} = \frac{FA}{TN + FA} \quad (53)$$

The range of both metrics is [0 ; 1] and the graphical representation is the two-dimensional ROC graphs (Figure 10), in which *HR* is plotted on the Y axis and *FAR* is plotted on X axis.

Analyzing the figure, it is possible to classify a model as:

a. *Most Conservative* if predictions are unconditionally produced at or above the threshold. It is plotted as (1,1) in the ROC graphs and it corresponds to a FLaIR model for which a threshold equal to 0 is assumed.
b. *Least Conservative* if predictions are unconditionally produced below the threshold. It is plotted as (0,0) in the ROC graphs and it corresponds to a FLaIR model for which a infinitive threshold value is assumed.
c. *Perfect Classifier* if it produces neither missed alarms nor false alarms and it is plotted as (0,1) in the ROC graphs.

For a fixed parameter vector $\underline{\theta}$ of the selected transfer function $\psi(.)$ of the FLaIR model, it is possible to produce points in the ROC graphs, associated with different threshold values, which describe a monotonic curve termed ROC curve. The area under a ROC curve, named *AUC* (Fawcett 2006), is a measure of the performance. When a model is a Perfect Classifier, *AUC* equals 1.

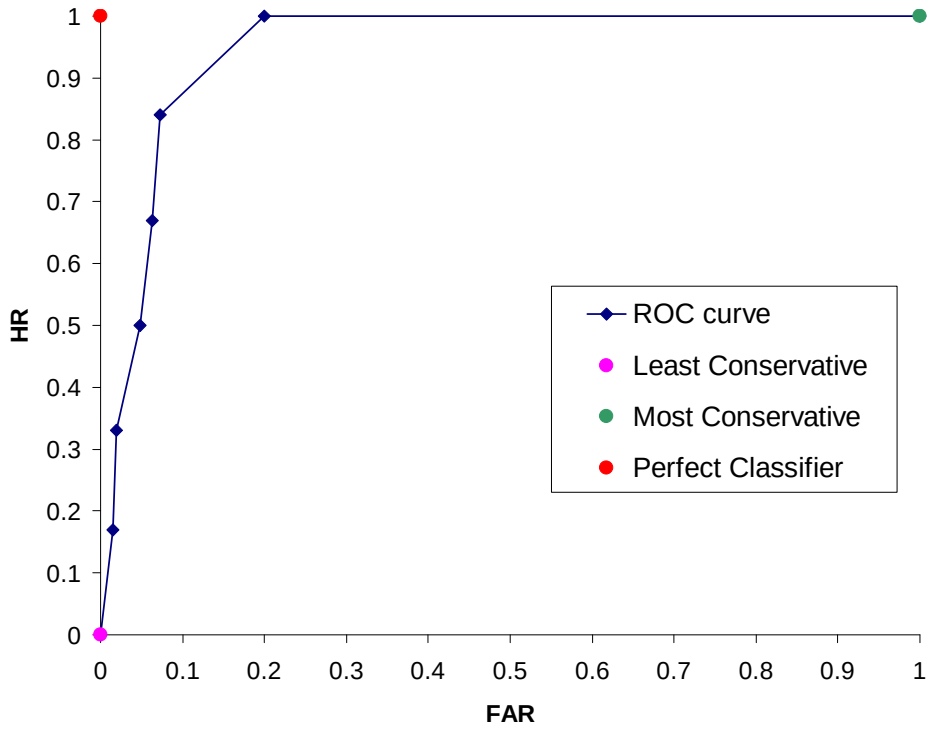


Figure 10. ROC space.

Then, the optimum parameters vector $\underline{\theta}$ of the FLaIR model corresponds to the maximum *AUC* value. The proposed 'ranking' and 'crossing' techniques (see section 2.7), tend to provide ROC curve with *AUC*=1. In fact, if the mobility function $Y_F(t,\underline{\theta})$ assumes its k highest values in correspondence with the k known historical landslide movements, and the upper limit $f_U(.)$ is considered as threshold value $Y_{F,cr}$, *FA* and *MA* are equal to 0 and, then, *HR = 1* and *FAR = 0*.

If two or more vectors $\underline{\theta}$ produce the same maximum value of *AUC*, the optimum will be characterized by the lower amount of false alarms (i.e., minimum FAR) in correspondence to the threshold for which there are no missed alarms (i.e., *HR*=1). If the same minimum *FAR* value is obtained, a second step is necessary, which consists in the evaluation of:

$$r = f_U - f_L \qquad 54$$

between the upper limit value $f_U(.)$ and the lower limit value $f_L(.)$ for different values of $\underline{\theta}$ (see section 2.7). The value of $\underline{\theta}$ which determines the maximum value of r is then selected as model parameter. This estimation allows the largest possible gap between the mobility function $Y_F(.)$ values related to no movement and those related to movement.

If a large number *k* of observed landslides for a study area is available, it is possible to divide the sample interval in two parts: k_1 and k_2. Model calibration is carried out by using the subset k_1, and the performance in ROC space is analyzed by considering the other interval.

3.4 EMPRICAL DYNAMIC MODELS FOR THE PREDICTION OF LANDSLIDES

A standard way to assess the model performances is to split the data set into a *training set* (which is used to determine the parameters of the model, that is the weights of the connections between neurons in the case of the NN models, and the parameter vector θ in the case of the linear model) and a *validation set* (which is used to compare the result of the prediction obtained with the model and the observed values recorded in the time series). Both the training and the validation sets consist of a series of patterns of input-output couples.

In the analyzed case study, each input-output couple is made up of the value of the cumulated horizontal displacement recorded at the day *x,* as the output, and the set of the rainfall values recorded in the 28 days before the day *x*, as the input. In order to evaluate the predictive performance of the model, it is important to use a set of sample patterns that were not used in the phase of training (or calibration) of the model.

If the model performs well with the same data used for its training, but has a poor predictive performance with the data belonging to the validation set, it is said to have a poor *generalization performance*.

4 DATASETS USED

This chapter gives a brief description of the data used to evaluate the models described in chapter 2.

4.1 LA FRASSE (SWITZERLAND)

The La Frasse landslide is located in Switzerland, some 20 km East of Lake Geneva (Figure 11). The landslide can be described as a major natural instability, both in terms of spatial extension and of the economical risks it generates. It intercepts over 1km the RC705, a cantonal road of strategic importance for the development of tourism. In fact, the road ensures the access to several ski resorts. This road is regularly damaged due to the recurrent crises of the lower part of the landslide such as in 1966, 1981 and 1994 where movements of up to 7m, 4m and 2m respectively were recorded. The motion of the La Frasse landslide furthermore induces the risk of the damming of the downhill Grande Eau River with significant potential damages for up to the 8'100 inhabitants of the city of Aigle.

Several studies have been carried out on the landslide, with the main purpose to analyse its behaviour and to propose mitigation measures (Lugeon et al. 1922; DUTI (Détection and Utilisation des Terrains Instables) 1986; Noverraz and Bonnard 1990; Vulliet and Hutter 1988; EPFL-NCG and Association technique Norbert De Cérenville Géotechnique + EPFL pour l'étude du glissement de La Frasse 2004; Laloui et al. 2004; Tacher et al. 2005; Planchat et al. 2009). The main features of the landslide are hereafter summarized.

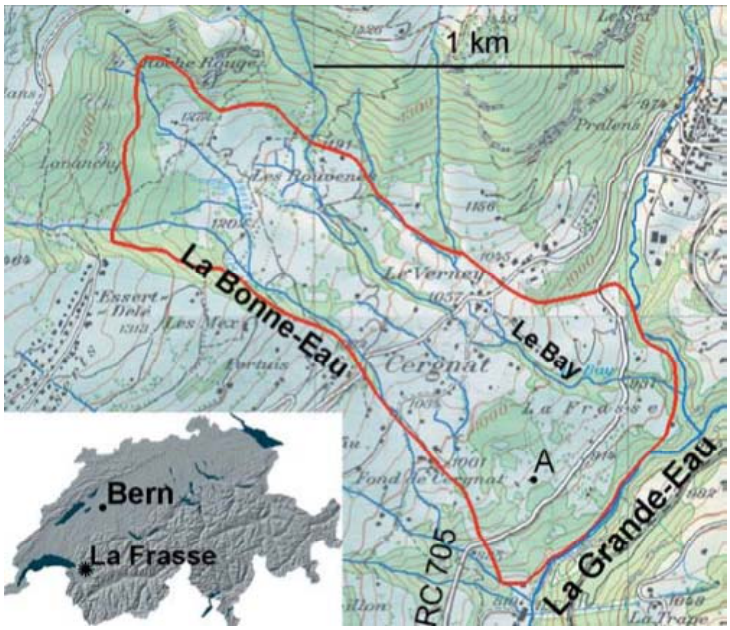


Figure 11.The La Frasse landslide: location within Switzerland and general view (Tacher et al. 2005).

Main features

Table 7 shows the general features of the case study. The landslide extends over a length of 2 km and covers an area larger than 2 km^2. It slides along a NW to SE axis, i.e., the line of the steepest slope, down to the Grande Eau River which continuously erodes its toe. Given the estimated thickness of the landslide of 80 m in its upper section and of 40 m near the toe, the total volume of the sliding mass reaches 73 million m^3. From the topographical survey (Figure 12), it is possible to distinguish an upper and a lower landslide. In its upper region, the La Frasse landslide experiences velocities of 5 to 10 cm/year. There, the instability is 80 m thick and follows a slope of 11°. Starting at 1000 m a.s.l. and going down to the foot of the landslide, the lower landslide experience velocities of 20 to 60 cm/year. The higher velocities can be explained by the absence of obstacles towards the Grande Eau River, the reduced thickness of the slide (only 40 m) and the steeper slope of 20°. The presence of a secondary scarp zone makes this morphological change obvious. The lowest quarter of the landslide is further divided into *Zone+* and *Zone++* based on their activity. *Zone++* was the most active part of the landslide until the construction of a pumping platform in 1994.

Table 7 General features of the la Frasse landslide.

Name	La Frasse landslide (Switzerland, Canton of Vaud)
Average coordinates	569 600/133 500
Active slide area	~ 1 km^2
Max / Min altitude	1500 / 800 m a.s.l.
Total length	>2000 m
Width	500 / 800 m
Average depth	80 m (upper part) 40 m (lower part)
Volume	~ 42 million m^3 (active slide) / ~ 73 million m^3 (total slide)
Mean inclination	11° (upper part) 20° (lower part)
Velocity	10–15 cm/year (upper part) / 20-60 cm/year (lower part)
Geological context	Triassic dolomites, Malm limestones, Eocene siltstones, flysch
Vegetation	Pastureland, wooded zones and swamps
Possible damage	Roads (RC 705 and RC 709), around 20 chalets, Aigle (VD)

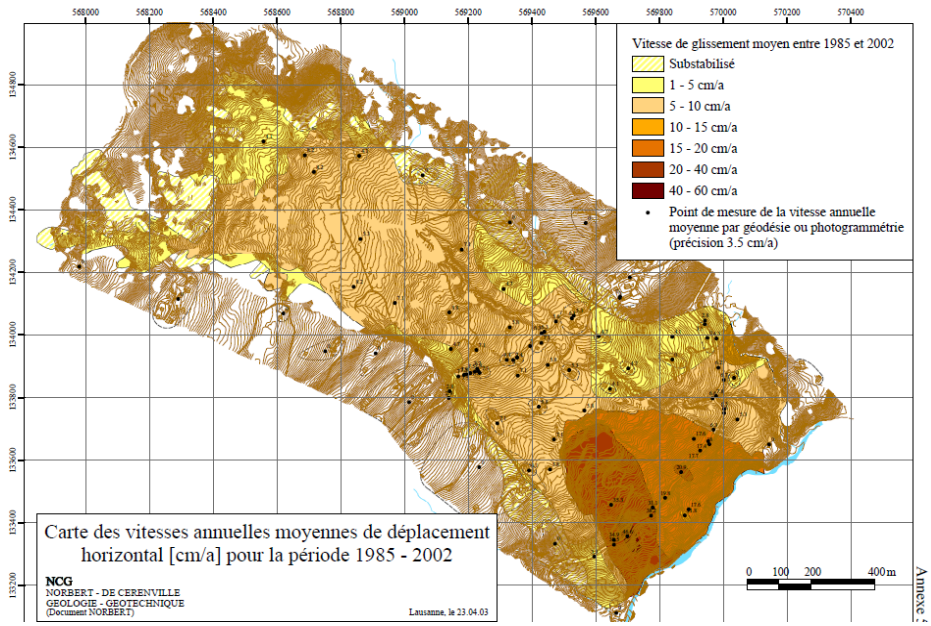


Figure 12. Kinematic analysis of the slope and division of the instability phenomenon (EPFL-NCG and Association technique Norbert De Cérenville Géotechnique + EPFL pour l'étude du glissement de La Frasse 2004)

Hydrogeology

At the landslide scale, due to the heterogeneity of the sliding material, the whole sliding mass can be considered as an aquifer. Local excess groundwater pressures (artesian levels) have been recorded not only at the main slip surface but also within the sliding mass. At the slip surface level, the hydraulic head fluctuations can reach up to tens of meters and occur over only a few hours. The variations are maximal in the vicinity of the limestone abutment due to the reduced flow section. They are less significant uphill due to gravitational drainage, and at the toe due to the consistent level of the Grande Eau River (Noverraz and Bonnard 1990).

The hydrological budget amounts to 7000 m^3 of water per day. There is no clear relationship between the acceleration phases and either gross rainfall or net infiltration even when sophisticated infiltration formulas are used (COUP model, Jansson and Karlberg 2001). A better correlation is found if past events up to two years before the day of calculations are considered and if the daily rainfall values below a certain threshold value are neglected. Even then, the contribution of direct surface infiltration does not exceed one third of the hydrological input (except after a prolonged rainfall event or during snowmelt). The rest is provided by the lateral fluxes from the regional geological context enclosing the La Frasse landslide which acts as a draining basin of nearly 3 km^2. Surprisingly, the bedrock seems not to be draining (nor feeding) the instability although it experiences lower groundwater levels than in the landslide. A dye tracing test has furthermore proved that the karstic network bypasses the site (Parriaux et al. 1987).

Triggering Mechanisms

The La Frasse landslide crises can be related to two types of triggering mechanisms. The first mechanism is classically related to the external hydrological factors. According to this theory, a crisis is prepared by a long period (probably a minimum of several months) of significant rainfall over the whole draining basin. While the hydraulic head progressively increases below the slip surface, the slope movement is finally triggered by the increase of the percolation forces within the sliding mass and the reduction of the shear strength at the slip surface. The lower part of the landslide reaches a new equilibrium soon after the rainfall event ceases, while the upper part starts to move (or not) after a time lag that is not yet known.

The necessity of an "internal" mechanism appeared after multiple accelerations were recorded in periods of low precipitation (for instance the 1966 crisis) and in order to explain the landslide behavior between crises. The theory of this internal mechanism predicts that the successive rise and drawdown of the groundwater level, creep deformation or the percolation forces could cause significant morphological changes, and in particular could foster steeper slopes. In practice, the occurrence of these two triggering scenarios is not exclusive while other factors such as the toe erosion by the Grande Eau River could also occasionally play a role.

Dataset

The landslide has been monitored by means of photogrammetry, the comparison between successive cadastral plans, topographical triangulations, GPS measurements, building crack

measurements, inclinometer readings, an extensive record of the landslide surface and in depth movements. A continuous displacement measurement device, ROBOVEC has moreover been installed in 2006. Located on the left bank of the Grande Eau River, it enables a real time monitoring of the movements in the RC 705 vicinity. Most of the inclinometers highlight a well defined and unique failure plane with little or no place for any viscous deformation. Secondary failure planes have also been observed in *Zone+* but their influence is supposedly less significant. These shallow slides were furthermore partially stabilized by the boring of a series of 10 m long anchored piles along the RC 705 in 1987.

For the purpose of the present study, it was chosen to consider the data collected during the years 1982-1995. Starting from 1995, different mitigation strategies were implemented on-site which have since altered the landslide behaviour in a significant way. Figure 13 depicts the cumulated displacement in the point A (Cf. Figure 11), which is located in the most active part of the landslide, along with the daily rainfall for the years 1982-1996.

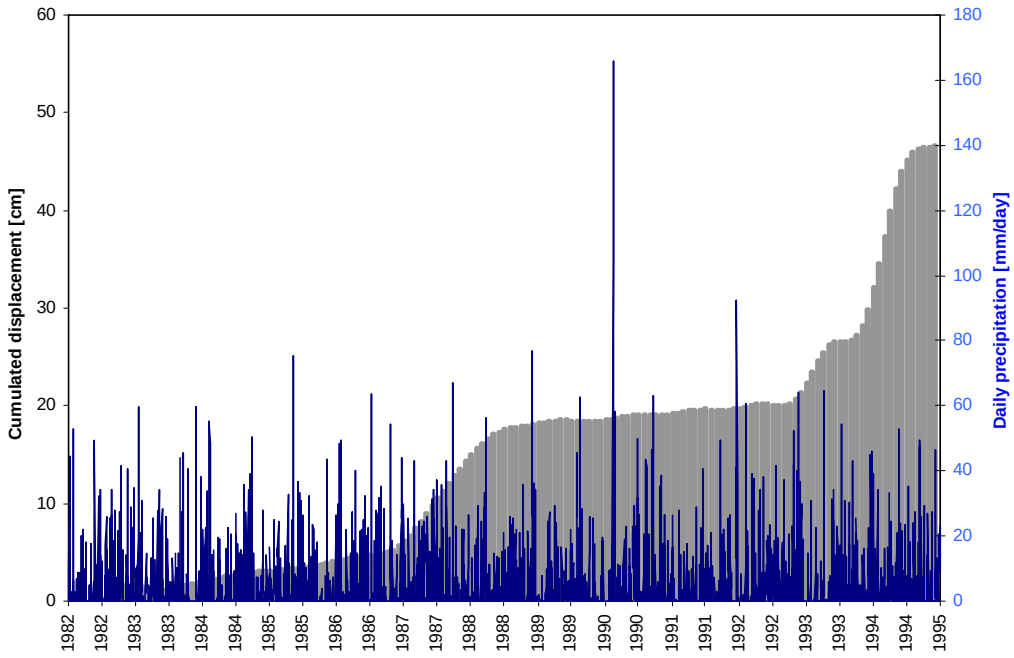


Figure 13. The La Frasse landslide: daily rainfall and cumulated displacement at point A (see Figure 11) for the years 1982 – 1995.

4.2 BARCELONNETTE (FRANCE)

General aspects

Rainfall is worldwide a recognised trigger of landslides. Hydrological triggering of mass movements can be divided in three general types: i) development of local perched water tables in the subsoil leading to shallow slope instabilities and possible gravitational flows, ii) long-lasting rise in permanent water tables leading to deep-seated slope instabilities, and iii) intense runoff causing channel-bed erosion and debris flows. Types i) and iii) are usually observed during heavy storms characterized by high rainfall intensities (hourly and daily rainfall); type ii) is usually observed through increasing water content in the subsoil due to previous rainfalls (weekly or monthly rainfall).

Many investigations have been carried out to determine the amount of precipitation needed to trigger slopes failures and associated mass movements. An important state-of-the-art on rainfall thresholds can be found in Guzzetti et al. (2007; 2008).

White et al. (1992) define a threshold as the minimum or maximum level of some quantity needed for a process to take place (trigger) or a state of change (landslide reactivation). For rainfall-induced landslides a threshold may be defined as the rainfall, soil moisture or hydrological conditions that, when reached or exceeded, are likely to trigger landslides (Guzzetti et al. 2007). Rainfall thresholds can be defined on physical (process-based, conceptual) or empirical (historical, statistical) bases (Corominas 2000; Crosta and Frattini 2001; Aleotti 2004; Wieczorek and Glade 2005; Guzzetti et al. 2008; and references therein).

The Barcelonnette Basin, South French Alps: physiographic and climatic characteristics.

The Barcelonnette Basin is located in the dry intra-Alpine zone (Figure 14) and extends from 1100 to 3000 m a.s.l. It is characterised by (1) a mountain climate with a marked interannual rainfall variability (725 ± 400 mm over the period 1928 -2009) and 130 days of freezing per year, (2) a continental influence with significant daily thermal amplitudes (> 20°) and numerous freeze-thaw cycles and (3) a Mediterranean influence with summer rainstorms yielding occasionally more than 50 mm/h. Heavy spring rains on melting thick snow layers also lead to high discharges. Meso-climatic differences on a small scale emerge due to the East-West orientation of the valley (Remaître 2006). The valley is drained by the Ubaye River which is fed by several torrents on the north- and south-facing slopes. It constitutes a geological window, baring the autochthonous Callovo-Oxfordian black marls, also called 'Terres Noires', under the allochthonous Autapie and Parpaillon flysch (Maquaire et al. 2003). Local slopes are characterised by specific morphology due to the geological setting.

Data: mass movement catalogues, meteorological data and methodology.

Collection of historical data in catalogues, newspapers, monographs, technical reports, bulletins and scientific papers for the period between 1850 and 2009 has been carried out. Detailed descriptions on the type and quality of information collected and the methodologies used to analyze the data can be found in Flageollet et al. (1999) and Remaître (2006).

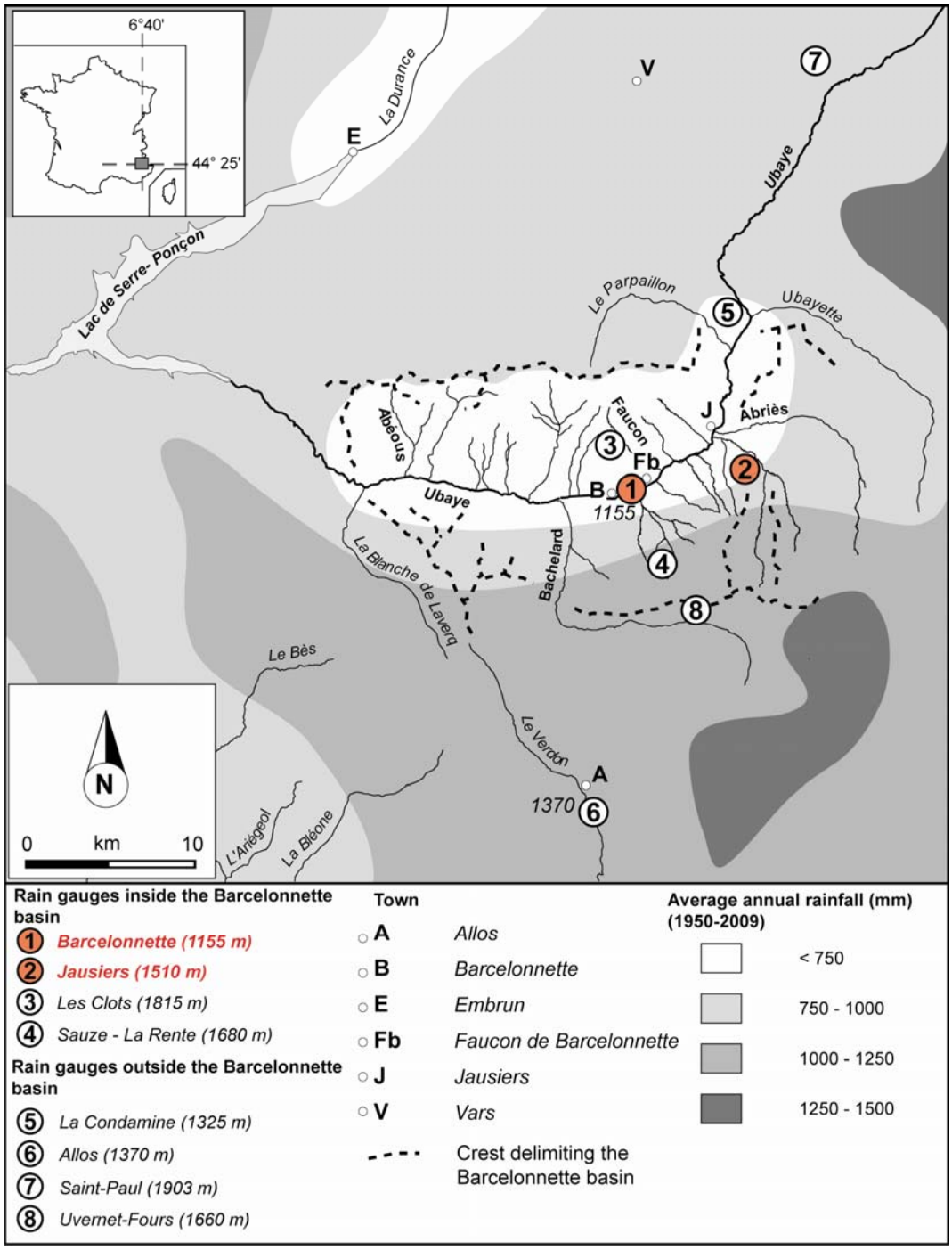


Figure 14. Mean annual precipitation (M A P) map of the South French Alps and location of meteorological stations in the surrounding of the Barcelonette basin. The two stations in red are those used in this study.

Over the period 1451-2010, about 600 references with exact date and location of triggering have been recorded. For each soil slide and debris flow event, the information collected and

stored in the Barcelonnette catalogue includes: (1) the precise or approximate location of the area affected by the soil slide or debris flow, (2) the precise time, date or period of failures, (3) the rainfall condition that resulted in slope failures, including the daily rainfall, the antecedent rainfall for various periods and for the Barcelonnette station from 1998: the total event rainfall, the rainfall duration, the mean rainfall intensity and peak rainfall intensities, (4) the landslide type, (5) a generic description of the rock types and associated soils.

There are 8 meteorological stations located inside or in the surroundings of the Barcelonnette basin (Table 8). In this work, only two of these stations were used (Figure 14), because (1) both stations exhibit a similar annual precipitation pattern (Figure 15), (2) their locations are close to the main debris flows and soil slides observed within the basin, and (3) they cover 50 years of data (1960-2009).

Table 8. Characteristics of the climatic stations in the Barcelonnette area.

Meteorological station	X coordinate	Y coordinate	Elevation (m)	Date of creation	Type of data
(1) Barcelonnette	44°23'30"N	6°40'18"E	1155	1925	Rainfall & temperature (Hourly)*
(2) Jausiers	44°24'48"N	6°43'24"E	1240	1961	Rainfall & temperature (Daily)
(3) Les Clots	44°24'52"N	6°39'48"E	1815	2001	Rainfall & temperature (Hourly)
(4) Sauze – La Rente	44°21'24"N	6°41'55"E	1680	1991	Rainfall & temperature (Hourly)
(5) La Condamine	44°27'39"N	6°44'46"E	1325	1955	Rainfall & temperature (Daily)
(6) St-Paul	44°35'30"N	6°50'30"E	1903	1971	Rainfall & temperature (Daily)
(7) Allos	44°13'36"N	6°37'54"E	1370	1971	Rainfall & temperature (Daily)
(8) Uvernet-Fours	44°19'06"N	6°41'30"E	1660	1955	Rainfall & temperature (Daily)

* Hourly data are available from 1998

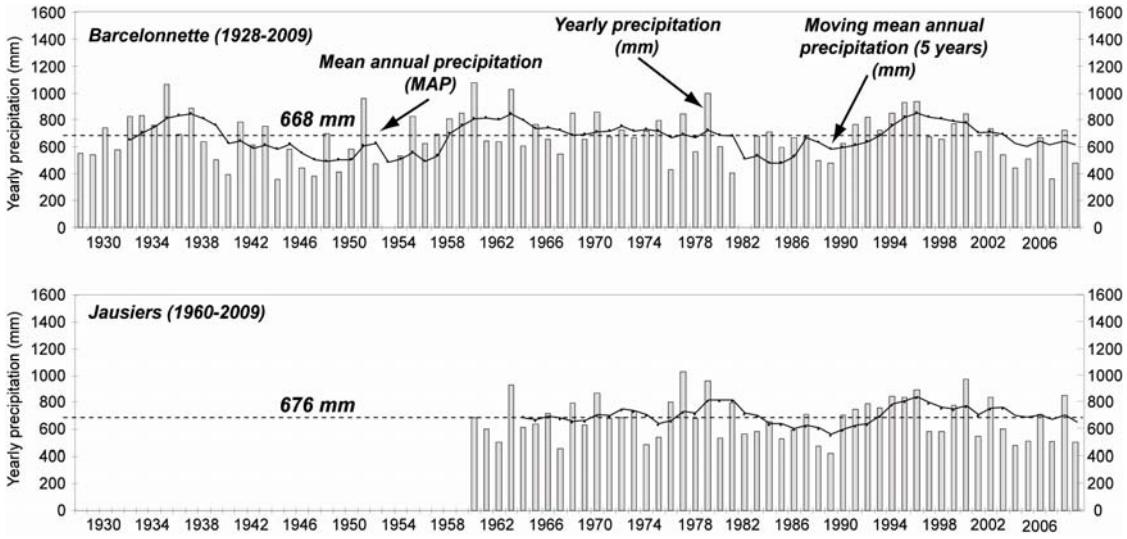


Figure 15. Annual precipitation for the Barcelonnette and Jausiers rain gauges.

Data are mainly available for 24 h-periods; data for 1 h-periods are only available for the Barcelonnette meteorological station (1998-2010).

For this work, the information on rainfall events was checked in order to identify: duplicate events, events for which the rainfall information was not considered as relevant or sufficiently documented. These events were excluded from the catalogue. Soil slides or debris flows that were not caused only by rainfall (e.g., related to rapid snow melting) were also excluded. At the end of the review process, a catalogue of 106 mass-movement events (53 debris flows and 53 soil slides, see Figure 16) was used for the daily rainfall analysis (section 5.2.1); 12 events (6 debris flows and 6 soil slides, see Figure 17) were used for the hourly rainfall analysis (section 5.2.2).

ID	LOCATION	TOWN	Station P.	Day	Month	Year	Label	P.J0	Moy.P.J0	Yearly.Rain	Mean.Yea.Rain	Pmois	M.Mens	Pcum6M	M.cum6	Pcum5M
1	Riou-Bourdoux	Saint-Pons	Barcelo	7	8	1950	DF_01	8,2	1,6	582,9	691,1	32,1	51,0	237,4	320,0	179,8
2	Riou-Bourdoux	Saint-Pons	Barcelo	14	8	1950	DF_02	0,0	1,6	582,9	691,1	32,1	51,0	236,5	318,8	189,9
3	Riou-Chanal	Barcelonnette	Barcelo	25	7	1952	DF_03	0,0	1,4	471,5	691,1	11,8	43,8	270,8	318,9	239,1
4	Sanières	Jausiers	Barcelo*	24	6	1955	DF_04	7,0	2,0	826,1	691,1	121,5	60,3	486,1	322,7	317,0
5	Général sur tout le Bassin	Barcelonnette	Barcelo	13	6	1957	DF_05	38,8	2,0	690,6	691,1	105,3	60,3	303,3	316,0	264,4
6	Sanières + Abriès	Jausiers	Jausiers	16	9	1960	DF_06	58,5	2,0	685,1	676,6	170,9	59,6	132,6	315,8	132,6
7	Sanieres	Jausiers	Jausiers	23	6	1962	DF_07	4,3	2,0	502,1	676,6	38,2	59,8	309,2	321,7	214,8
8	Riou-Bourdoux	Saint-Pons	Barcelo	11	1	1963	DF_08	7,3	1,6	1027,6	691,1	77,0	50,0	337,8	366,3	329,4
9	Faucon	Faucon	Jausiers	28	6	1963	DF_09	3,6	2,0	929,0	676,6	72,0	59,8	379,2	322,4	313,2
10	Général sur tout le Bassin	Barcelonnette	Barcelo	15	11	1963	DF_10	43,8	2,6	1027,6	691,1	277,0	77,3	577,1	372,2	499,9
11	Faucon	Faucon	Jausiers	18	7	1964	DF_11	13,5	1,4	613,1	676,6	50,0	42,5	375,2	316,1	366,5
12	Riou-Bourdoux + Marquises	Saint-Pons	Barcelo	4	7	1965	DF_12	22,7	1,4	763,0	691,1	69,9	43,8	220,2	262,3	247,4
13	Buriane + La Combe	Faucon	Barcelo	10	8	1967	DF_13	19,7	1,6	544,5	691,1	67,2	51,0	275,5	318,3	202,3
14	Poche + R.-Bdx + La Combe + La Tourache	Jausiers	Jausiers	11	9	1970	DF_14	27,4	2,0	454,8	676,6	42,4	59,6	314,8	316,8	257,0
15	Riou-Bourdoux + La Combe	Saint-Pons	Barcelo	14	7	1972	DF_15	34,6	1,4	724,4	691,1	70,9	43,8	340,7	320,2	242,3

ID	Pcum3M	M.cum3	Pcum2M	M.cum2	Pcum30	M.Cum30	Pcum14	M.Cum14	Pcum7	M.Cum7	Pcum6	M.Cum6	Pcum5	M.Cum5	Pcum4	M.Cum4	Pcum3	M.Cum3	Pcum2	M.Cum2
1	128,6	161,6	45,8	101,1	22,4	43,1	4,0	17,6	4,0	9,2	4,0	8,4	0,0	7,8	0,0	6,5	0,0	5,7	0,0	4,9
2	121,2	163,5	50,6	99,2	31,5	42,6	15,3	21,6	11,3	12,5	3,1	10,5	3,1	8,9	3,1	6,8	3,1	5,5	3,1	3,3
3	80,4	169,1	33,9	112,6	16,3	49,1	11,8	21,0	10,3	8,2	10,3	5,8	10,3	5,1	6,8	3,8	2,1	2,8	0,0	2,0
4	156,6	175,4	144,0	120,7	95,7	62,8	15,3	28,9	15,3	15,2	15,3	13,1	15,3	11,4	4,1	8,8	0,0	6,1	0,0	4,2
5	146,8	170,8	106,0	118,9	68,6	63,5	44,7	28,7	30,5	14,3	30,5	12,7	30,5	9,9	22,9	8,4	22,9	6,1	12,2	3,5
6	132,6	137,2	132,6	88,9	77,8	48,6	70,8	21,7	56,9	13,0	56,9	11,5	56,9	10,6	56,9	8,6	56,9	7,0	53,2	4,4
7	140,8	181,6	64,1	123,7	45,4	62,6	20,1	26,9	3,1	11,9	3,1	10,4	3,1	8,5	3,1	7,1	3,1	5,2	3,1	2,8
8	223,5	196,5	150,6	116,8	129,0	48,3	95,0	22,9	28,0	11,5	27,0	10,6	9,2	8,4	2,2	6,6	2,2	6,1	2,3	4,5
9	163,6	183,6	91,2	121,5	67,0	61,0	24,6	27,6	24,6	13,2	24,6	11,9	24,6	10,4	17,8	7,8	7,4	6,5	7,4	4,8
10	318,9	208,7	247,5	156,7	208,6	79,1	195,7	43,0	24,6	19,4	24,6	17,8	24,6	14,4	11,8	12,8	11,8	10,2	6,1	6,2
11	174,3	176,0	141,1	113,1	63,3	49,2	24,6	20,9	7,2	9,8	7,2	9,0	7,3	7,6	7,3	6,4	7,3	4,3	7,3	2,7
12	151,4	176,4	128,9	118,0	65,3	60,2	27,2	27,4	26,6	10,7	26,6	8,5	26,6	6,3	23,7	5,0	13,6	3,8	8,8	2,3
13	165,0	162,7	92,8	100,3	44,6	43,6	36,8	20,4	36,8	12,2	36,8	11,3	36,4	10,6	13,2	6,8	13,2	5,7	13,2	3,7
14	166,0	138,0	123,8	86,0	49,2	52,2	44,6	21,6	6,3	9,5	6,3	8,0	6,3	5,9	6,3	4,6	6,3	3,8	6,3	2,4

Figure 16. View on a sample of the database. Label "Land_" corresponds to soil slides; label "DF_" corresponds to debris flows. P0 is the total amount of rainfall for the day of occurrence of the event.

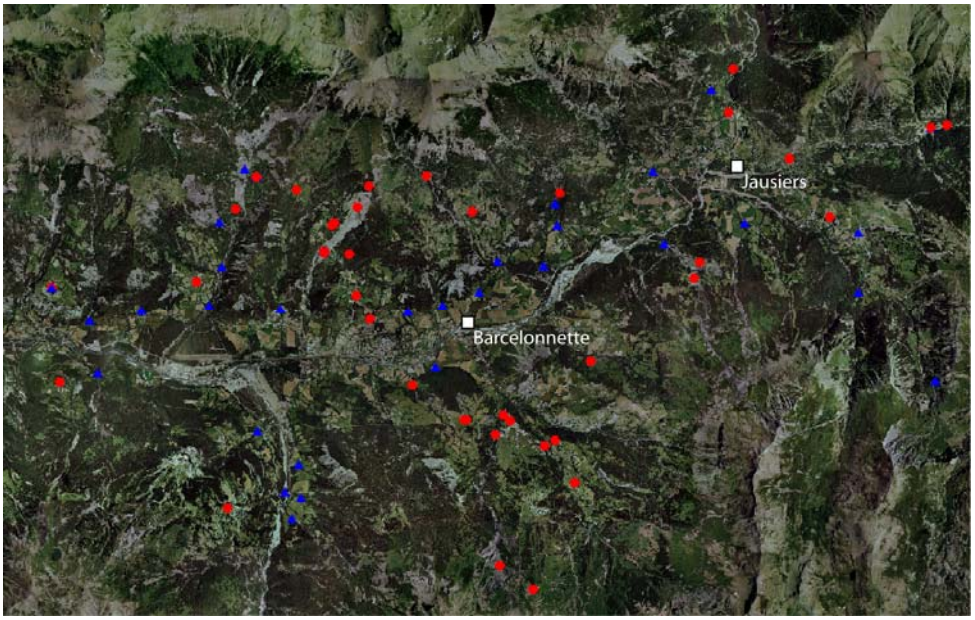


Figure 17. Orthophoto of the Barcelonnette basin and location of the mass movement events. The blue triangles correspond to debris flows, the red hexagons correspond to soil slides.

4.3 SOUTH-EASTERN NORWAY

The dataset comprises the Nedre Romerike area in South-Eastern Norway. The Norwegian landslide database includes 30 earth slides (not induced by human activity) in the Nedre Romerike area. All these events were triggered during autumn 2000. Many of these landslides occurred near rivers and lake shores. The terrain is characterised by flat open areas that are cut by steep river ravines. The terrain was formed from river and marine sediments after the last ice age. The ground consists mainly of a thin layer of soils of grain sizes varying from gravel to clay. There are large areas of desalinated marine quick clays. In South-Eastern Norway, major historical landslides have occurred in these quick clay deposits leading to the destruction of property and loss of lives (Jørstad 1968; Karlsrud and Lillevik 1984; Karlsrud et al. 1985; Furseth 2006). The area is ideal for farming and comprises a significant fraction of the farm land in Norway. Due to the high population density in this area, there is a strong pressure on authorities for reduction of landslide risk in the area. The dataset of landslide events in South-Eastern Norway during autumn 2000 has been described in detail by Jaedicke and Kleven (2008). The distribution of landslides is shown in Figure 18. The time-uncertainty of these events ranges between 1 day and 64 days. Daily data from the station Skedsmo-Hellerud (1972-2003) was used for estimating precipitation conditions. This data was retrieved from eKlima (http://eklima.no), the web portal of the Norwegian Meteorological Institute.

A typical event from this crisis is shown in Figure 19. Even though the return period of the 2-week antecedent precipitation barely exceeded 10 years, the accumulated precipitation over the previous 50 days corresponded to a much larger return period: approximately 150 years (based on the analyses of events that occurred on 30 November 2000, see Figure 20).

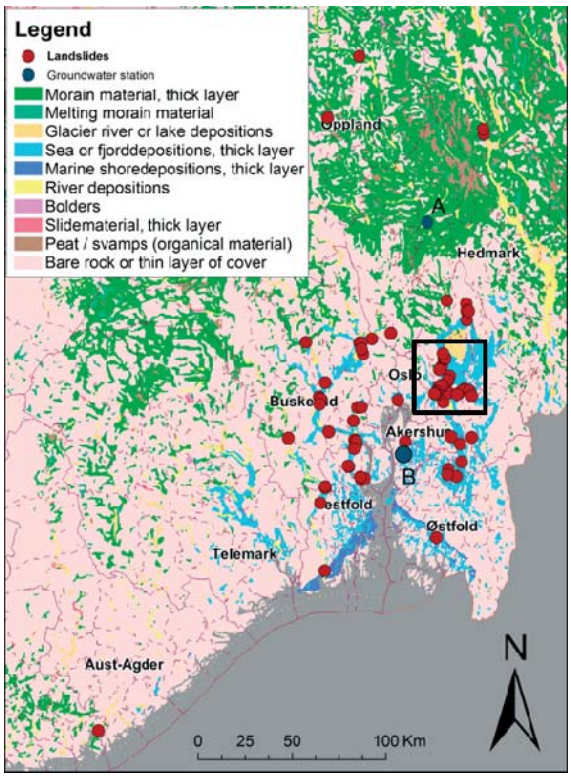


Figure 18. **South-Eastern Norway: landslide inventory during autumn 2000 and surficial deposits (Jaedicke and Kleven 2008). The landslide events used for estimation of thresholds in the present study are inside the black square (east of Oslo) in the Nedre Romerike area.**

Figure 19. Landslide at Dorr, Minnesund, South-Eastern Norway triggered on 22 November 2000 (Jaedicke and Kleven 2008).

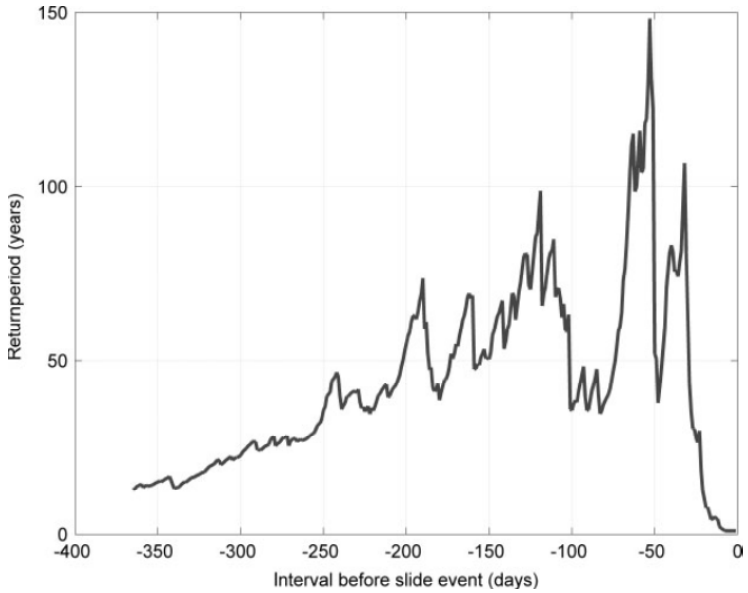


Figure 20. Return period of antecedent precipitation for events occurring on 30 November 2000 (Jaedicke and Kleven 2008).

4.4 SARNO, SATRIANO AND VERZINO (ITALY)

Datasets containing three landslide sites have been provided; two of them are located in the Calabria region, the third one in the Campania region:

1) Sarno (Campania region);
2) Satriano (Calabria region);
3) Verzino (Calabria region).

The datasets for each case consist of precipitation data and dates of historical movement. These data are necessary for applying empirical models. The precipitation data consists of daily rainfall data. The criteria for selecting the rain gauge were the proximity to the landslide site, and the completeness of the time series. The following paragraphs describe the selected landslide sites.

4.4.1 Sarno

Sarno is located at the toe of the Pizzo d'Alvano massif which is located in the Campania Apennines of Southern Italy, about 30 km from Naples (Figure 21). The historical mobility date is 5 May 1998.

GENERAL INFORMATION

Region:	Campania
Municipality:	Sarno
Province:	Salerno
Hydrographic Basin:	Tuostolo - Scalandrone
Cartographic datum	I.G.M: F.448116

Figure 21. Landslide site in Sarno, Italy.

On 5 May 1998, more than 100 slides were triggered on the Pizzo d'Alvano Mountain, due to a meteorological event characterised by heavy and prolonged rainfall. Many slides transformed into mudflows which hit the urban areas of four small towns, including Sarno. In Sarno, the affected area was about 2 square kilometres (Figure 22). As a consequence of these events, 159 people died. The rain gauge considered for this site is the Santa Maria la Foce station. The dataset contains the daily rain data recorded since 1 January 1919 to 31 December 1998.

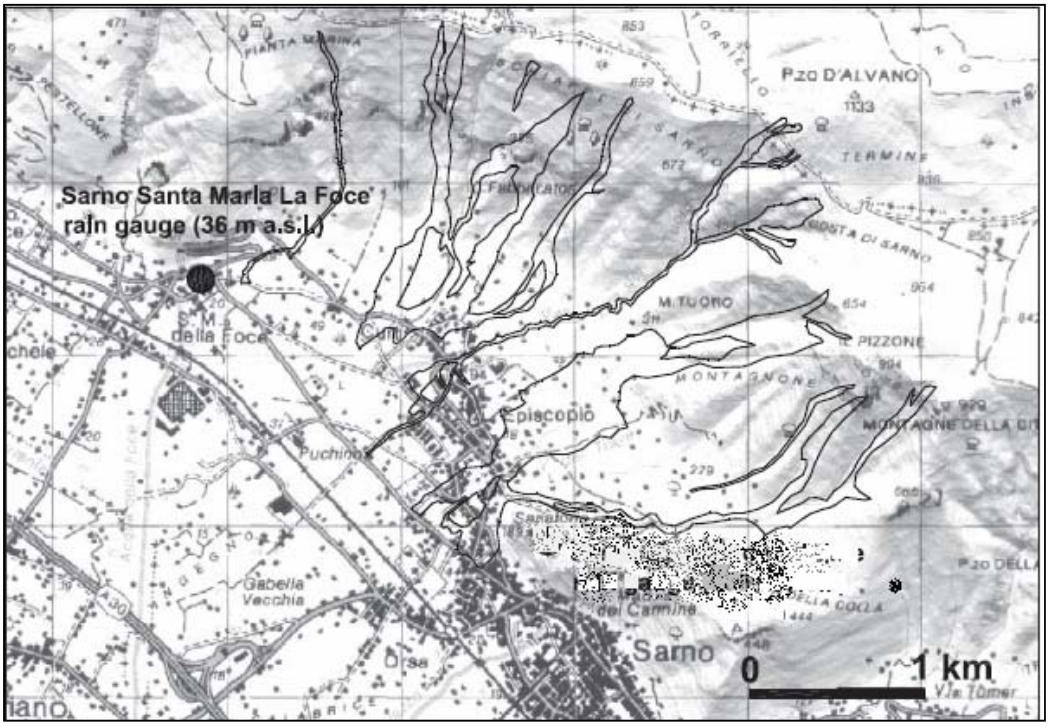


Figure 22. Rain gauge location and main mudflows in the Sarno area.

The bedrock is mainly limestone, dolomitic limestone and marly limestone. The Pizzo d'Alvano slopes are mantled by pyroclastic soils, derived from the explosive phases of the Somma-Vesuvius volcanic activity, both as primary air-fall deposits and as secondary deposits. The first ones are constituted by pumice and ash; the secondary deposits have been re-worked by sheet erosion and by mass-wasting processes. These deposits are present as debris and colluvium in a vast area; air-fail deposits covered a wide area reaching distances up to 50 km. The thickness of the pyroclastic cover varies from less than 1 m up to 15-20 m. Pumiceous and ashy deposits belonging to at least 5 different eruptions were recognized. From the older to the younger, they are: *Ottaviano Pumice* (8000 years B.P.), *Avellino Pumice* (3800 years B.P.), *79 A.D. Pumice*, *472 A.D. Pumice*, *1631 A.D. Pumice*.

4.4.2 Satriano

Satriano is a village located close to the city of Catanzaro, in the Calabria region (southern Italy). The historical movement dates are on 18 October 1951 and 2 January 1973.

GENERAL INFORMATION

Region:	Calabria
Municipality:	Satriano
Province:	Catanzaro
Hydrographic Basin:	Ancinale basin
Cartographic datum	I.G.M: F.242 – III SO C

Figure 23. Landslide site in Satriano, Italy.

The landslide movement int the Satriano area is a slide evolving in a flow. The precipitation data was recorded by the Chiaravalle Centrale rain gauge, from 1 January 1951 to 30 June 2000.

In the area outcrops, on the structure of igneous acid coarse rocks of the Paleozoic, a stratigraphic sequence constituted from the bottom layer by:

1. Sandy conglomerates with cobblestones, generally rounded, of granite and metamorphic rocks.
2. Coarse sands, non fossiliferous, with conglomeratic horizons
3. Clays and marls, of colour from grey-azure to white, with thin intercalations of sands and silts with rich microfauna. There are also sands, gravels, conglomerates of the Pleistocene and Holocenic warpes.

Two faults with throw and direction NE-SW have been identified. The high primary permeability in sands and conglomerates allows groundwater circulation and storage in these deposits.

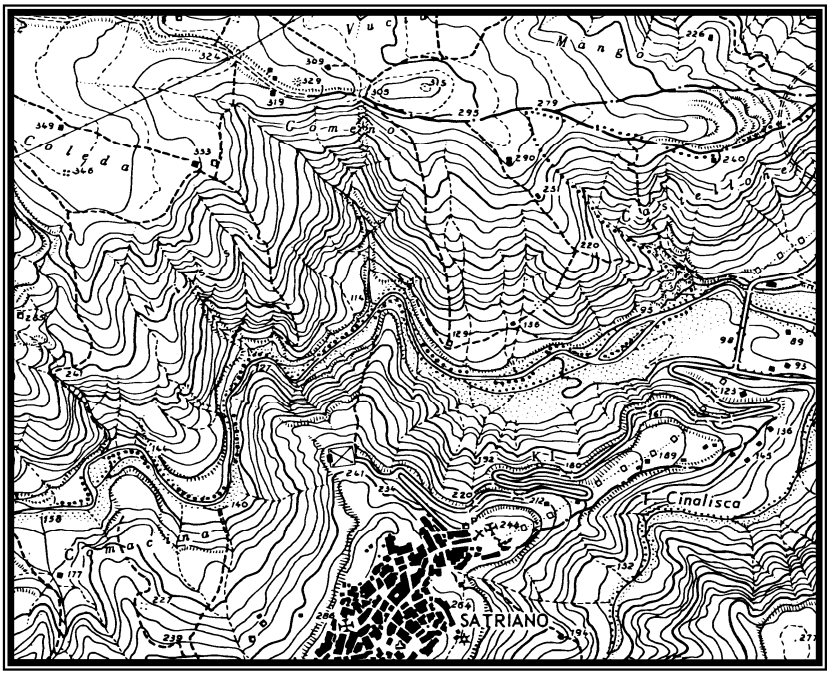


Figure 24. Map of the Satriano area.

4.4.3 Verzino

Verzino is a village 49 km from Crotone, in the Calabria region (southern Italy). The historical mobility date is 8 January 1991. The landslide was a complex movement.

GENERAL INFORMATION

Region:	Calabria
Municipality:	Verzino
Province:	Crotone
Hydrographic Basin:	Vitravo basin
Cartographic datum	I.G.M: F.237 – I NE D

Figure 25. Landslide site in Verzino, Italy.

The rain gauge considered is the Verzino rain gauge. The dataset consists of daily rainfall data recorded from 1 September 1921 to 26 March 2000.

The area presents a sequence of sedimentary soils. From the lower to the upper layer there are: conglomerates, coarse sandstones and conglomerates that, proceeding to high, pass to a clayey-siltic unit, while at the south of the Verzino outcrop, there are white clays, marls and diatomites. This unit consists of gypsum, gypsum-sandstones and passing, proceeding to high, to clays with sandstones and limestones. In the upper zone white limestones prevail. In the Verzino area, the upper-Miocenic soils have a synclinal structure at large scale; in fact along the creeks, the metamorphic substratum and the inferior sedimentary units outcrop. Groundwater circulation is influenced by soil heterogeneity. A first hydrogeological block is represented by formations with conglomerate-sandstones and clays; a second one is represented by formations of gypsum, clays with sandstones and limestones. There is shallow groundwater recognized by several springs.

Figure 26. Map of the Verzino area.

4.5 WESTERN NORWAY

The dataset from Western Norway corresponds to the Norangselva catchment. The location of the catchment is shown in Figure 27 with the precipitation regions of Norway.

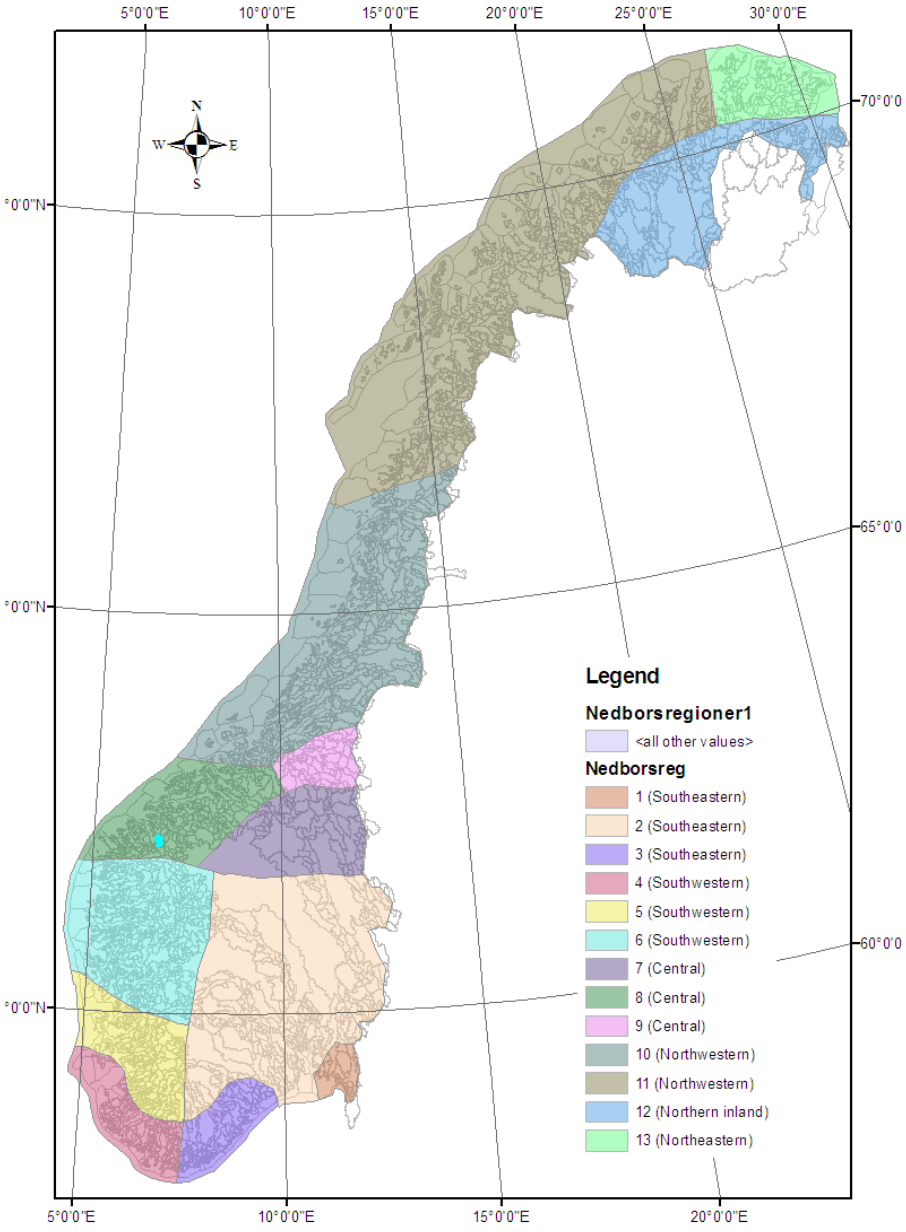


Figure 27. Precipitation regions of Norway and location of the case study area, the Norangselva catchment (highlighted area in precipitation region No. 8).

The Norangselva catchment has an area of 56 km^2 and is located in the county of Møre og Romsdal covering the municipalities of Ørsta and Stranda. Figure 28 shows a hillshade map of the catchment and the distribution of landslide incidents in the dataset. All the landslides were reported by the road authorities. In consistence with this, note that the locations of the incidents follow the road.

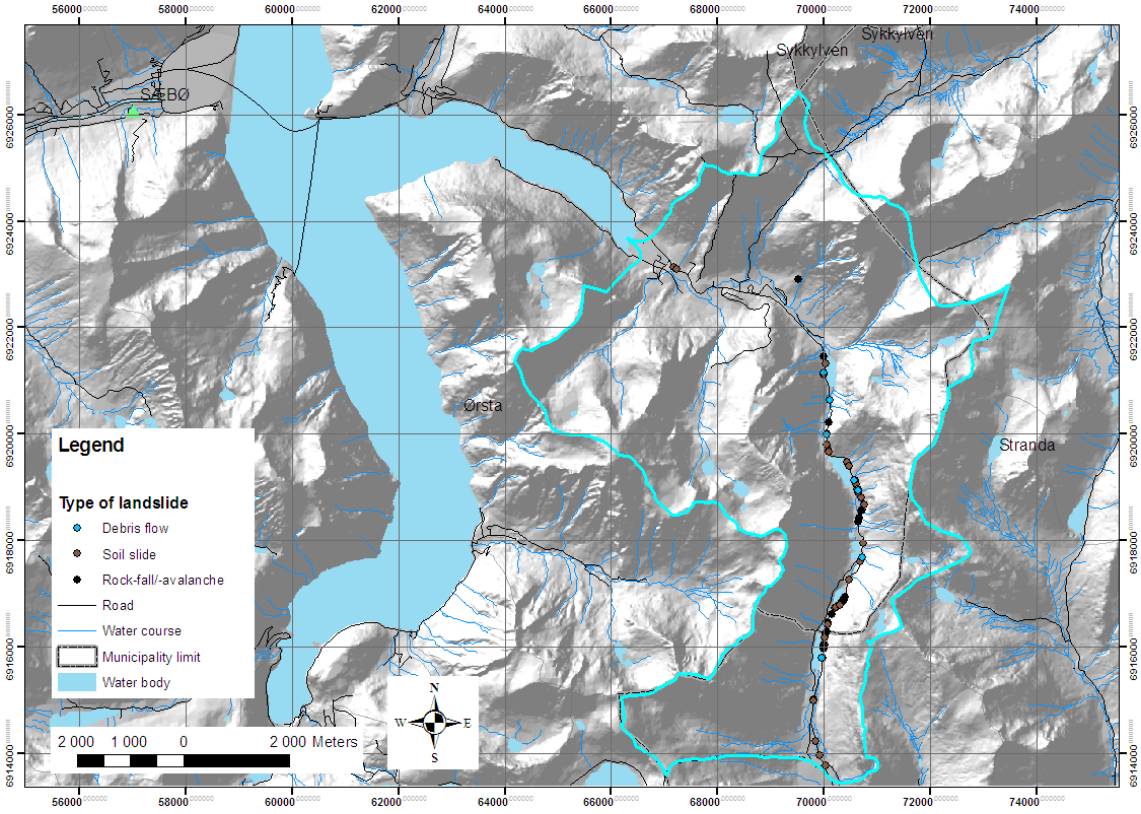


Figure 28. Landslide incidents at the Norangselva catchment (catchment No. 097.4Z). Area of catchment: 56 km^2; County: Møre og Romsdal; Municipalities: Ørsta and Stranda.

Soil slides and debris flows incidents at the catchment are associated with remobilization of deposits along channels and ruptures on moraine or weathered materials. Rock slides and rock falls occur due to rupture of gneiss in steep slopes. Figure 29 shows the distribution of surficial deposits and Figure 30 shows the bedrock geology of the area.

The dataset comprises 90 landslides from 1892 to 2005. The distribution according to their type is as follows:

- Debris flows: 13
- Soil slides: 54
- Rock falls and rock avalanches: 23

A summary of the dataset is shown in Table 9.

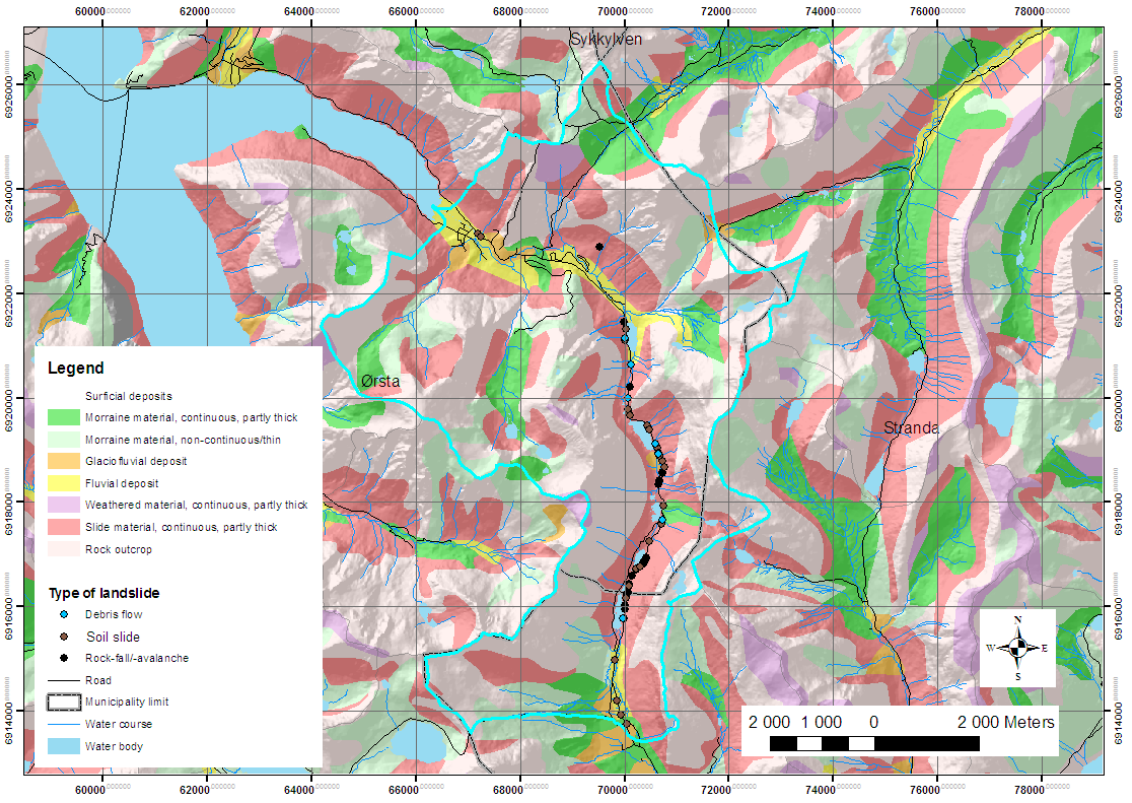


Figure 29. Surficial deposits at the Norangselva catchment.

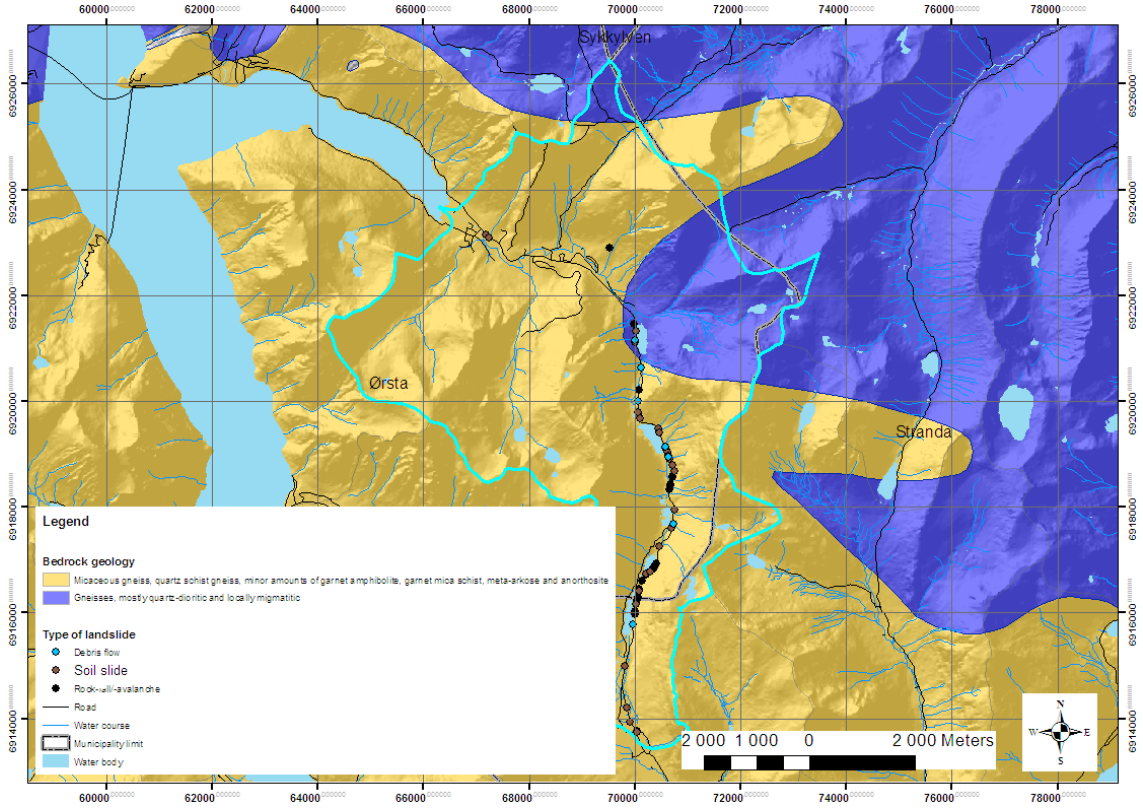


Figure 30. Bedrock geology at the Norangselva catchment.

Table 9. Landslide incidents at Norangselva catchment. Note: all events, except 1892, 1908, and 1926, were reported originally by the Norwegian Public Roads Administration (http://www.vegvesen.no/en/Home).

Year	M	D	H	Min	Land slides	Debris flows (#)	Soil slides (#)	Rock falls and Rock slides (#)	Dist1 (§) (km)	Dist2 (*) (km)
1892	-	-	-	-	1			1 (?)		
1900	1	1	0	0	1	1 (30)				
1908	-	-	-	-	1			1 (1000000)		
1926	-	-	-	-	1			1 (?)		
1992	12	15	18	0	1		1 (1300)			
1993	1	11	23	0	4		4 (?)		6.1	
1993	1	27	5	15	1		1 (500)			
1993	10	14	7	15	1			1 (10)		
1993	12	12	7	0	4		4 (100-1000)		2.0	
1993	12	18	17	0	2		2 (35-600)		0.8	
1993	12	31	20	0	4		4 (125-1440)		3.8	
1994	10	6	13	0	1		1 (125)			
1994	12	25	9	0	1		1 (900)			0.05
1994	12	27	11	0	1		1 (400)			
1995	5	24	6	30	1		1 (350)			
1995	8	1	22	0	8		7 (50-375)	1 (375)	3.1	
1997	1	30	12	0	2		2 (?)		1.9	
1997	2	3	20	0	1		1 (?)			0.5
1997	2	5	8	0	1		1 (?)			
1997	2	12	13	30	1		1 (950)			
1997	10	23	18	0	1		1 (?)			
1998	2	2	17	0	1		1 (?)			0.0
1998	2	3	17	0	1		1 (?)			
1998	2	4	12	0	1		1 (?)			
1998	2	28	18	0	1		1 (?)			
1998	3	3	22	0	1		1 (?)			0.0
1998	3	4	22	0	1		1 (?)			
1999	2	5	18	0	1		1 (?)			
1999	2	15	10	0	1		1 (?)			
1999	3	29	13	0	1		1 (560)			1.8
1999	3	30	5	0	1		1 (1000)			
1999	11	4	0	0	2		2 (?)		0.3	
1999	12	16	0	0	4		4 (?)		6.3	
2000	5	22	0	0	4		4 (?)		6.3	
2001	7	10	14	0	1			1 (?)		
2001	11	2	0	0	1	1 (?)				5.5

2001	11	2	10	0	1	1 (?)				
2001	11	2	12	0	1		1 (?)			
2001	11	10	0	0	3			3 (?)	0.2	0.2
2001	11	10	6	30	1			1 (?)		
2002	10	24	17	0	2			2 (?)	0.06	
2003	2	16	8	0	1			1 (?)		
2003	4	21	18	30	3			3 (?)	3.2	
2003	5	2	17	30	1			1 (3)		
2003	8	15	12	0	8	8 (5-100)			3.6	
2003	10	9	0	0	2	2 (30, ?)			2.1	
2004	3	19	8	0	1			1 (?)		
2004	4	3	14	55	1			1 (?)		
2005	1	2	9	45	1			1 (?)		0.0
2005	1	2	10	20	1			1 (?)		
2005	5	16	19	0	1			1 (?)		
2005	8	18	19	0	1			1 (?)		
Σ					90	13	54	23		

Notes:
M: month; D: day; H: hour;
#: number of incidents (range of volumes in m^3)
§: maximum distance between incidents occurring at the same date and time
*: maximum distance between consecutive incidents occurring within 72 hours

5 MODEL EVALUATION

5.1 LA FRASSE (SWITZERLAND)

The case study was tackled with two methods: 1) the linear black-box models for time series prediction; 2) their nonlinear counterparts implemented through the utilization of neural network models.

The data set includes data of 161 monthly cumulated displacements for the period 1982 – 1995. The first 128 patterns of the whole data set were used as the training set, whereas the remaining 33 patterns (including the main crisis in 1994) were used as the validation set. The model providing the best results was in this case the ARX 441, which is an ARX model with $n_a = 4$, $n_b = 4$ and $n_k = 1$ (for all the 28 input channels, i.e. the daily rainy in the previous 28 days). These parameters were estimated minimizing the sum of squares of the right-hand side minus the left-hand side of the expression (49), with respect to *a* and *b*.

In Figure 31, the results of the one-step-ahead forecasts obtained with the ARX441 model on the training set are shown. The percentage of the output variation that is explained by the model is also computed and displayed in the figure. It is given by the expression:

$$fit = 100\left(1 - \frac{|yh - y|}{|y - mean(y)|}\right) \qquad 55$$

where *yh* is the model output (i.e. the output predicted by the model) and *y* is the measured output of the system.

Figure 32 depicts the one-step-ahead forecasts of the same model on the validation set. In this case the model behaves well, which means that it has good generalization ability. Figure 33 shows a zoom of the last part of the graph showed in Figure 32, where the slight difference can be observed between the forecasted and the measured values.

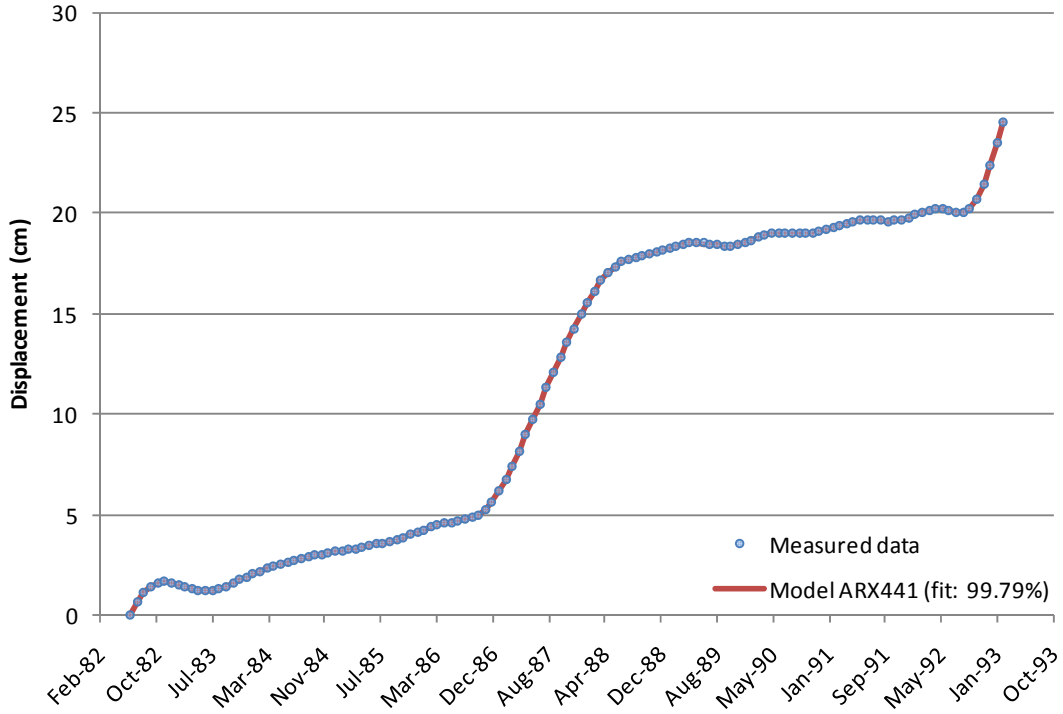


Figure 31. Model ARX 441: Comparison of measured and predicted values for the training set.

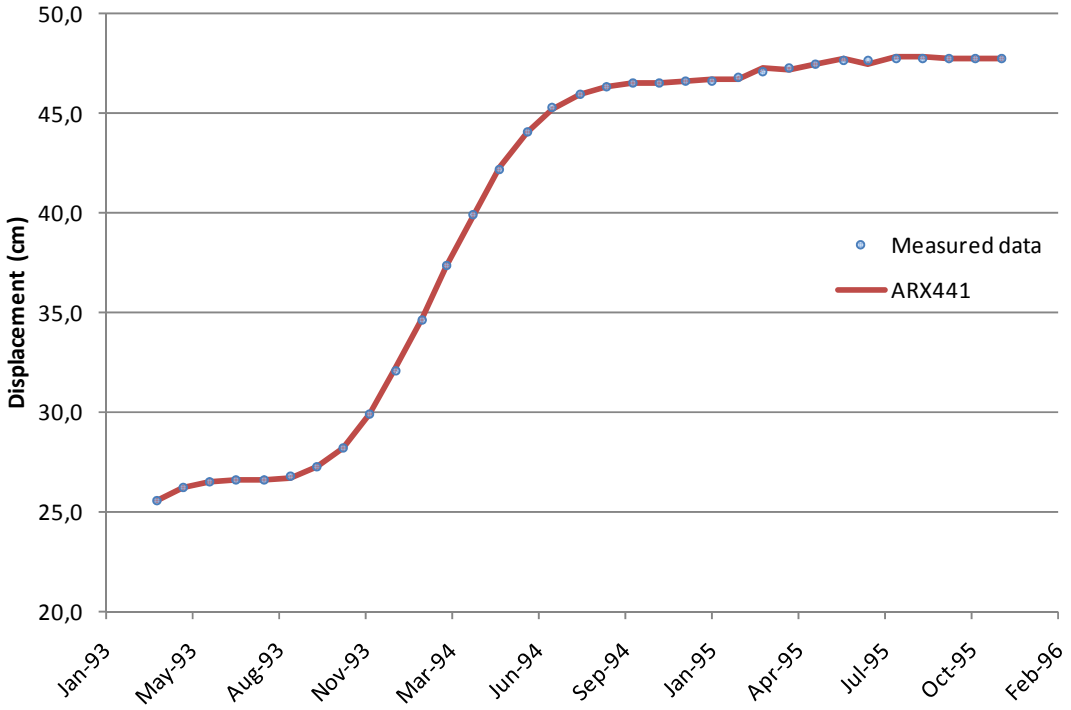


Figure 32. Model ARX 441: Comparison of measured and predicted values for the validation set.

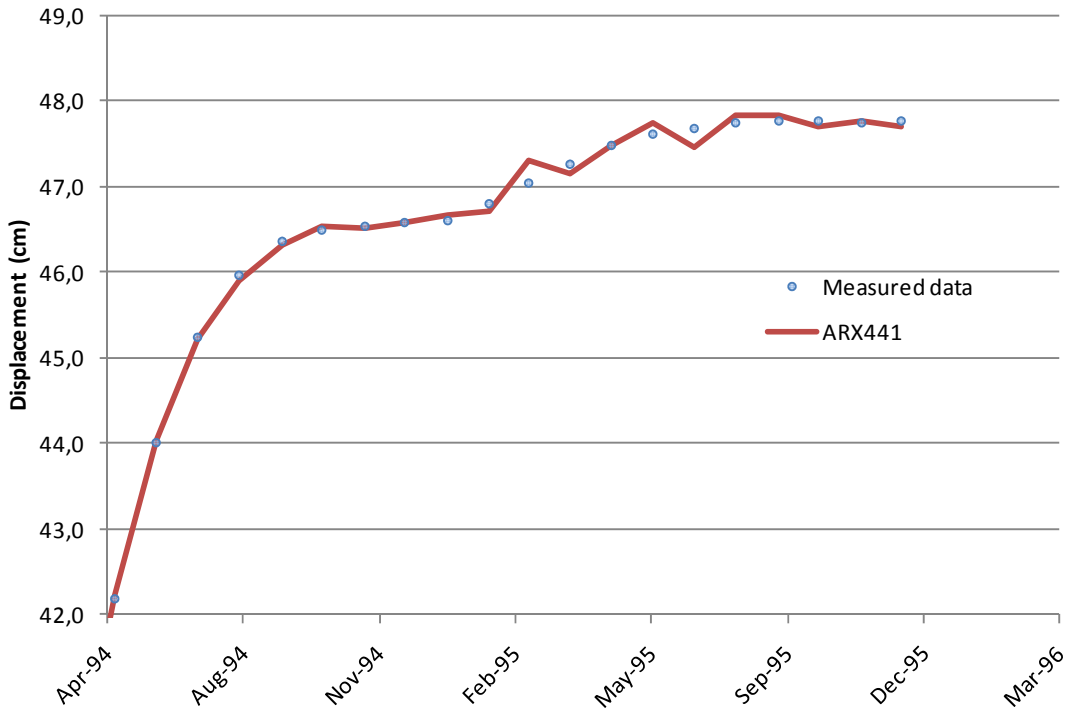


Figure 33. Zoom of the last part of the graph shown in Figure 32.

The second approach, based on the nonlinear versions of the time series forecasting models, was also attempted. Since the time series do not exhibit marked nonlinearity features, these models did not yield very good results.

In the case study under analysis, the models yielding the best results were the NNARX and NNARMAX models (see Figure 9).

Different structures of the models were tried. As an example, Figure 34 and Figure 35 show the results of a the one-step-ahead forecasts obtained with a model NNARMAX 2121 (n_a=2; n_b=1; $_{nc}$=2; n_k=1), respectively on the training set and on the validation set. The lower part of the figures also show the evolution of the prediction error. It is possible to see that in this case the model does not have a good generalization performance.

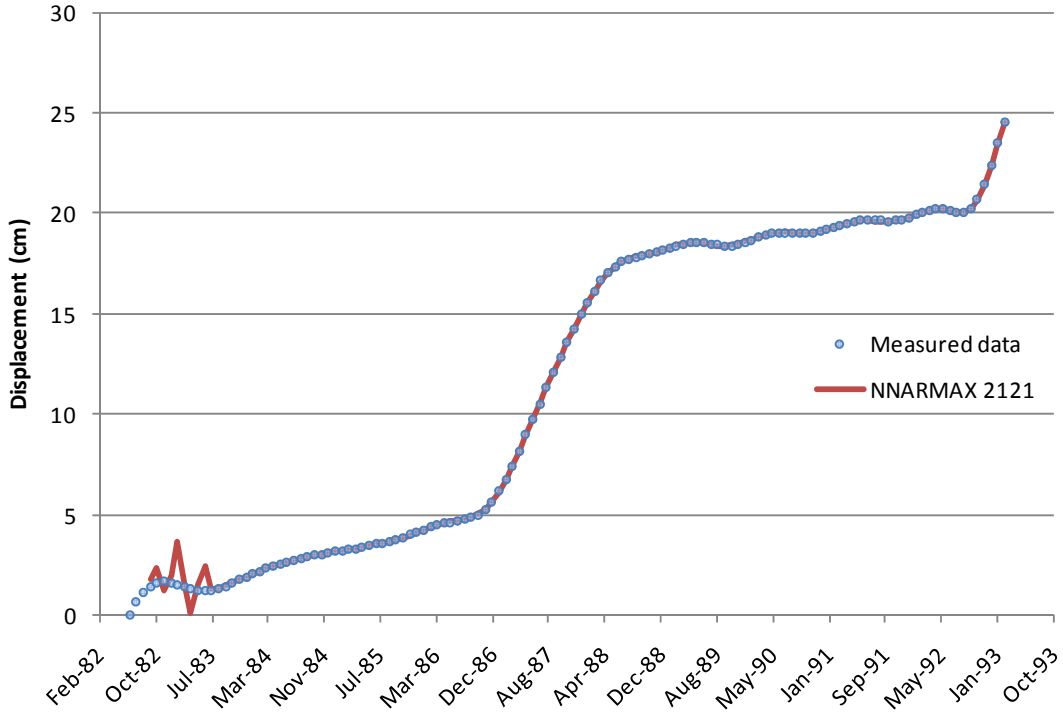


Figure 34. Model NNARMAX 2121: Comparison of measured and predicted values for the training set.

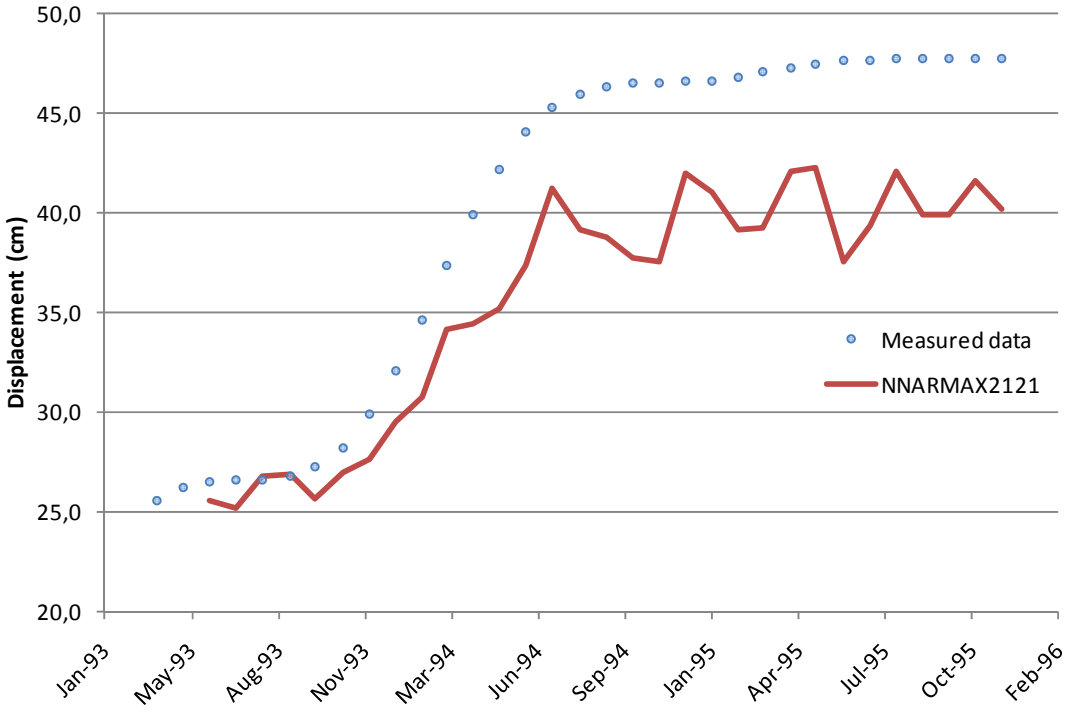


Figure 35. Model NNARMAX 2121: Comparison of measured and predicted values for the validation set.

5.2 BARCELONNETTE (FRANCE)

The present work is focused on rainfall thresholds defined on empirical bases. The objective of this work is to analyze the relationships between two types of mass movements (shallow soil slides and debris flows) and different patterns of rainfall at different temporal scales.

5.2.1 Antecedent precipitation analysis

Antecedent rainfall analysis is of primary interest for landslide hazard assessment; in fact antecedent rainfall controls groundwater levels and soil moisture conditions which are two of the main triggering factors of slope failures (Crozier 1986; Wieczorek 1996). Antecedent precipitation can be used in some cases to determine when landslides are likely to occur.

When using antecedent rainfall measurements to predict landslide occurrence, a key difficulty is the definition of the period over which to accumulate precipitation. Usually, such analyses reveal a significant scatter according to the period in consideration (e.g., Guzzetti et al. 2007).

In this work, several periods have been investigated: 180, 150, 120, 90, 60, 30, 14, 7, 6, 5, 4, 3 and 2 days before the event. Using the daily rainfall and the antecedent rainfall for a 30-day period, four main types of rainfall event can be distinguished (Figure 36):

(1) **Type A:** this situation is characterised by heavy daily rainfall (recorded by most of the rain gauges of the Barcelonnette basin) following a 30-day dry period. For most of the observed events, this climate situation corresponds to violent generalised summer storms;

(2) **Type B:** this situation exhibits the same pattern as Type A, except that the rainfall event has been recorded in a single rain gauge while the neighbouring rain gauges have little or no precipitation. In this case, the summer storm is concentrated in a small area (typically a single crest);

(3) **Type C:** this situation is characterised by heavy cumulative rainfall distributed over a very rainy period of 30 days. This climate situation characterises either the progressive saturation of the topsoil, the rising of a permanent groundwater table and the build-up of positive pore pressures. Cumulative rainfall over a short period is sometimes enough to trigger a debris flow, even though little rainfall is observed on the date of occurrence;

(4) **Type D:** this situation characterises absence of rainfall events for a given day of occurrence of a landslides or a debris-flow event. This can be related to two main explanations: (a) the triggering is due to a rapid snow melt without any liquid rain, (b) the rainfall occurs as a localised hailstorm not recorded by the rainfall station.

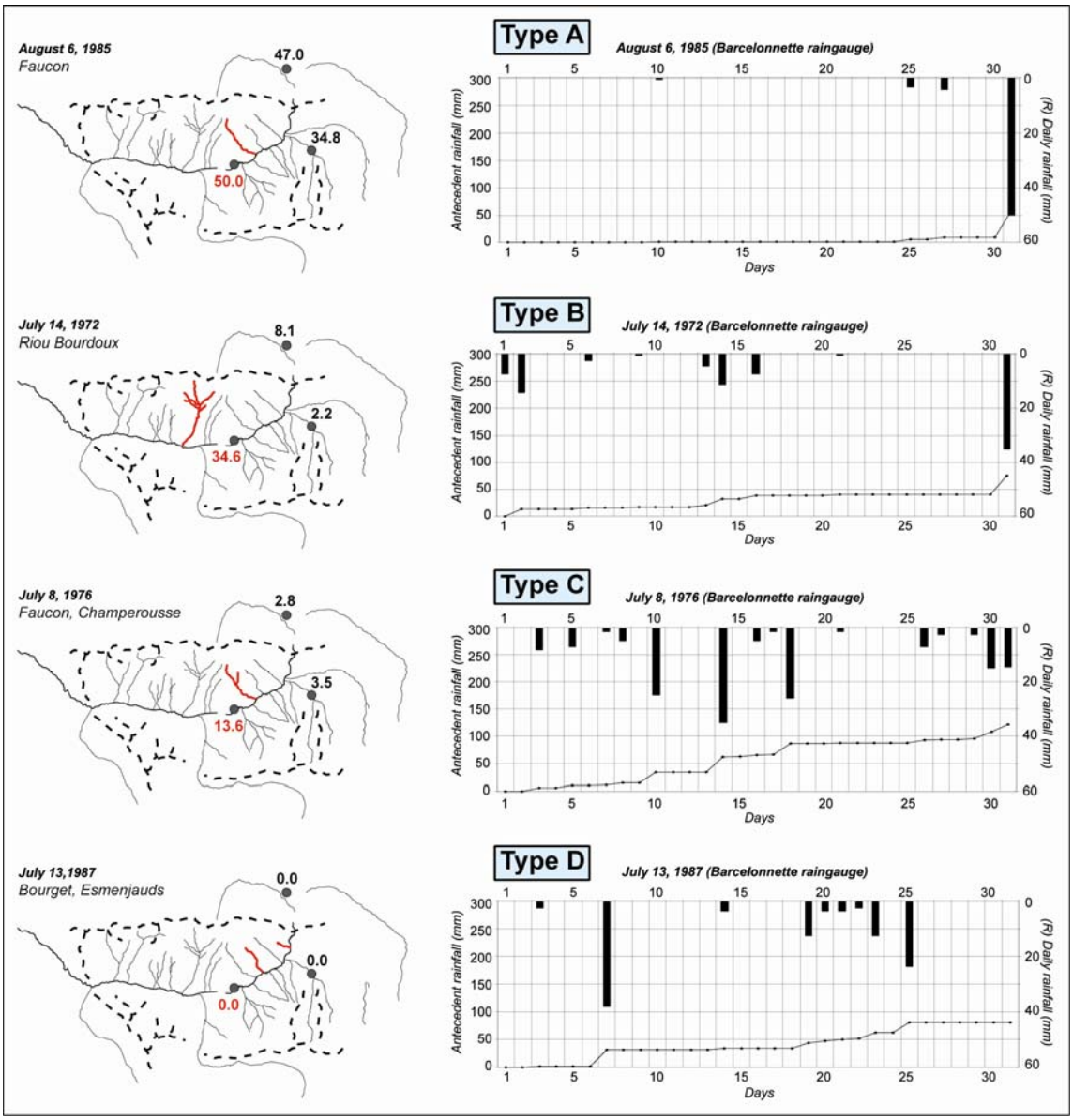


Figure 36. Four different types of rainfall pattern defined on the base of daily rainfall and antecedent rainfall (30 days).

Considering the two main types of mass movements, the rainfall patterns are quite different for soil slides and debris flows. Type A and B are mostly associated to debris flow events, while Type C is more or less associated to the occurrence of soil slide events. Figure 37 shows all the considered antecedent rainfall periods for all the events. A clear distinction can be observed between soil slides and debris flows. Most of the debris flows occurs after a 7-day dry period while soil slide triggering or reactivation needs at least 40 mm of antecedent precipitation during the last 7 days (Figure 37). For a larger time scale, results show that debris flows never occurred when the antecedent rainfall for the 180-day period exceeds 410

mm. On the contrary, antecedent rainfall for a 180-day period has to be at least 205 mm in order to trigger a soil slide.

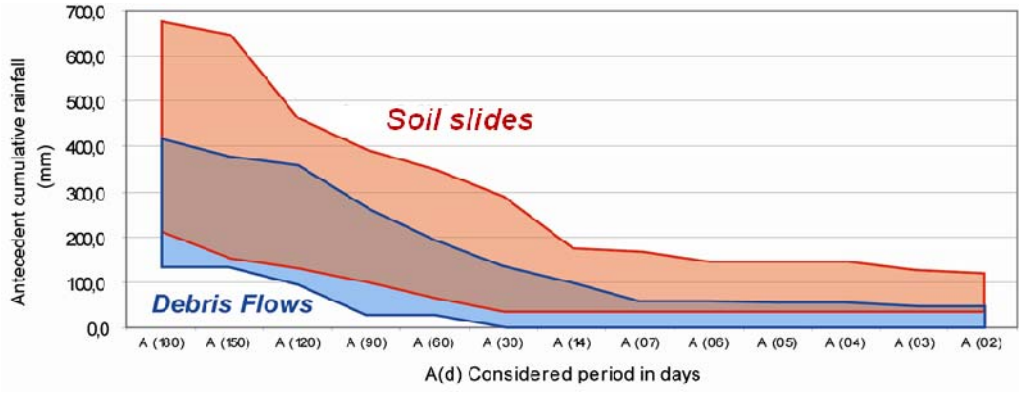


Figure 37. Antecedent cumulative rainfall for debris flows and soil slides in the Barcelonnette Basin for different periods: 180, 150, 120, 90, 60, 30, 14, 7, 6, 5, 4, 3 and 2-days periods before the event.

Although, the correlation between events and the rainfall of the triggering date is quite good, no threshold can be established due to the inaccurate knowledge on rainfall variability, especially in the upper part of hillslopes.

5.2.2 I-D model

5.2.2.1 Visual classification

Twelve events (six debris flows and six soil slides) occurred after 1998; for these events, hourly rainfall data were available from the Barcelonnette station. All the available Intensity-Duration (ID) data was plotted in a single graph (Figure 38):

(1) For debris flows, the rain intensity ranges from 2.7 to 30 $mm.h^{-1}$, and the total duration of the event ranges from 1 to 7 hours;

(2) For soil slides, the rain intensity ranges from 0.5 to 1.8 $mm.h^{-1}$, and the total duration of the event ranges from 8 to 30 hours.

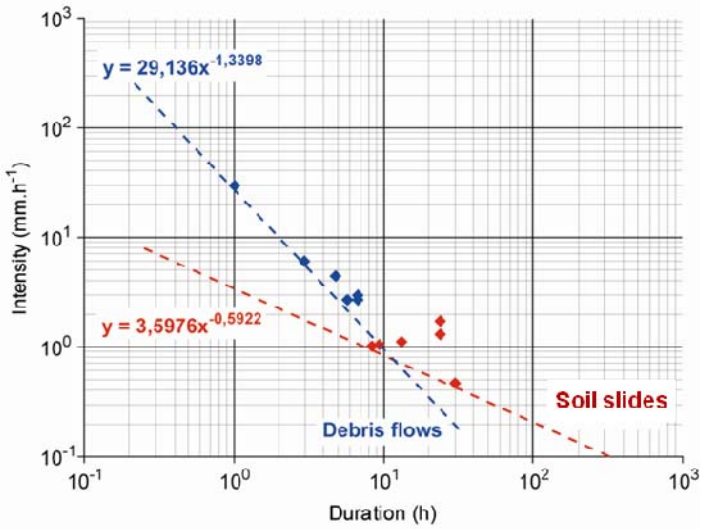


Figure 38. Rainfall intensity duration (ID) for 12 events in the Barcelonnette basin: blue points correspond to debris flow events, red points correspond to soil slide events.

For the purposes of threshold estimation, and in order to better constrain the threshold within a range of durations, some authors (e.g., Moody and Martin 2001; Godt et al. 2006; Cannon et al. 2008) have suggested that storms should be characterised using both mean intensities and also peak intensities for a selection of durations. Such analyses have been performed for the twelve mass movement events (Figure 39).

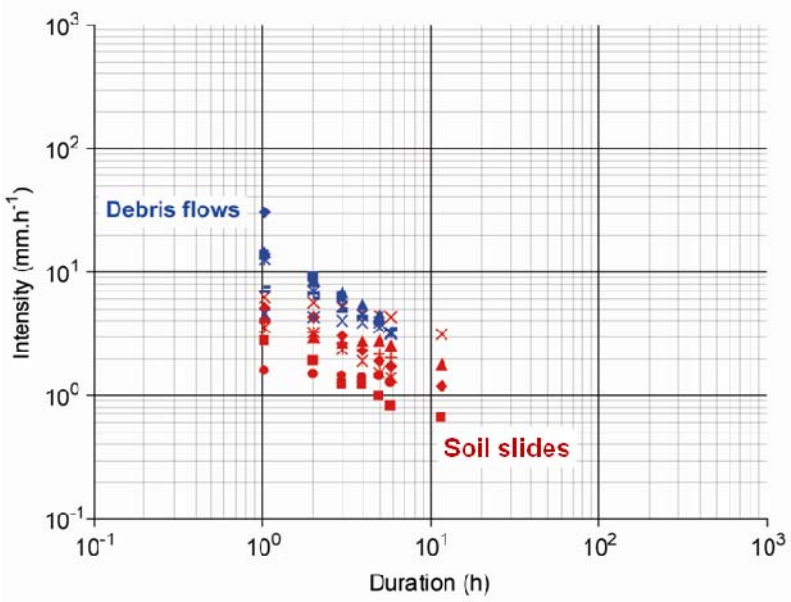


Figure 39. Peak intensities for a selection of duration (12, 6, 5, 4, 3, 2 and 1-hour period) for 12 events in the Barcelonnette basin: blue points correspond to debris flow events, red points correspond to soil slide events.

From a general point of view, the general shape of the curves are quite similar, and show a clear descending trend of the minimum level of average rainfall intensity with increasing rainfall duration. Plotted peak intensities show a similar trend than Figure 37 with a clear distinction between rainfall patterns for debris flows and soil slides:

- Events characterised by high rainfall intensity (average and peak) and short episode duration (e.g., mostly the result of localized convective (hail)storms) will trigger mostly debris flows and shallow soil slides in relatively permeable soils (e.g., moraines, scree or poorly sorted slope deposits);
- Long rainfall periods characterised by low to moderate average and peak rainfall intensity (e.g., the result of multiple and successive storms during a period of several weeks or months) can trigger or reactivate shallow and deep-seated landslides in low permeability soils and rocks (e.g., black marls).

It is possible to assess the performance of these thresholds by analysing all rainfall events (i.e., both, the landslide triggering and the non-triggering). The hourly data consists of 874 rainfall events. The classification statistics (see Table 4 and Equations 13 and 14 for definitions) of the thresholds estimated in Figure 38 are shown in Table 10.

Table 10. Classification statistics for intensity-duration thresholds.

Parameters of classification statistics →	*FP* (false alarms)	*TN*	*FPR*	*TP* (predicted landslides)	*FN* (missed events)	*TPR*
Soil slide threshold – all events	269	589	0.31	12	4	0.75
Soil slide threshold – soil slides	276	590	0.32	5	3	0.63
Soil slide threshold – debris flows	274	593	0.32	7	0	1.00
Debris flow threshold – all events	107	751	0.12	8	8	0.50
Debris flow threshold – soil slides	113	753	0.13	2	6	0.25
Debris flow threshold – debris flows	109	758	0.13	6	1	0.86

Note: see Table 4 and Equations 13 and 14 for definitions of parameters for classification statistics.

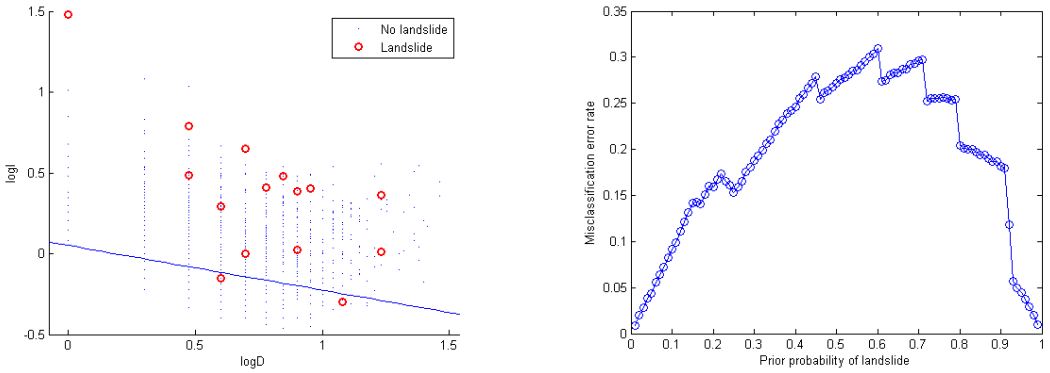
5.2.2.2 Discriminant analyses

Based on the hourly rainfall data (i.e., from 1998), the discriminant analyses classification technique was used for these three datasets and the Intensity-Duration model: (i) all the landslide events (both soil slides and debris flows), (ii) debris flows only, and (iii) soil slides only.

a. All landslide events.

Figure 40b shows the variation of the misclassification error rate vs. the prior probability of landslides in the discriminant analyses of all landslide events using the Intensity-Duration model. The threshold corresponding to 90% prior probability of landslides is shown in Figure 40a. The corresponding equation is:

$$I = 1.125D^{-0.277} \quad (56)$$



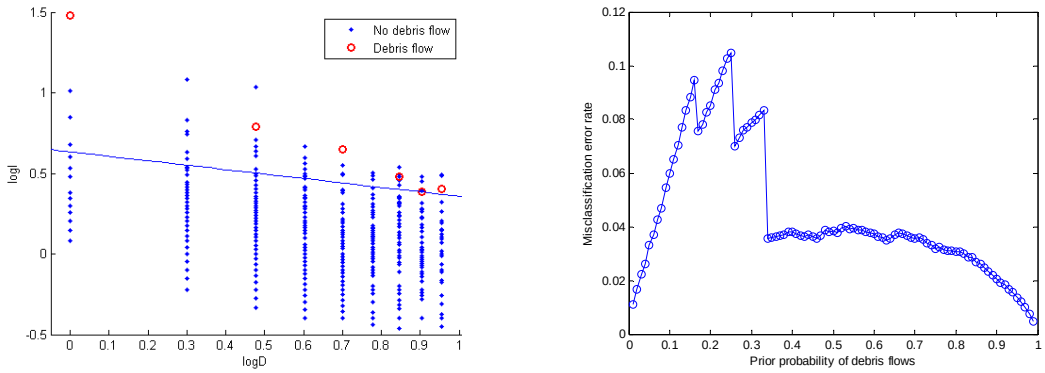
(A) (B)

Figure 40. Results of discriminant analyses using all landslide events (soil slides and debris flows) at Barcelonnette. (A) Intensity-Duration plot showing the threshold for a 90% prior probability of landslides and the rainfall events (blue dots and red circles are non-triggering and landslide-triggering rainfall events, respectively). (B) Misclassification error rate vs. prior probability of landslides.

b. Debris flows.

Figure 41b shows the variation of the misclassification error rate vs. the prior probability of debris flows in the discriminant analyses using the Intensity-Duration model. The threshold corresponding to 34% prior probability of debris flows is shown in Figure 41a. The corresponding equation is:

$$I = 4.297D^{-0.275} \quad (57)$$



(A) (B)

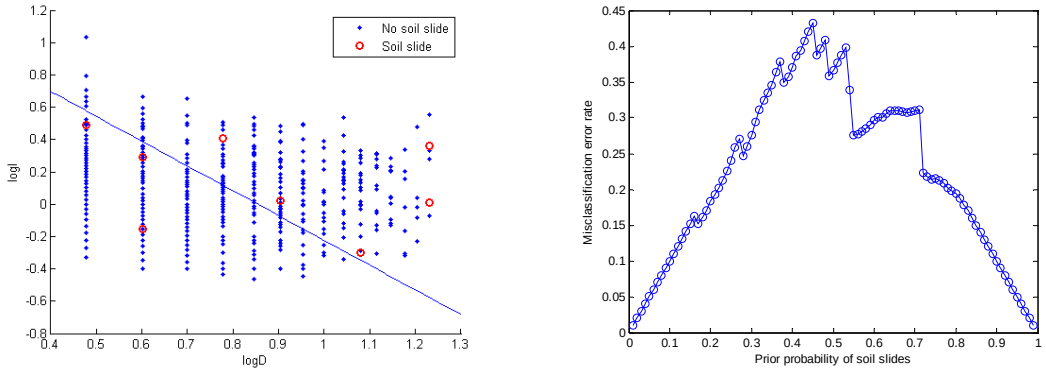
Figure 41. Results of discriminant analyses for debris flow events at Barcelonnette considering only durations ranging from 1 to 9 hours. (A) Intensity-Duration plot showing the threshold for a 34% prior probability of debris flows and the rainfall events (blue dots and red circles are non-triggering and debris flow-triggering rainfall events, respectively). (B) Misclassification error rate vs. prior probability of debris flows.

c. Soil slides.

Figure 42b shows the variation of the misclassification error rate vs. the prior probability of soils slides in the discriminant analyses using the Intensity-Duration model. The threshold corresponding to 50% prior probability of soil slides is shown in Figure 42a. The corresponding equation is:

58

$$I = 20.33D^{-1.53}$$



(A) (B)

Figure 42. Results of discriminant analyses for soil slide events at Barcelonnette considering only durations ranging from 3 to 17 hours. (A) Intensity-Duration plot showing the threshold for a 50% prior probability of soil slides and the rainfall events (blue dots and red circles are non-triggering and soil slide-triggering rainfall events, respectively). (B) Misclassification error rate vs. prior probability of soil slides.

A set of parameters in the Intensity-Duration model were estimated by varying the prior probabilities of soil slides from 0.01 to 0.99. These parameters were used to obtain the corresponding ROC curve for the Intensity-Duration model, which is shown in Figure 43. The area under this ROC curve, AUC, is 0.6833, and the minimum distance to the point of perfect classification (0,1) is 0.4608.

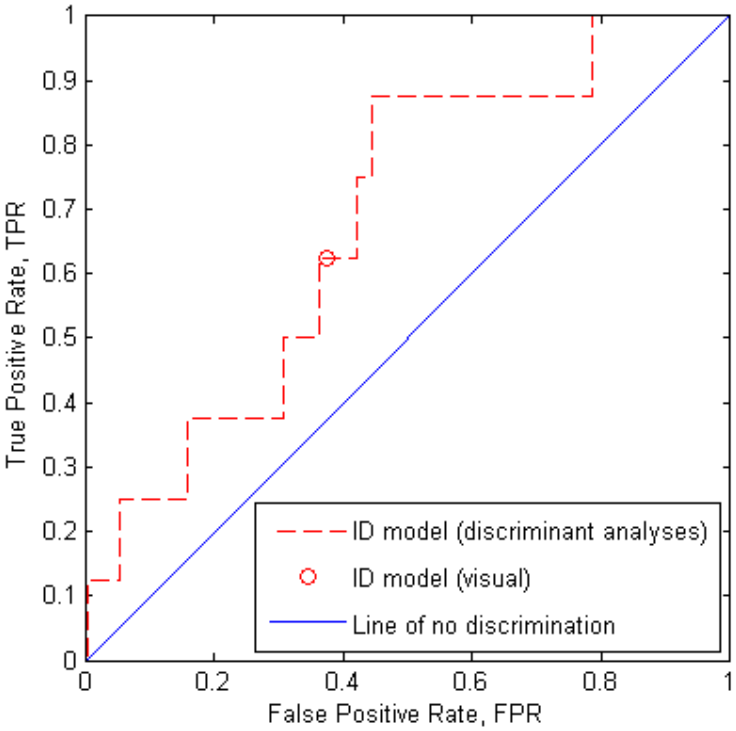
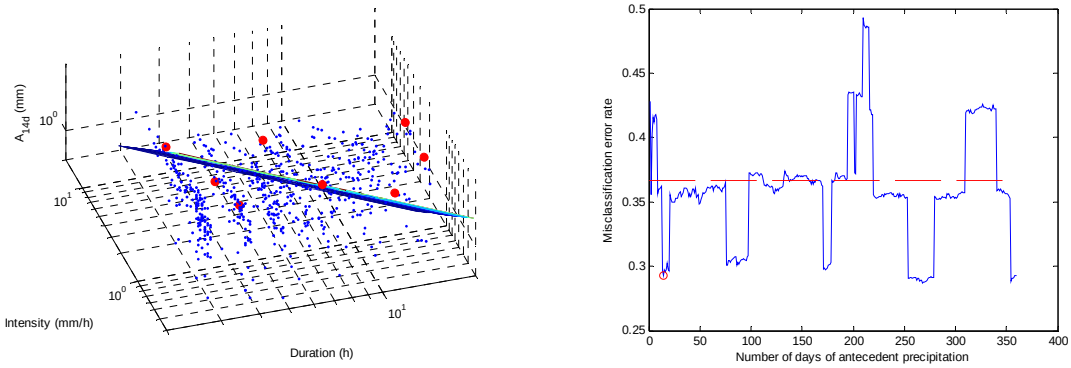


Figure 43. Receiver Operating Characteristic (ROC) curve for the Intensity-Duration model obtained using discriminant analyses. The circle marker shows the classifier for the visual estimate of the model as calculated in section 5.2.2.1.

5.2.3 I-A-D model

The Intensity-Duration threshold equation for debris flows (Equation 56) was considered satisfactory with a misclassification error rate lower than 0.04. Furthermore, the power law form of the Intensity-Duration model seems to be adequate for characterisation of triggering conditions for debris flows (e.g., see section 2.4). However, the corresponding threshold for soil slides produced a relatively low performance, evidenced by misclassification error rates larger than 0.35 for 50% of prior (Figure 42b). An improvement in the threshold for soil slides was attempted by using the Intensity-Antecedent Precipitation-Duration model presented in section 2.5.

Applying a discriminant analysis to the hourly dataset from Barcelonnette, and using a prior probability of 50%, the variation of the misclassification error rate vs the number of days of antecedent precipitation (n) was estimated. The result is shown in Figure 44b. It can be noted that the minimum misclassification error rate occurs for n = 14 days. The corresponding threshold is shown in Figure 44a.



(A) (B)

Figure 44. Results of discriminant analyses for soil slides events at Barcelonnette considering only durations ranging from 3 to 17 hours and using the I-A-D model. (A) Intensity-Antecedent precipitation-Duration (I-A-D) plot showing the threshold for a 50% prior probability of soil slides and the rainfall events (blue dots and red circles are non-triggering and soil slide-triggering rainfall events, respectively). (B) Misclassification error rate vs. number of days of antecedent precipitation (for a 50% prior probability of soil slides). The red dashed line is the misclassification error rate for the corresponding I-D model.

The equation of the threshold for the I-A-D model for n = 14 days (Figure 44a) is:

$$I = 53.82 A_{14d}^{-0.3340} D^{-1.5092} \quad (59)$$

A more comprehensive assessment of the performance of the I-A-D model can be obtained by estimating ROC curves for several values of n. In this case, n = 1, 2, 3, … , 360 days. The ROC curve for each case was obtained by varying the prior probabilities from 0.01 to 0.99 using 0.01 increments. For each ROC curve, the area AUC and the minimum distance to the perfect classifier (FPR = 0, TPR = 1) were obtained. The variations of AUC and minimum distance to perfect classificer vs n are shown in Figure 45 and Figure 46, respectively.

An overall perspective of the performance of the I-A-D model for varying values of n can be made by combining Figure 45 and Figure 46. The result is presented in Figure 47. It can be seen that values of n between 20 and 30 days have the best performance in terms of AUC, while n between 170 and 175 days are the best in terms of distance to the perfect classifier. The parameter AUC is given priority for qualifying the performance of the model, so a threshold model for n = 23 days was selected.

A comparison of the ROC curves for the ID model and the IAD model (for n = 23 days) is presented in Figure 48.

The IAD threshold for n = 23 days and for the set of parameters that produces the minimum distance to the perfect classifier is shown in Figure 49. The equation of this threshold plane is:

$$I = 181.2 A_{23d}^{-0.6788} D^{-1.5163} \quad (60)$$

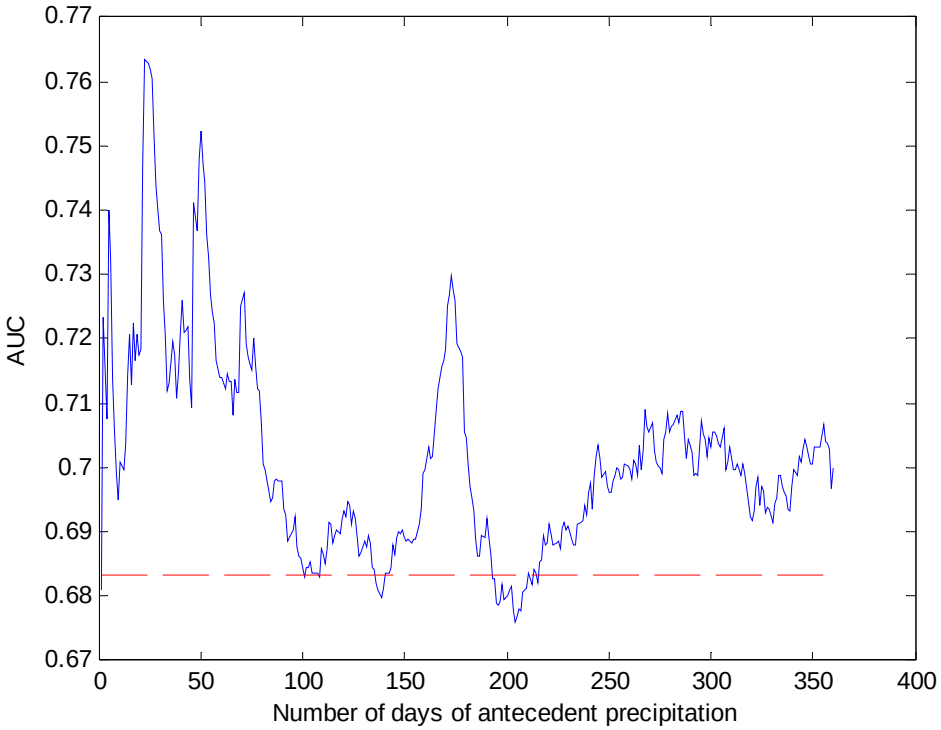


Figure 45. **Results of discriminant analyses for soil slides at Barcelonnette considering only durations ranging from 3 to 17 hours and using the I-A-D model: area under the ROC curve (AUC) vs. number of days of antecedent precipitation. The red dashed line is the value of AUC for the I-D model (section 5.2.2.2 and Figure 43). The performance of the model improves as AUC increases.**

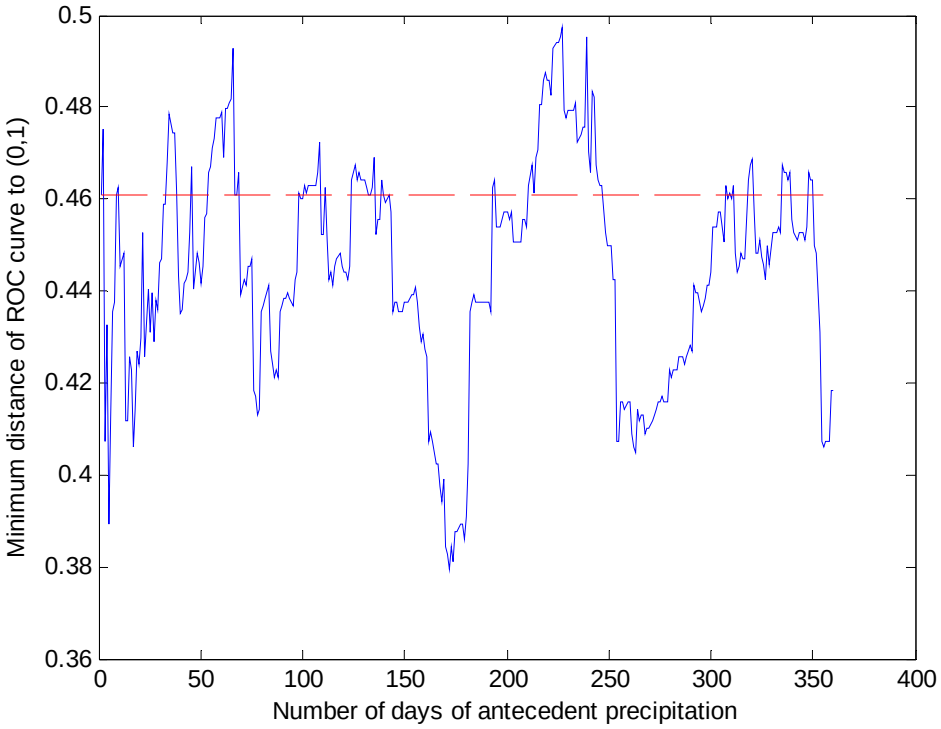


Figure 46. **Results of discriminant analyses for soil slides at Barcelonnette considering only durations ranging from 3 to 17 hours and using the I-A-D model: minimum distance of ROC curve to perfect classifier (FPR = 0, TPR = 1) vs. number of days of antecedent precipitation. The red dashed line is the value of minimum distance for the I-D model (section 5.2.2.2 and Figure 43). The performance of the model improves as the distance decreases.**

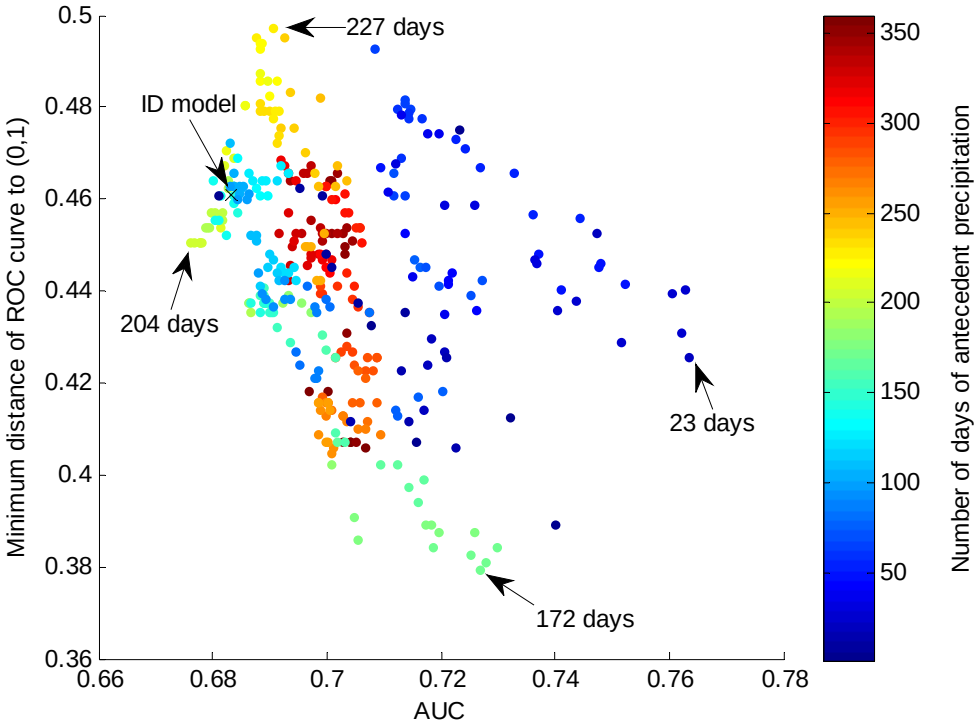


Figure 47. **Results of discriminant analyses for soil slides at Barcelonnette considering only durations ranging from 3 to 17 hours and using the I-A-D model: area under the ROC curve vs. minimum distance to perfect classifier (FPR = 0, TPR = 1). The colours of the markers correspond to *n*, (number of days for antecedent precipitation) and are explained by the colour bar to the right. For comparison, the performance of the I-D model (section 5.2.2.2) is represented by the X marker.**

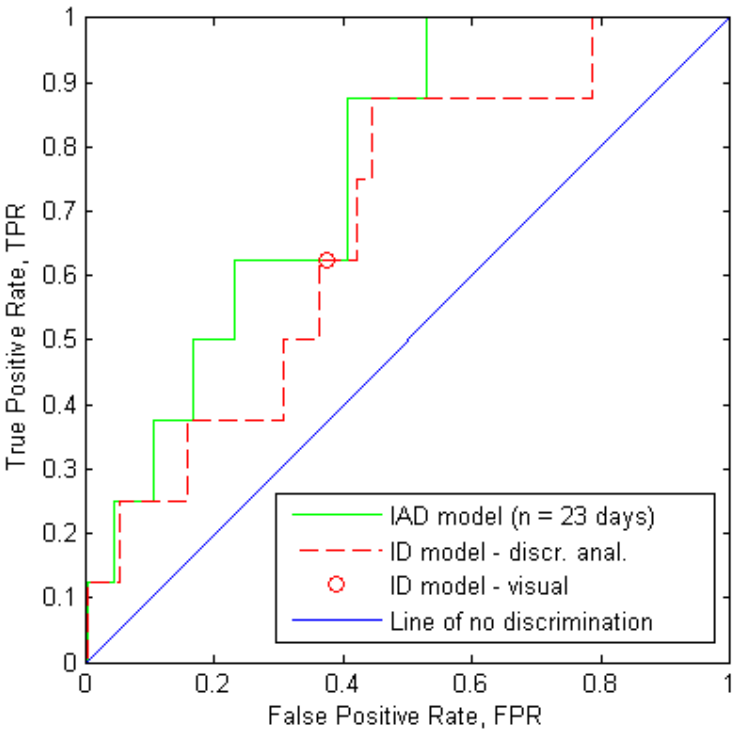


Figure 48. Comparison of ROC curves for ID model and for IAD model (*n* = 23 days).

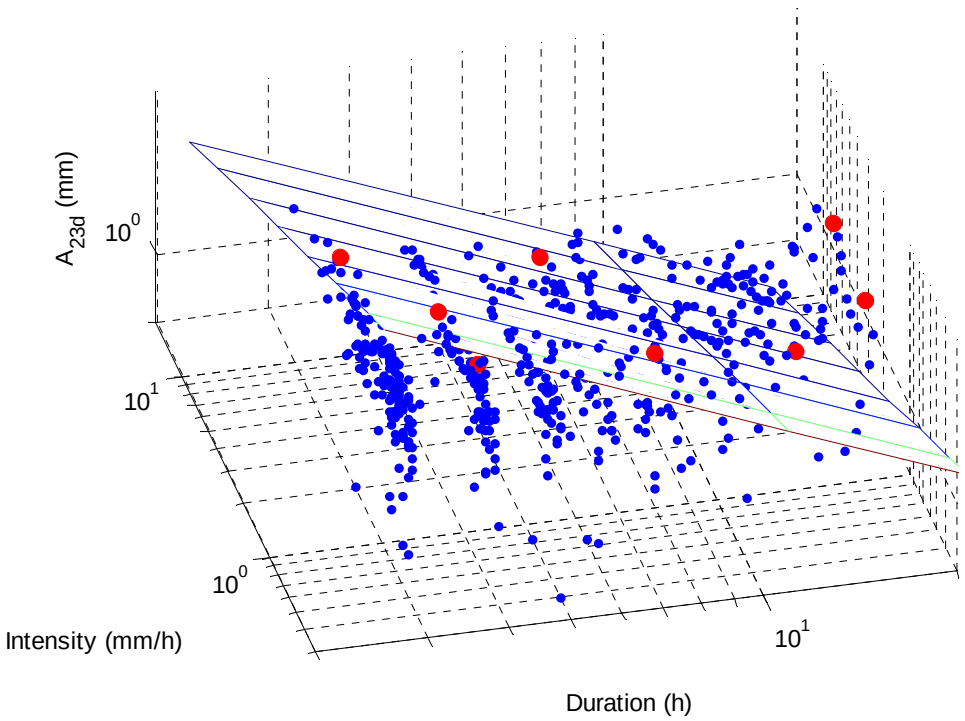


Figure 49. Intensity-Antecedent Precipitation-Duration (IAD) threshold for soil slides in Barcelonnette for *n* = 23 days of antecedent precipitation. The blue and red markers are non-triggering and triggering precipitation events, respectively.

5.2.4 FLaIR model

The FLaIR model has been applied to Barcelonnette data by:
a) Selection of landslide events;
b) FLaIR calibration and evaluation.

From the whole dataset of historical movements, only the records associated to landslide events (glissement de terrain) were selected. Moreover, the period 1950-1980 was considered for FLaIR calibration, while the period 1981-2009 was used for model validation.

Selection of landslide events

As an initial step, for both the calibration and validation periods, a landslide event was chosen from the historical database only if the field "CAUSE" was rainfall.

In a second step, the following operations were carried out for every landslide movement:

1. The start date related to the associated rainfall event is identified;
2. For each day between the start date of a rainfall event and the landslide occurrence, the antecedent cumulated rainfalls, with durations ranging from 1 to 30 days, are evaluated;
3. For each duration, the maximum rainfall height is computed;
4. For each one of the durations, the values estimated on step 3 are compared with the corresponding mean values of the annual maximum.

A movement date was definitively selected if, for all the considered durations, the maximum rainfall heights estimated at point 3 were greater than 50% of the mean values of the annual maximum rainfall corresponding to the same durations (point 4). The aim is to select only events with similar characteristics and that have been triggered by rainfall.

Table 11 shows all the landslide events, both the included in the analyses (black font) and those that were excluded (red font). After the selection, 6 events were within the calibration period and one in the validation period.

Figure 50 shows the daily rainfall histogram. The precipitation events on landslide days are indicated by red bars.

Table 11. List of landslide events (black font for events included in the analyses, and red font for those excluded).

		Maximum rainfall heights (mm)				
Landslide Date	**Start date of rainfall event**	**1 day**	**5 days**	**10 days**	**15 days**	**30 days**
CALIBRATION PERIOD						
13/3/1951	10/3/1951	49.8	87.6	95.1	95.1	176
26/5/1956	22/05/1956	13.1	27.2	38.7	40.3	45.2
18/10/1960	05/10/1960	36.8	78.7	106.1	120.9	237.4
12/9/1964	12/9/1964	13.1	13.1	37.8	43.5	64.6
2/3/1966	21/02/1966	40.5	62.2	72.8	81.3	94.1
21/3/1971	18/3/1971	52.6	109.5	140	140	140
22/10/1977	07/10/1977	43.3	106.6	107.8	115.3	144
14/1/1978	11/01/1978	29.7	86.9	86.9	86.9	95.8
VALIDATION PERIOD						
21/6/1983	21/6/1983	3.6	16.3	16.3	20.3	38.6
12/3/1986	12/3/1986	13.1	13.1	13.1	13.2	70.6
9/3/1994	27/02/1994	24.8	39	44.3	43.3	56.4
10/11/2000	06/11/2000	32.6	64	80.8	80.8	185
	50% of annual maximum expected value	**25.10**	**44.15**	**55.35**	**65.68**	**86.20**

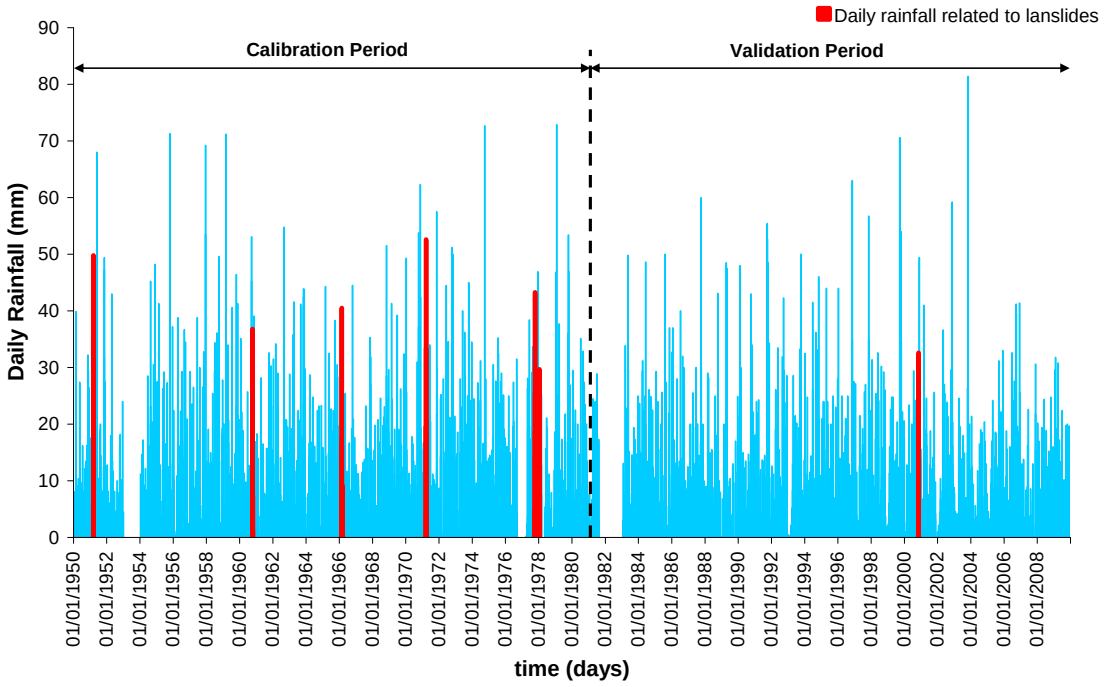


Figure 50. Barcelonnette: daily rainfall histogram.

FLaIR calibration and evaluation

A threshold P_0 (section 2.7) equal to 30 mm for daily rainfall was applied (this value is close to the minimum value of the daily precipitation for the considered events in Table 11).

The Gamma function was used as the transfer function $\psi(.)$:

$$\psi(t) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} t^{\alpha-1} e^{-\frac{t}{\beta}} \qquad \alpha > 0 \quad \beta > 0 \qquad (61)$$

For calibration, the ranking technique (section 2.7) was used. The minimum value of false alarms with no missed alarms was obtained for $\alpha = 1.1$ and $\beta = 1$ day, for which FA = 23 and $Y_{F,cr} = 21.92$. So, according to section 3.2, this set of parameters was adopted.

Figure 51 shows the transfer function $\psi(.)$, and Figure 52 shows the associated ROC curve, which was evaluated by setting $Y_{F,cr}$ equal to the $Y_F()$ values estimated in the days of landslide events (Table 12) and then by assessing for each threshold the metrics FA and HR. The corresponding AUC has a maximum value equal to 0.953. Figure 53 shows the computed mobility function.

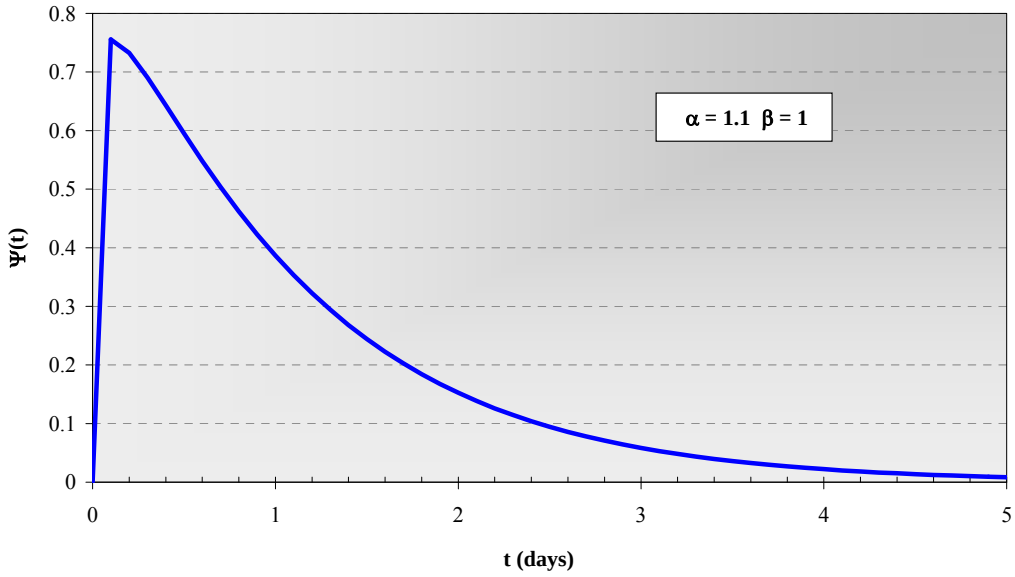


Figure 51. Barcelonnette: transfer function of the FLaIR model.

Table 12. Calibration period: values of the mobility function and metrics of the ROC space for the historical events.

Date→	**13/03/1951**	**18/10/1951**	**02/03/1966**	**21/03/1971**	**22/10/1977**	**14/01/1978**
$Y_{F,\,cr}$	**22.08**	**22.23**	**21.92**	**26.37**	**24.13**	**23.21**
FA	23	20	23	4	13	18
MA	1	2	0	5	4	3
CA	5	4	6	1	2	3
TN	294	297	294	313	304	299
FAR	0.073	0.063	0.073	0.013	0.041	0.057
HR	0.833	0.667	1.000	0.167	0.333	0.500

Regarding the validation period, the threshold $Y_{F,cr} = 21.92$ provides a correct alarm on the date of the historical event (10/11/2000), when $Y_F = 22.01$ (Figure 54). Moreover, *FA*=16, *TN* = 290 and consequently *FAR*=0.052. This implies that the FLaIR model is characterized by a good performance during the validation period in terms of the *FAR* value.

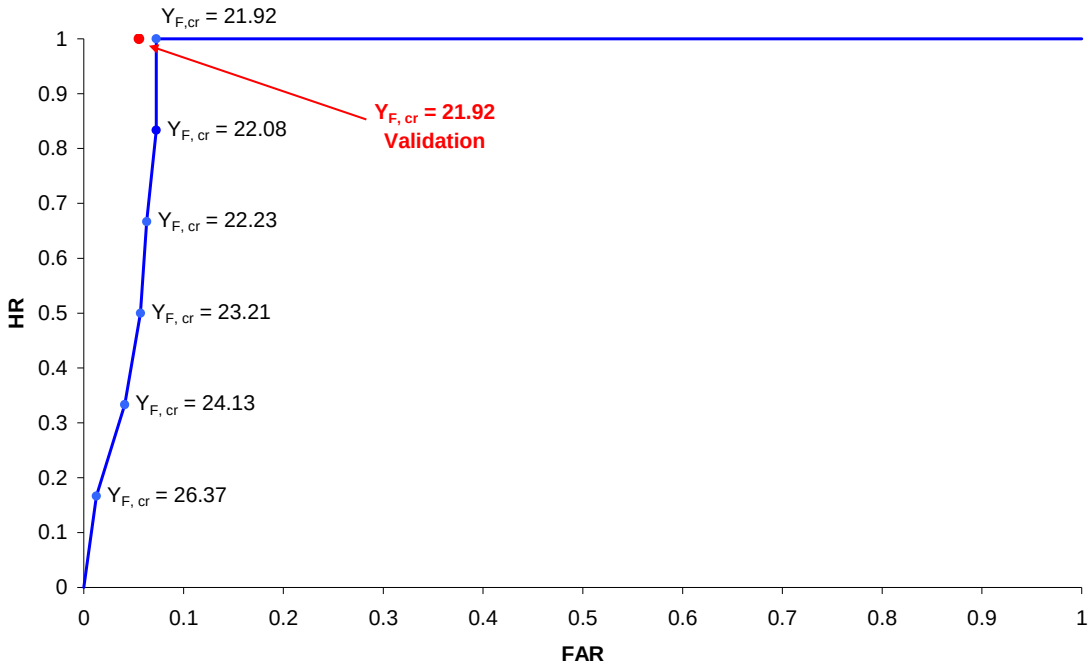


Figure 52. Barcelonnette: ROC curve.

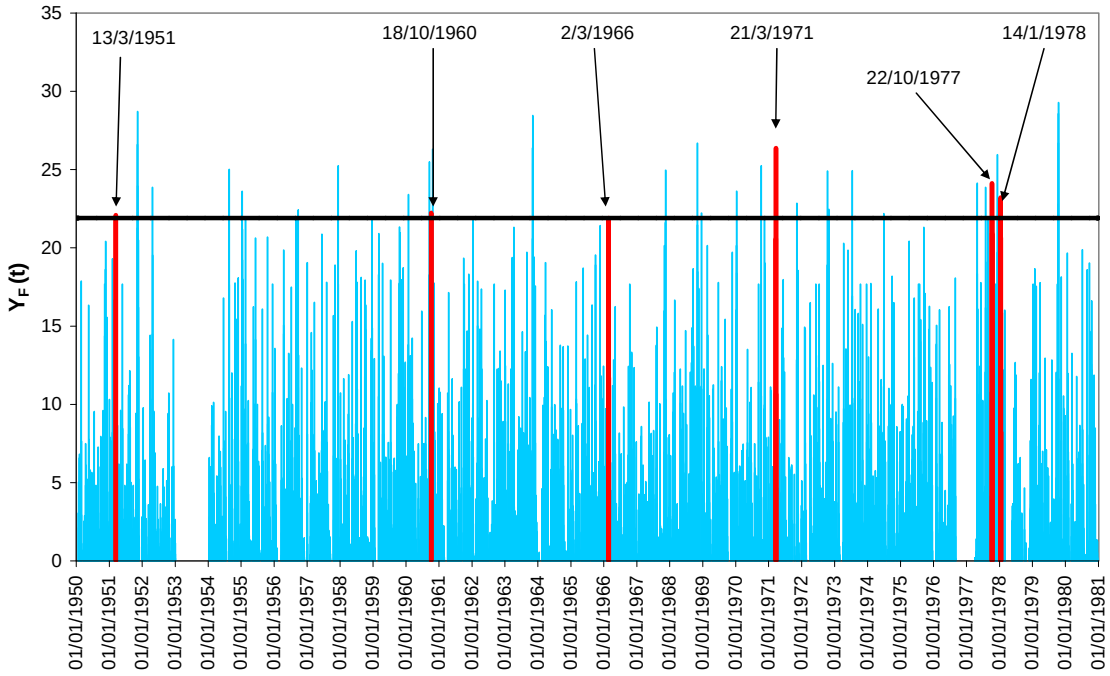


Figure 53. Barcelonnette – calibration period: mobility function ($Y_{F,cr}$ = 21.92).

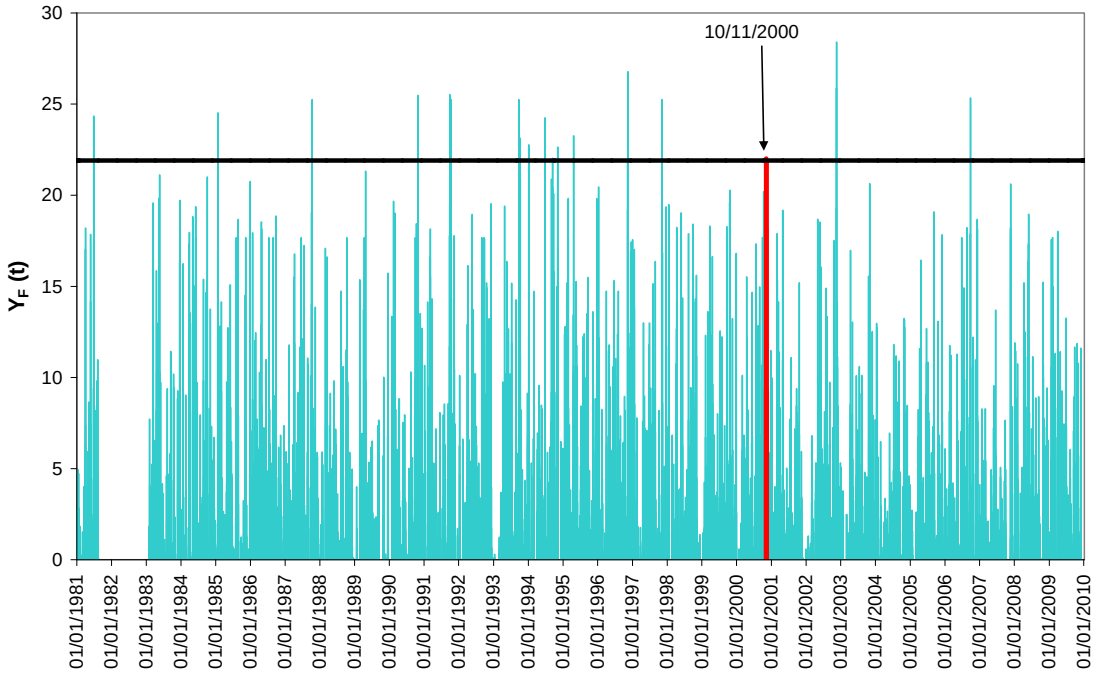


Figure 54. Barcelonnette – validation period: mobility function ($Y_{F,cr}$ = 21.92).

5.3 SOUTH-EASTERN NORWAY

The procedure described in section 2.10 and the time-uncertainty of each event in the dataset that ranged between 1 and 64 days (see section 4.3) were used to compute the distribution of probabilities in time. The resulting distribution is shown in Figure 55. For the purposes of calibration of thresholds, the probabilities in Figure 55 were classified in three ranges: low, medium and high (see Figure 56) with boundaries at 0.1 and 0.5.

Discriminant analyses were performed for antecedent precipitation ranging between 1 day and 360 days using the model described in section 2.5 and Equation 3. The optimum thresholds (minimum misclassification error) were obtained for 46-day antecedent precipitation (see Figure 57). The thresholds were also obtained for three levels of time-certainty: "low", "medium" and "high" (see Figure 56). Additionally, the analyses were performed taking into account variability in the prior probabilities of landslide events. Based on the conceptual framework of thresholds by Cepeda et al. (2010), these prior probabilities can be considered to correspond to susceptibility levels ranging from low (low prior) to high (high prior). The calibrated thresholds are shown in Figure 58 and the threshold values for a 50% prior probability are shown in Figure 56.

Figure 55. Probability of earth slides vs time in the Nedre Romerike area during autumn 2000. The vertical axis is the probability of occurrence of at least one landslide in one day estimated based on the associated time-uncertainty of each event in the landslide inventory.

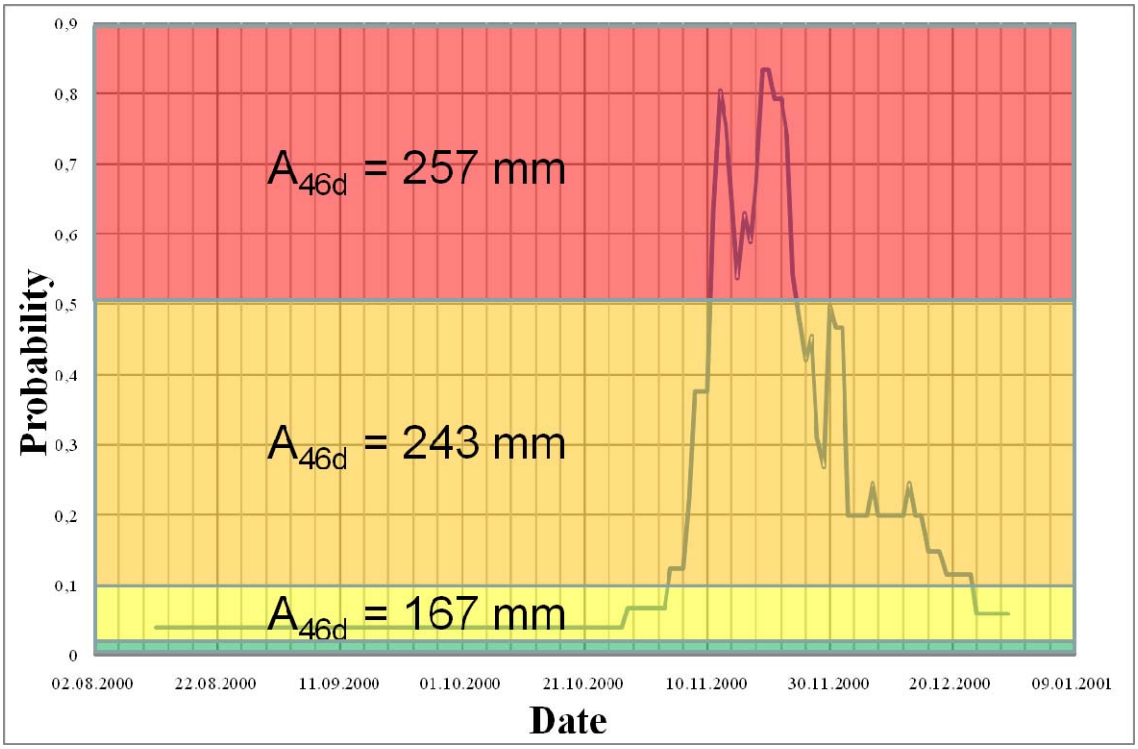


Figure 56. Ranges of probabilities of earth slides vs time in the Nedre Romerike area during autumn 2000. Yellow is low probability, orange is medium and red is high.

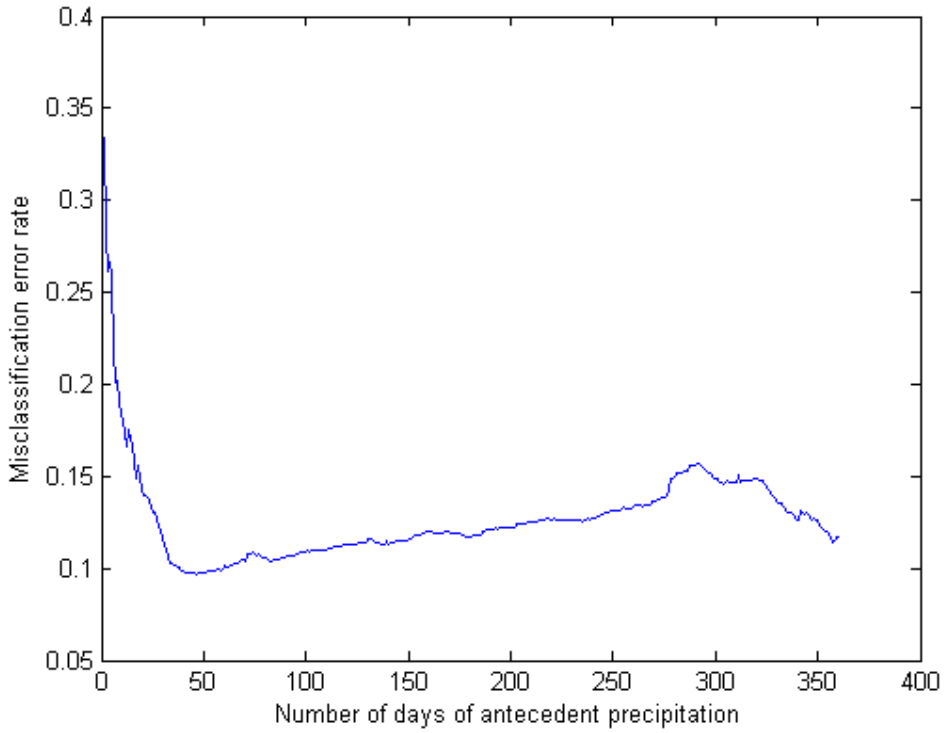


Figure 57. Misclassification error rate for threshold model based on Equation 3 using the dataset from the Nedre Romerike area. The minimum error occurs for 46 days of antecedent precipitation.

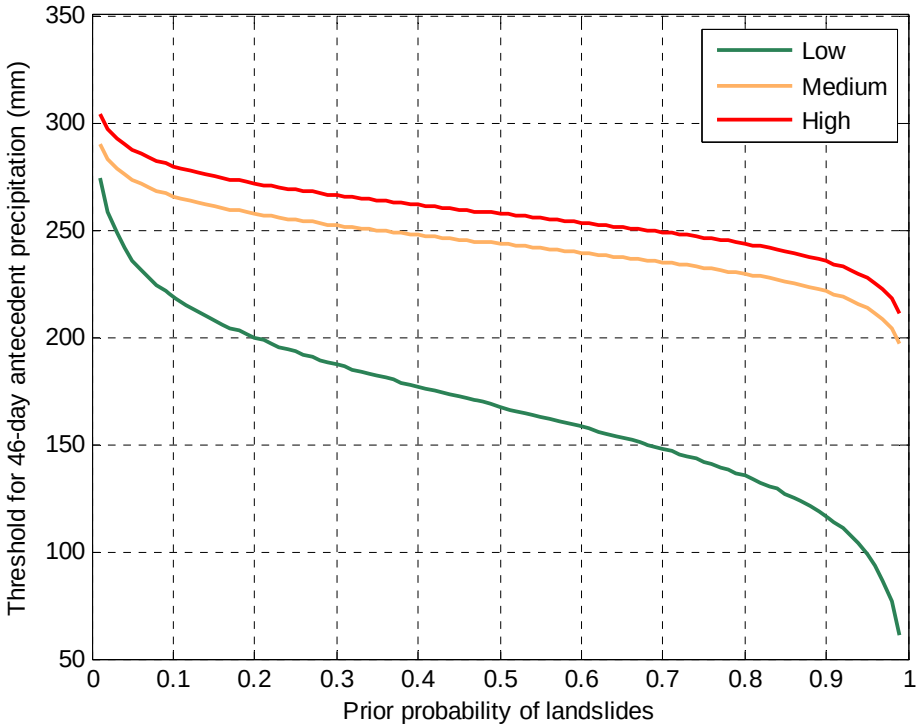


Figure 58. Thresholds for precipitation-induced earth slides in the Nedre Romerike area. The curves "low", "medium", and "high" indicate the time-certainty of landslide occurrence. The horizontal axis (prior probability of landslides) can be calibrated for landslide susceptibility levels (i.e., low prior probability for low susceptibility zones, and high prior probability for high susceptibility zones).

There was also an attempt to use a threshold model as a linear combination of two different numbers of days of antecedent precipitation according to Equation 5 in section 2.5. The values of n and p in Equation 5 were varied from 1 to 360 days. The misclassification error rates for each combination of n and p are shown in Figure 59. The blank area in Figure 59 corresponds to combinations that were highly correlated (R^2 >= 0.8), and thus were not evaluated in the threshold model. A linear threshold for n = 10 days and p = 54 days is shown in Figure 60.

For practical purposes, and considering the relatively low misclassification error rates, it is considered that the antecedent precipitation threshold based on Equation 3 (Figure 56 and Figure 58) is adequate for characterising triggering conditions for the Nedre Romerike area.

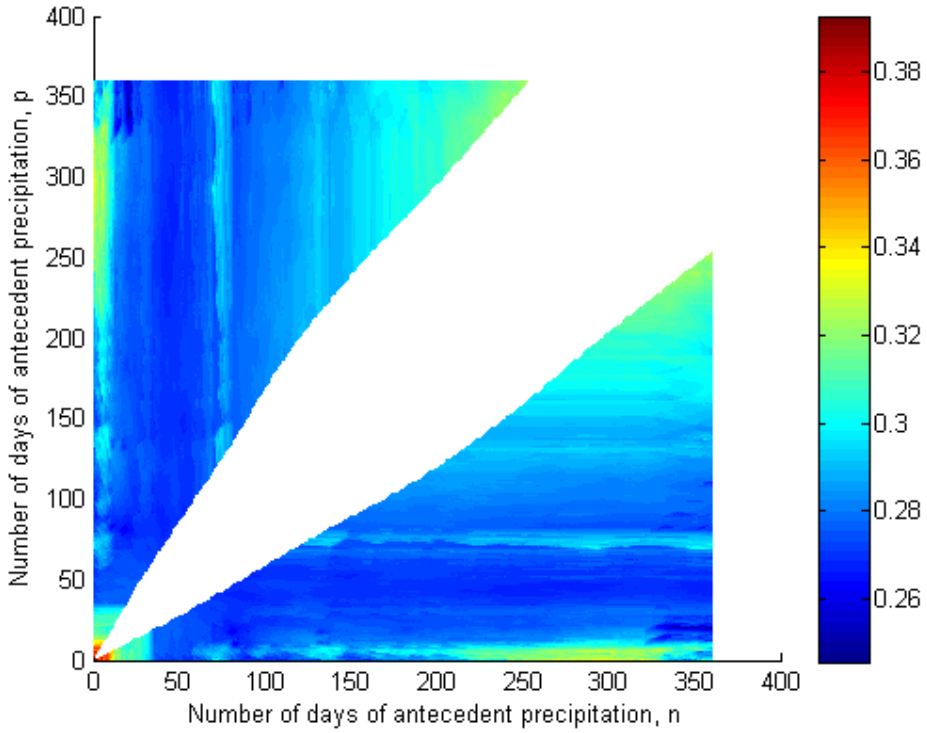


Figure 59. **Misclassification error rates for the Nedre Romerike dataset applying threshold model of Equation 5 and varying *n* and *p* from 1 to 360 days. The colour bar to the right indicates the correspondence between colours and values of errors.**

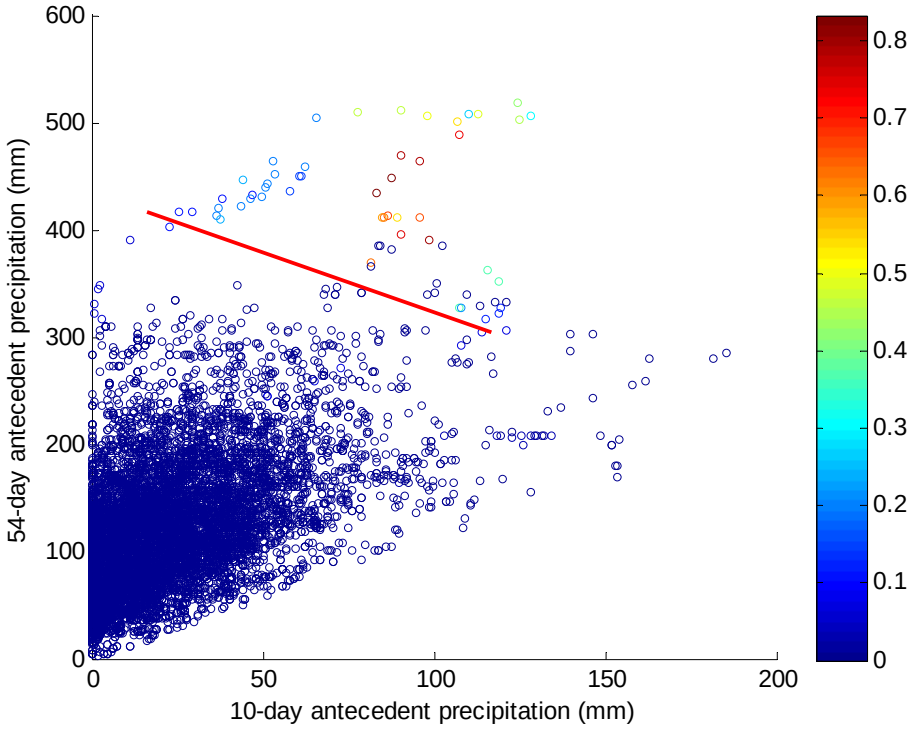


Figure 60. **Linear threshold for the Nedre Romerike area according to Equation 5 for n = 10 days and p = 54 days. The colour bar to the right corresponds to the probability values based on the uncertainty of time occurrence (Figure 55).**

5.4 SATRIANO, VERZINO AND SARNO (ITALY)

The FLaIR model has been applied to the three datasets described in section 4.4:
1) Satriano (Calabria region);
2) Verzino (Calabria region);
3) Sarno (Campania region).

For each study area, datasets with daily rainfall heights and historical movement dates were analyzed, in order to calibrate the model by using the technique reported in sections 2.7 and 3.2. Once the expression of the transfer function $\psi(.)$ is selected, the following results are presented:

a) the feasible region of the parameters, characterized by the maximum value of AUC (section 3.2)
b) the transfer function $\psi(.)$ related to the parameters which maximize $r = f_U - f_L$ (section 3.2)
c) the mobility function.

5.4.1 Satriano

As reported in section 4.4, daily rainfall data of the Chiaravalle Centrale rain gauge was used. The data ranges from 1 January 1951 to 30 June 2000.

The historical movement dates are:
1) 18th October 1951;
2) 2nd January 1973.

In Figure 61 the daily rainfall histogram is shown. The precipitation events that triggered landslides are also indicated. For the historical movements, the maximum rainfall heights, cumulated over antecedent days (from 1 to 30) are reported in Table 13.

Table 13. Satriano: Maximum rainfall heights for the landslide events.

		Maximum rainfall heights (mm)				
Landslide Date	**Start date of rainfall event**	**1 day**	**5 days**	**10 days**	**15 days**	**30 days**
18/10/1951	15/10/1951	436.1	1020.8	1034.6	1369.9	1426.3
2/1/1973	27/12/1972	98.7	395.9	491.4	578.8	883.7

The gamma function is selected as transfer function $\psi(.)$:

$$\psi(t) = \frac{1}{\beta^{a}\Gamma(\alpha)} t^{\alpha-1} e^{-\frac{t}{\beta}} \qquad \alpha > 0 \quad \beta > 0 \tag{62}$$

In this case, equation 62 is characterized by a limited feasible region of the parameters, for which AUC is equal to 1 (Figure 62); for $\alpha = 0.9$ and $\beta = 20$ days, the difference $r = f_U - f_L$ assumes its maximum value, equal to 0.02, and the associated $\psi(.)$ is in Figure 63. Figure 64

shows the mobility function; $Y_F(.)$ is 64.04 on 18 October 1951 and 31.71 on 2 January 1973; consequently, $Y_{F,cr}$ is set equal to 31.71.

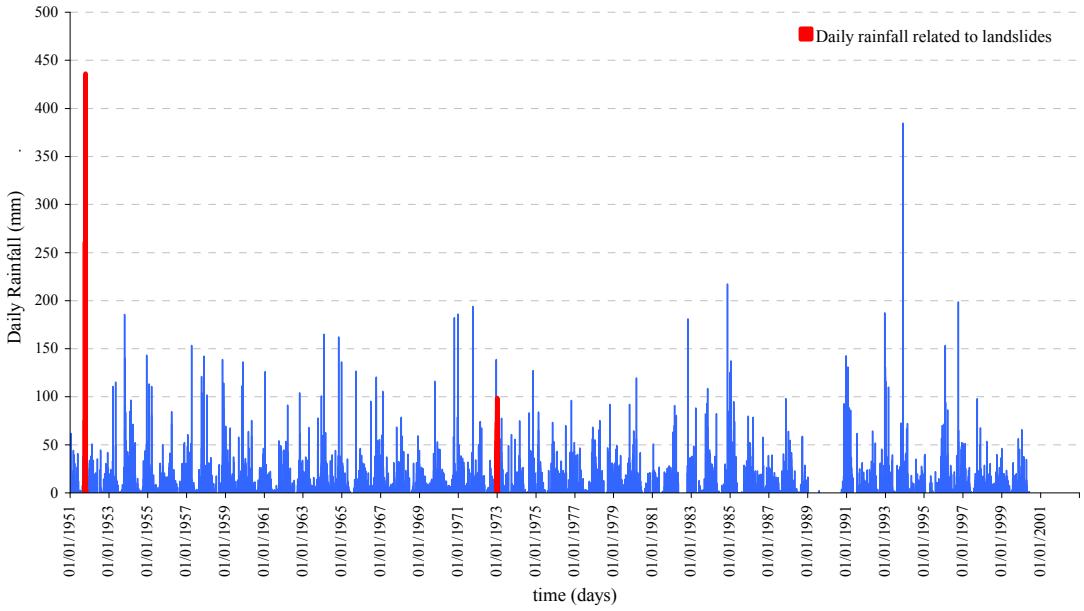


Figure 61. Satriano: daily rainfall histogram.

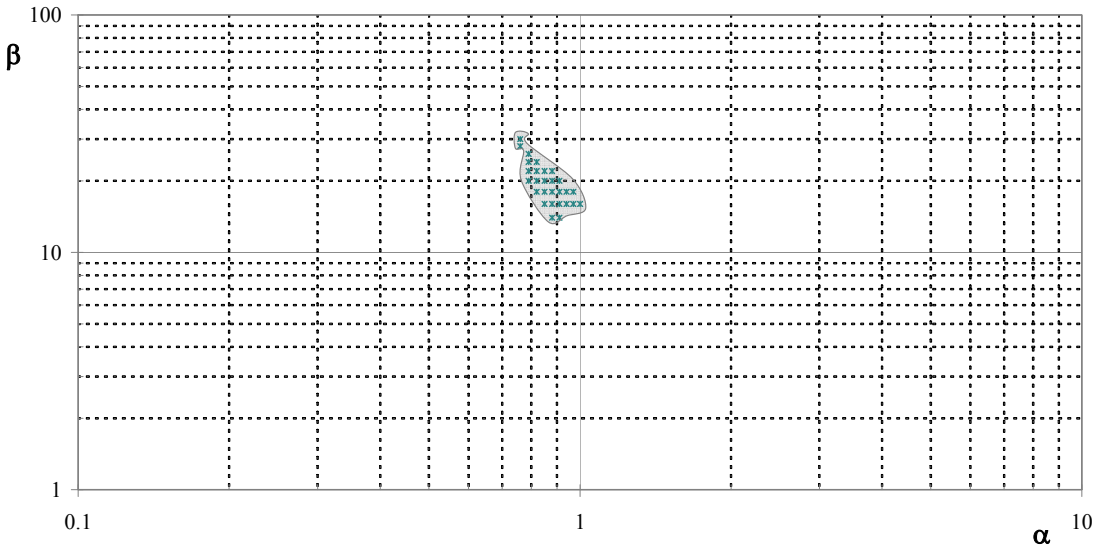


Figure 62. Satriano: feasible region of the FlaIR model parameters (AUC = 1).

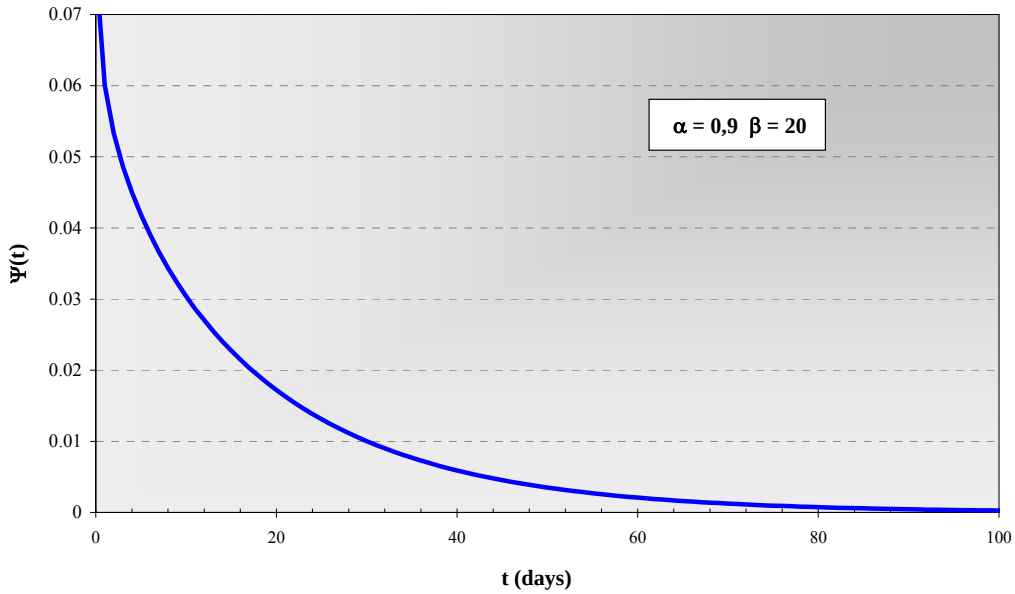


Figure 63. Satriano: transfer function of the FlaIR model.

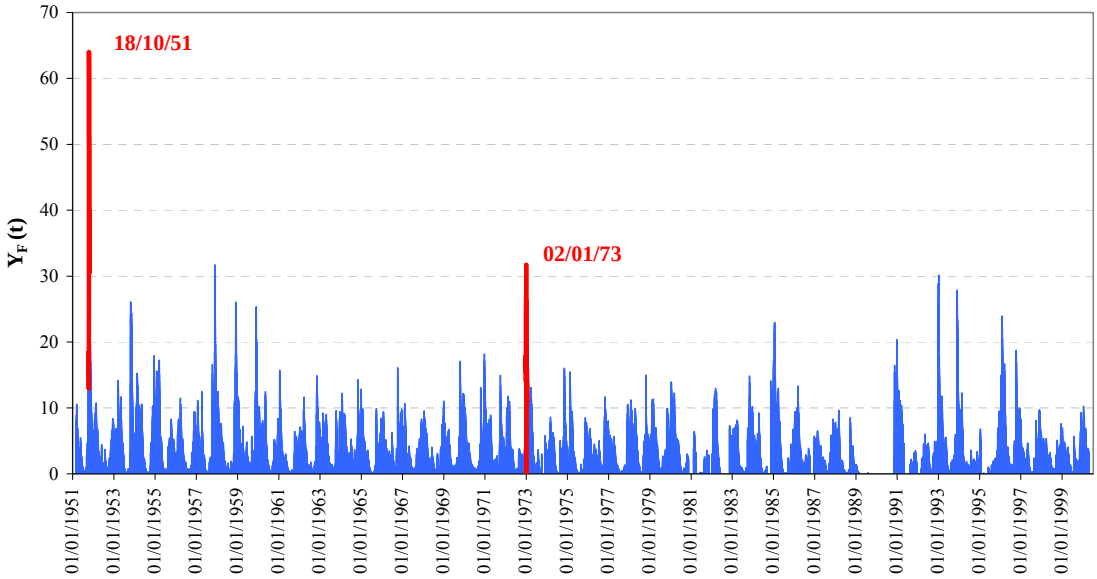


Figure 64. Satriano: mobility function.

5.4.2 Verzino

As reported in section 4.4, daily rainfall data from the Verzino rain gauge was used. The data ranges from 1 September 1921 to 26 March 2000.

The historical movement date is 8 January 1991.

Figure 65 shows the daily rainfall histogram and highlights the precipitation on the day of the landslide. The maximum rainfall heights, accumulated over several antecedent days (from 1 to 30) are presented in Table 14.

Table 14. Verzino: maximum rainfall heights for the landslide events.

		Maximum rainfall heights (mm)				
Date	**Start of rainfall event**	**1 day**	**5 days**	**10 days**	**15 days**	**30 days**
8/1/1991	25/12/1990	184.1	403.2	407.2	432.5	602.1

The transfer function $\psi(.)$ is the gamma function, characterized by the feasible region of the parameters, for which AUC=1, as shown in Figure 66; for $\alpha = 0.9$ and $\beta = 6$ days, the difference $r = f_U - f_L$ assumes its maximum value, equal to 0.05, and the associated $\psi(.)$ is shown in Figure 67. Figure 68 shows the mobility function; $Y_F(.)$ reaches the value 49.79 on 8 January 1991, which is assumed as the threshold $Y_{F,cr}$.

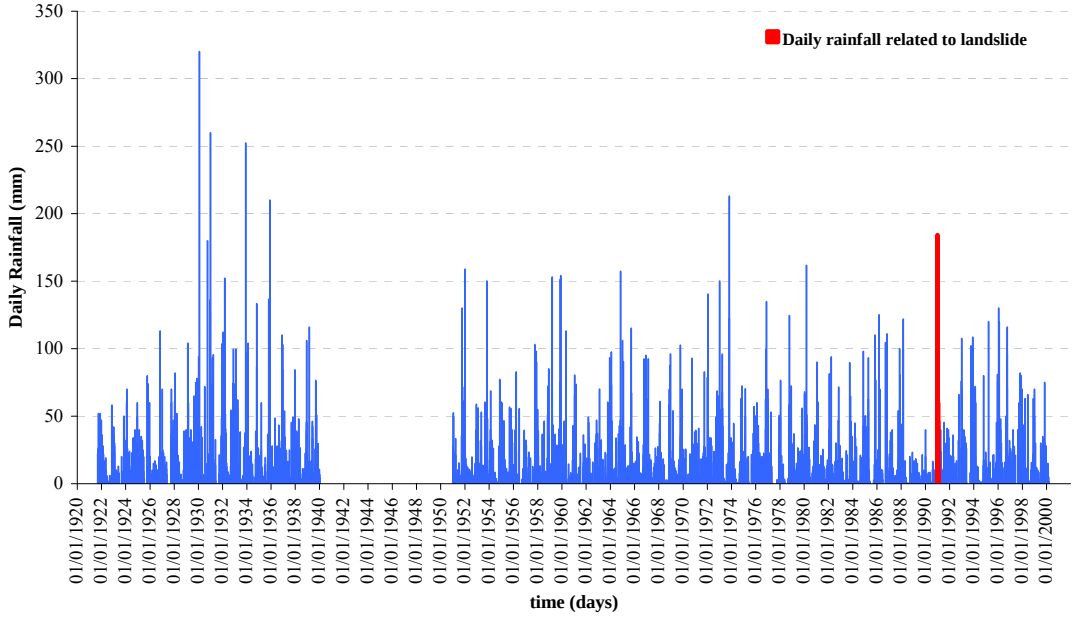


Figure 65. Verzino: daily rainfall histogram.

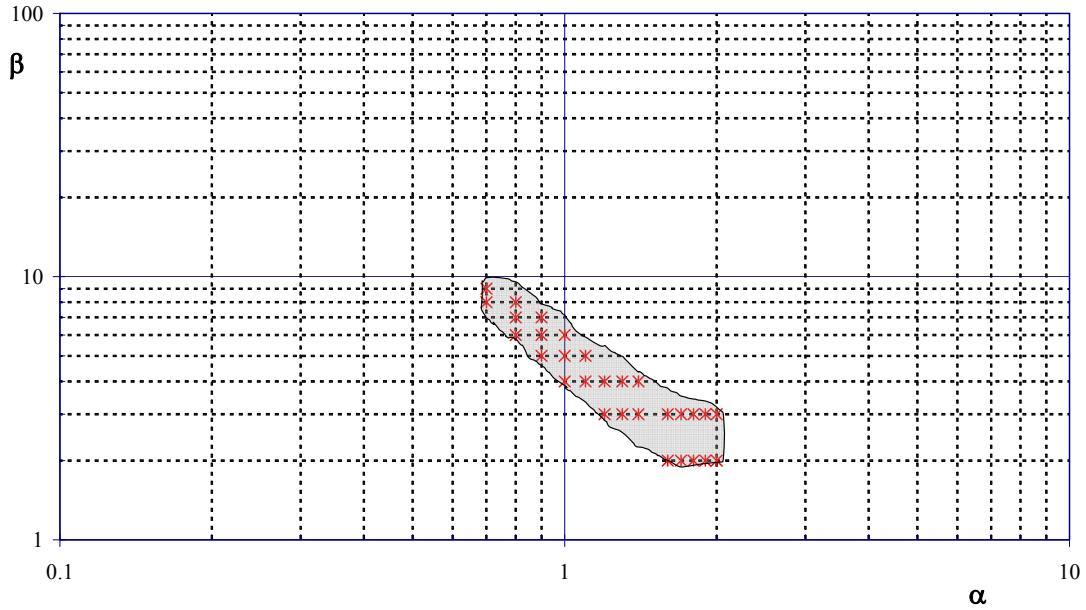


Figure 66. Verzino: feasible region of the FlaIR model parameters (*AUC* = 1).

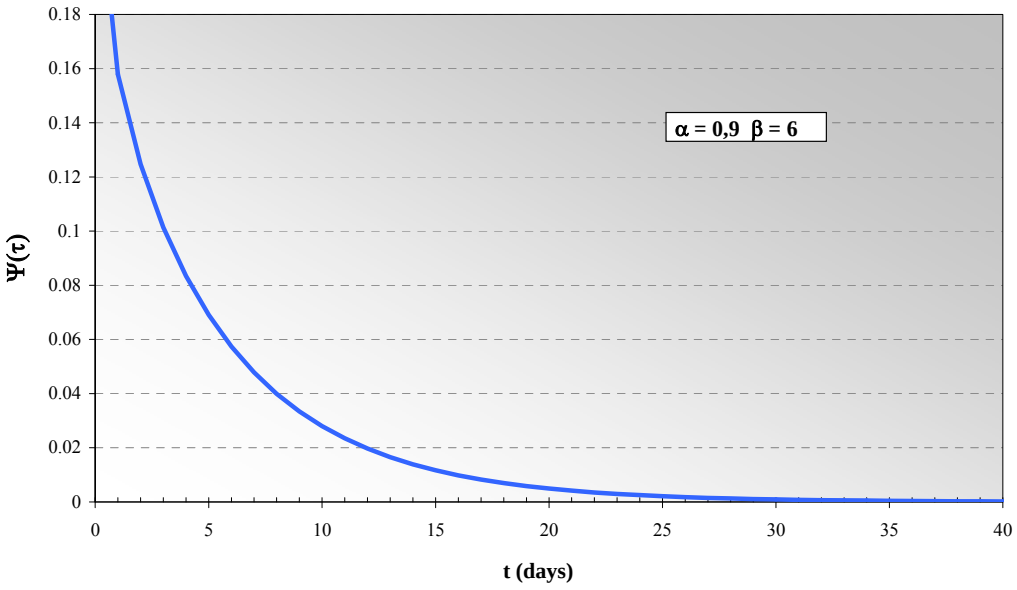


Figure 67. Verzino: transfer function of the FlaIR model.

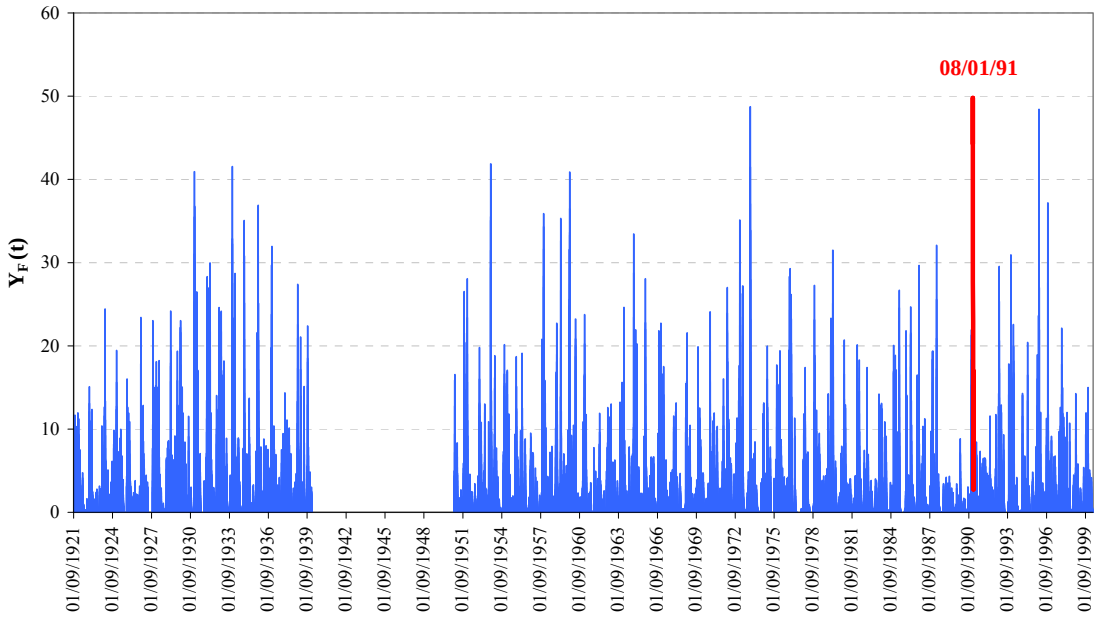


Figure 68. Verzino: mobility function.

5.4.3 Sarno

As reported in section 4.4, daily rainfall data from the Sarno - Santa Maria la Foce rain gauge was used, ranging from 1 January 1919 to 31 December 1998. The historical movement date is 5 May 1998.

In Figure 69, the daily rainfall histogram is shown, in which the precipitation related to landslide is indicated by the red bar. For this historical movement, the maximum rainfall heights, accumulated over several antecedent days (from 1 to 30) are reported in Table 15.

Table 15. Sarno: maximum rainfall heights for the landslide events.

		Maximum rainfall heights (mm)				
Date	**Start of rainfall event**	**1 day**	**5 days**	**10 days**	**15 days**	**30 days**
5/5/1998	4/5/1998	72.6	112.6	157.4	157.4	219.2

For this case study, the combination of two exponential functions was assumed as $\psi(.)$:

$$\psi(t) = \omega\beta_1 \exp(-\beta_1 t) + (1-\omega)\exp(-\beta_2 t) \quad \beta_1 \geq \beta_2 > 0,\ 0 \leq \omega \leq 1 \qquad (63)$$

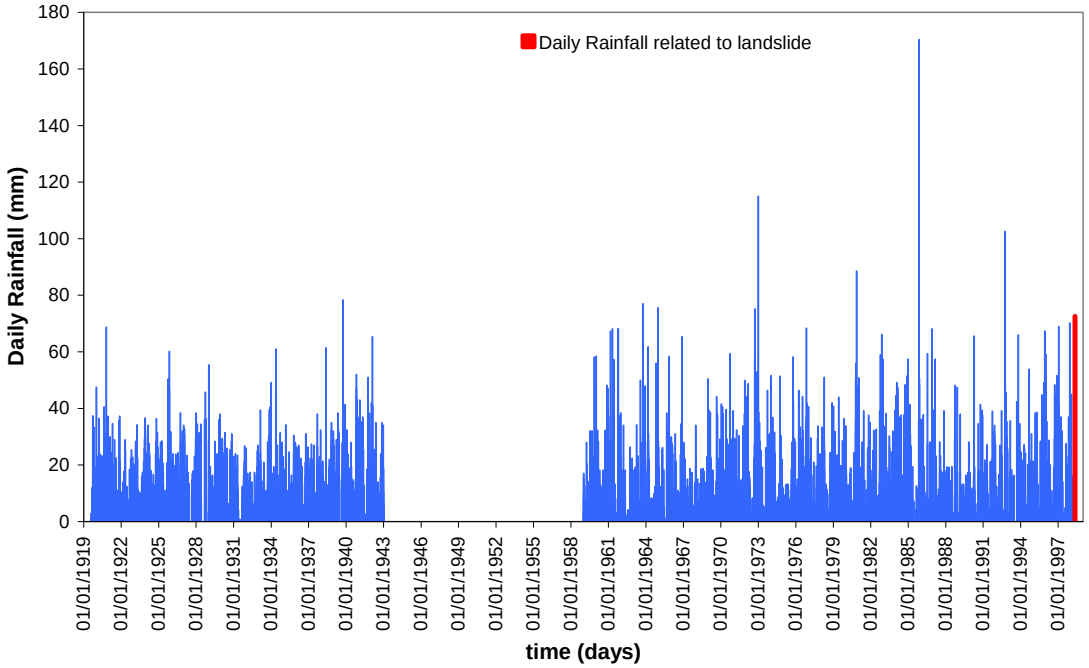


Figure 69. Sarno: daily rainfall histogram.

Moreover, due to the soil characteristics, evapotranspiration significantly reduces the water content during the summer season, and consequently $Y(.) = 0$ is assumed at the end of the dry season, namely on 1 October.

Furthermore, two values for threshold P_0 (see section 2.7) were considered: 75 mm and 90 mm. The parameters estimated by the ranking technique are shown in Table 16. In both cases, the maximum value of the mobility function is reached on 5 May 1998.

Table 16. Sarno: estimation of parameters using the ranking technique.

	Threshold P_0 (mm)	
	75	**90**
ω **(–)**	0.1	0.1
$1/\beta_1$ **(days)**	0.75	0.75
$1/\beta_2$ **(days)**	150	80
$Y_{F,cr}$	7.81	8.83

Figure 70 shows the associated transfer functions, while in Figure 71 and Figure 72 show the mobility functions $Y_F()$ for P_0 equal to 75 mm and 90 mm, respectively.

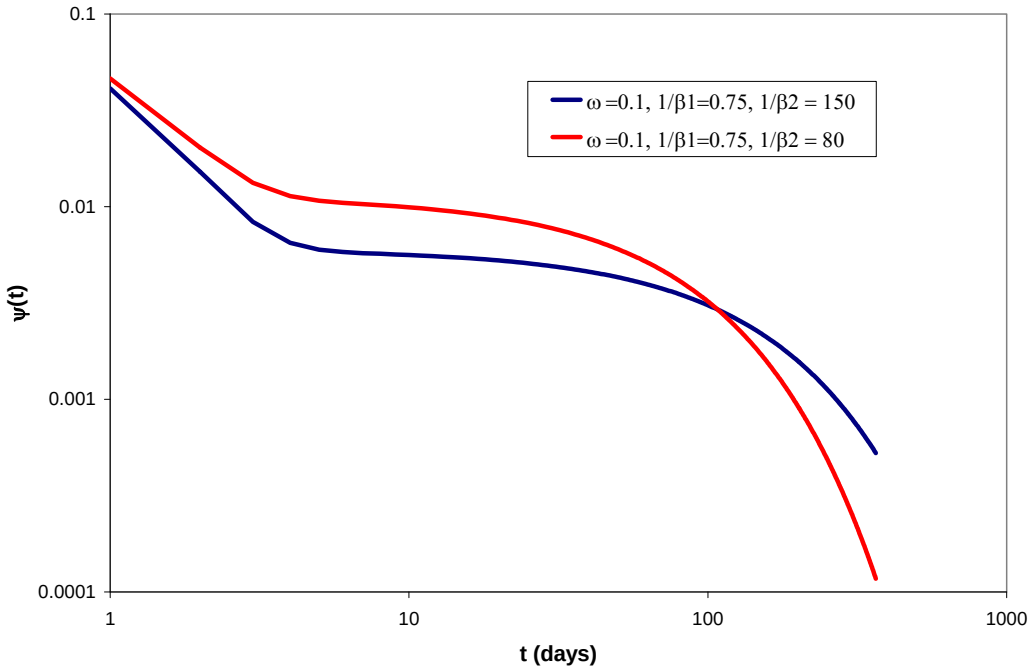


Figure 70. Sarno: transfer functions.

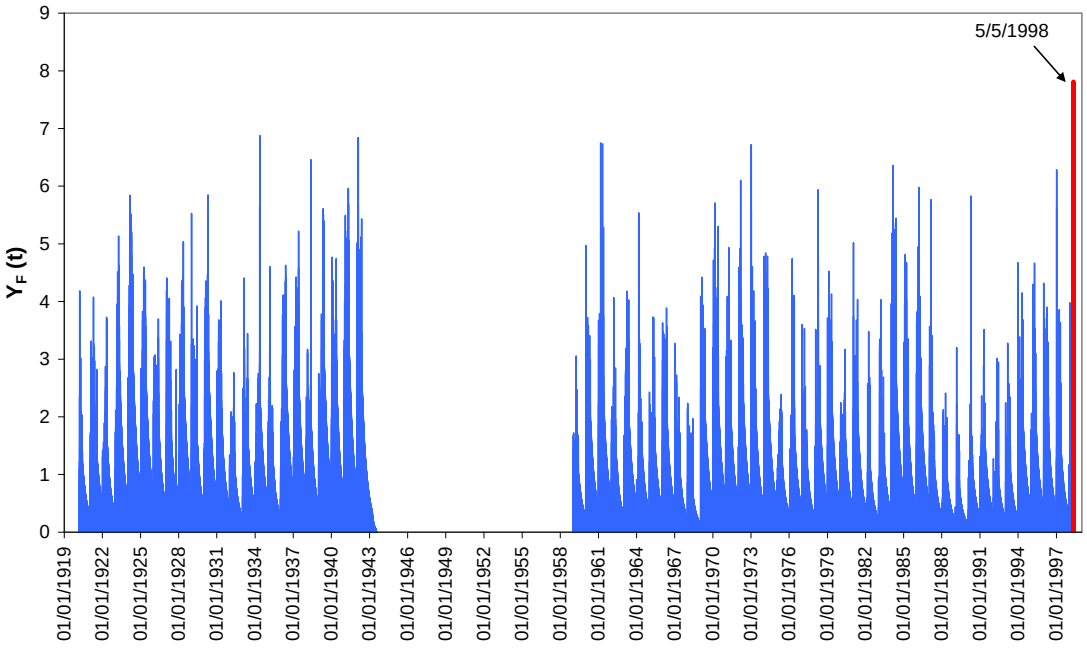


Figure 71. Sarno: mobility function for daily rainfall threshold equal to 75 mm.

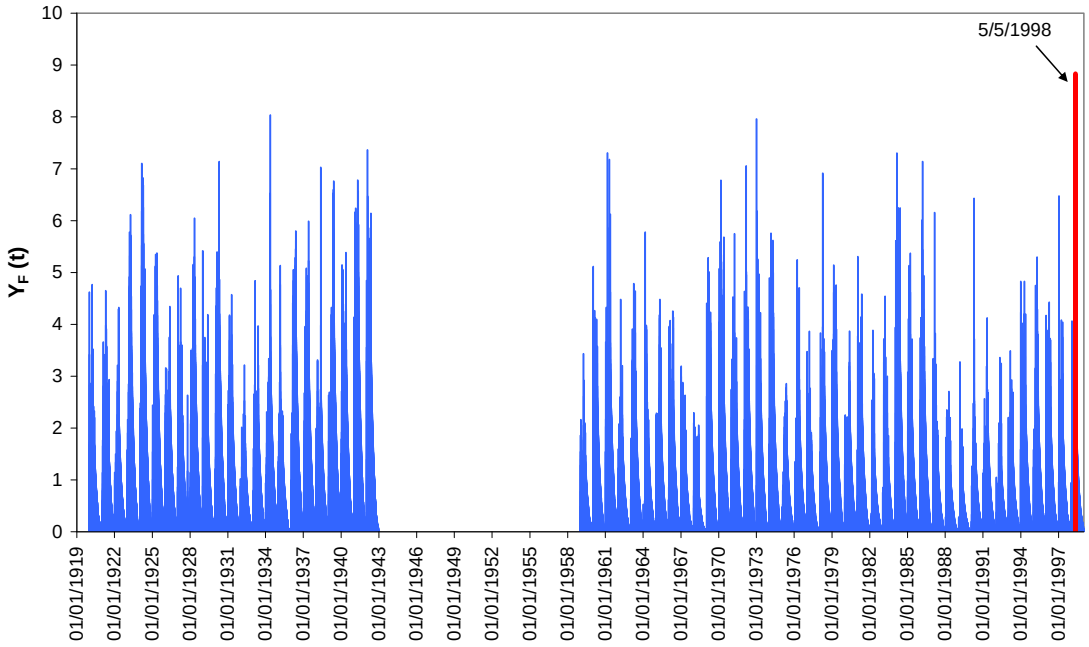


Figure 72. Sarno: mobility function for daily rainfall threshold equal to 90 mm.

In conclusion, for all the investigated case studies of the AMRA dataset, the FLaIR model can be considered as a Perfect Classifier, according to the ROC analysis. In fact, it is possible to determine transfer functions $\psi(.)$ and threshold values, which are able to produce forecasts without missed or false alarms, as long as back analyses of available data are carried out.

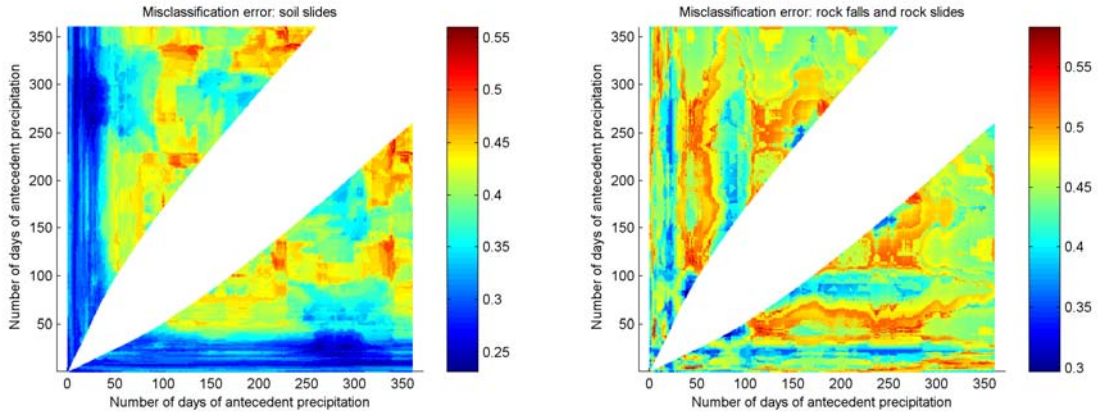
5.5 WESTERN NORWAY

Thresholds were estimated for debris flows and soil slides using a model of antecedent precipitation based on Equation 3. The predictors considered in the analyses were antecedent precipitation values ranging between 1 and 360 days. The critical number of days of antecedent precipitation for debris flows and soil slides were 1 and 7 days, respectively. The calibrated thresholds are shown in Table 17. Note that the calibration has considered not only landslide-triggering precipitation, but also non-triggering events (last two columns in Table 17). Correct and incorrect predictions are shown on the green and orange columns, respectively. It is a well known fact that hourly precipitation is required to adequately characterise triggering conditions for debris flows (e.g., Sandersen et al. 1996), so the threshold for debris flows given in Table 17 could be significantly improved if hourly data were available.

Table 17. Thresholds for debris flows and soil slides on the Norangselva catchment.

Type of landslide	Number of days of antecedent precipitation	Threshold (mm)	Days with landslides		Days without landslides	
			Predicted	Missed	Predicted	False alarms
Debris flows	1	17	3	0	4564	547
Soil slides	7	54	25	5	3681	1403

Discriminant analyses were also performed for this dataset by computing linear combinations of pairs of antecedent precipitation (e.g., 2-day with 30-day, 5-day with 45-day, etc.) from 1-day to 360-day antecedent precipitation, using the threshold model presented in Equation 5. From the results of the discriminant analysis at each pair, the misclassification error was computed (see section 3.2). The misclassification error is a measure of the number of missed events and false alarms. The results for soil slides are shown in Figure 73a. It can be seen that durations of less than about 40 days consistently produce the lowest misclassification errors.



(A) (B)

Figure 73. Misclassification errors for (A) soil slides and (B) rock falls and rock slides on the Norangselva catchment using the discriminant analysis method. The blank areas are combination pairs of antecedent precipitation that are strongly correlated (i.e., for the purposes of this paper, having an R-squared value larger than 0.8), and, consequently, were not analysed.

The lowest misclassification error in Figure 73a corresponds to n = 26 days and p = 266 days. The corresponding thresholds for this combination of days are shown in Figure 74. The two thresholds shown correspond to two different values of prior probabilities. It can be noted that the thresholds indicate that an increase in long-term antecedent precipitation produces an increase in required short-term antecedent precipitation for triggering, which is the opposite behaviour that should be expected (see remark at the end of section 2.5). Hence, it is concluded that the linear threshold model based on Equation 5 does not produce a consistent threshold and it is preferred to use the thresholds based on Equation 3 as presented in Table 17.

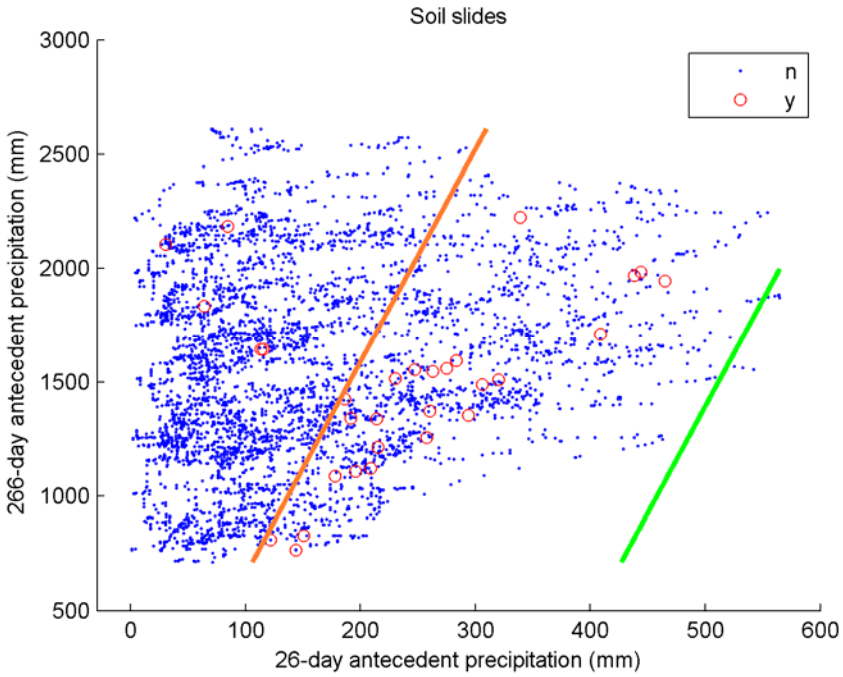


Figure 74. Linear thresholds based on Equation 5 for soil slides in the Norangselva catchment.

Figure 73b shows the result of the discriminant analyses for rock slides and rock falls. Compared to Figure 73a, a consistent pattern of low misclassification errors cannot be identified. Furthermore, the plot indicates that there is a predominance of higher misclassification errors. This low predictive capacity of precipitation for establishing thresholds in rock falls and avalanches confirms previous observations made by Sandersen et al. (1996), which were based on the analysis of 31 rock falls and rock slides from all Norway. Improvements in thresholds for this type of landslides can be foreseen if other factors influencing significant effects (e.g., freeze-thaw) are also included.

6 CONCLUSIONS

The following general conclusions can be drawn from the analysed case studies and experiences from the reported studies, including those summarized in the appendices:

- The analysis of hourly data, when available, indicates that the duration of triggering storms for initiating debris flows is less than 12 hours (e.g., see Barcelonnette case).
- In soil slides, the antecedent moisture conditions play an important role in the triggering conditions, and therefore should be accounted for in the threshold models. These antecedent conditions can be characterised by the soil moisture (see Appendix C) or other equivalent proxies, such as long-term precipitation and water supply (see Barcelonnette, South-Eastern Norway, and Appendices B and C).
- The contribution of snow melt to the of the water supply conditions on the slope should be evaluated, as it can be significant for characterising both antecedent and triggering conditions (see Appendix C).
- Freeze-thaw effects should be incorporated in the assessment of thresholds for rock falls and rock slides in order to reduce the misclassification errors for this type of landslides (e.g., see Western Norway).
- The variations of thresholds in different locations are mainly due to differences in lithological conditions, and in the case of soils the variations in permeability and depth of sediments. The soil moisture conditions (and the corresponding proxies in the threshold models) in areas with deep deposits of fine-grained soils (low permeability) are affected by long term precipitation or water supply (e.g., see South-Eastern Norway). In areas with shallower deposits of coarse-grained soils (high permeability), the response of the ground water conditions to precipitation and water supply is faster, and therefore the thresholds are controlled almost solely by short-term precipitation or water supply (e.g., see thresholds for debris flows in Barcelonnette and in Western Norway).

The following particular conclusions can be drawn for each individual case study:

- La Frasse landslide (Switzerland):
  - Empirical dynamic models can take as an input precipitation and observed displacement and be calibrated to produce forward predictions of landslide displacement. The models applied to La Frasse landslide have proved to be satisfactory.
  - The application of neural networks to dynamic empirical models seems to be promising, but the results applied to the La Frasse landslide seem to show some oscillations and increasing trends in the error of the model with time, which should be investigated further.
- Barcelonnette (France):
  - Debris flows are triggered by storms lasting between 1 and 9 hours, and are adequately predicted using an Intensity-Duration threshold.
  - Soil slides are triggered by storms with durations between 3 and 17 hours. The prediction of soil slides requires the use of a model that incorporates antecedent precipitation. An Intensity-Antecedent precipitation-Duration model was used, showing an improvement when compared to the Intensity-

Duration threshold. The critical antecedent precipitation was estimated equal to 23 days.
    - The FLaIR model achieved a very good performance corresponding to an area under the ROC curve (AUC) equal to 0.953.
- Western Norway:
    - Thresholds at the Norangselva catchment are defined by 1-day and 7-day antecedent precipitation for debris flows and soil slides, respectively.
    - Thresholds for rock falls and rock slides at the Norangselva in Western Norway are weakly predicted by the precipitation predictors. Other factors (e.g., freeze-thaw effects) need to be accounted for in order to improve the prediction performance of a threshold for this type of landslides.
    - Thresholds for debris flows in the dataset from the Norangselva catchment in Western Norway would be expected to improve with the use of hourly precipitation, which allows a better correlation with triggering of debris flows. However, representative observations are scarce due to the localised nature of short-duration episodes of rainfall, mainly associated with storms during the summer months.
- Satriano, Verzino and Sarno (Italy):
    - The FLaIR model was calibrated with sets of parameters that achieved perfect classification performance (AUC = 1) for these three cases.
- South-Eastern Norway:
    - A new method was proposed for incorporating uncertainty in time of occurrence in the estimation of thresholds.
    - The critical duration of antecedent precipitation for triggering of earth slides was estimated at 46 days.

7 REFERENCES

Ahmad, R. Developing early warning systems in Jamaica: rainfall thresholds for hydrogeological hazards. 2003.

Ref Type: Generic

Aleotti, P. (2004). "A warning system for rainfall-induced shallow failures." *Engineering Geology*, 73(3-4).

Aleotti, P., Baldelli, P., Bellardone, G., Quaranta, N., Tresso, F., Troisi, C., and Zani, A. (2002). "Soil slips triggered by October 13-16, 2000 flooding event in the Piedmont Region (Northwest Italy): critical analysis of rainfall data." *Geologia tecnica & ambientale*, X(1), 15-25.

Aleotti, P., and Chowdhury, R. (1999). "Landslide hazard assessment: summary review and new perspectives." *Bulletin of Engineering Geology and the Environment*,(58).

Arboleda, R. A., and Martínez, Ma. M. L. (1996). "1992 Lahars in the Pasig-Potrero River System." *Fire and mud: eruptions and lahars of Mount Pinatubo, Philippines*.

Bacchini, M., and Zannoni, A. (2003). "Relations between rainfall and triggering of debris-flow: case study of Cancia (Dolomites, Northeastern Italy)." *Nat. Hazards Earth Syst. Sci.*, 3(1/2), 71-79.

Barbero, S., Rabuffetti, D., and Zaccagnino, M. (2004). "Una metodologia per la definizione delle soglie pluviometriche a supporto dell'emissione dell'allertamento." 211-217.

Baum, R. L., Godt, J. W., Harp, E. L., McKenna, J. W., and McMullen, S. R. (2005). "Early warning of landslides for rail traffic between Seattle and Everett, Washington, USA." A.A. Balkema, New York, 731-740.

Bolley, S., and Oliaro, P. (1999). "Analisi dei debris flows in alcuni bacini campione dell'Alta Val Susa." *Geoingegneria Ambientale e Mineraria*,(Marzo), 69-74.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). "Classification and regression trees." Wadsworth International Group, Belmont, Calif.

Burges, C. J. C. (1998). "A tutorial on Support Vector Machines for pattern recognition." *Data Mining and Knowledge Discovery*, 2(2), 121-167.

Caine, N. (1980). "The rainfall intensity - duration control of shallow landslides and debris flows." *Geogr. Ann. Ser. A-Phys. Geogr.*, 62(1-2).

Calcaterra, D., Parise, M., Palma, B., and Pelella, L. (2000). "The influence of meteoric events in triggering shallow landslides in pyroclastic deposits of Campania, Italy." *Landslides in research, theory and practice*, 209-214.

Campbell, R. H. (1975). "Soil slips, debris flows, and rainstorms in the Santa Monica Mountains and vicinity, southern California." *Rep. No. 851*, United States Geological Survey.

Cancelli, A., and Nova, R. (1985). "Landslides in soil debris cover triggered by rainstorms in Valtellina (central Alps - Italy)." *Proceedings 4th International Conference and Field Workshop on Landslides*, 267-272.

Cannon, S. H. (1988). "Regional rainfall-threshold conditions for abundant debris-flow activity." *Landslides, Floods, and Marine Effects of the Storm of January 3-5, 1982, in the San Francisco Bay Region, California*, S. D. Ellen, and G. F. Wieczorek, eds., United States Government Printing Office, Washington, 35-42.

Cannon, S. H., and Ellen, S. D. (1985). "Rainfall conditions for abundant debris avalanches, San Francisco Bay region, California." *California Geology*, 38(12), 267-272.

Cannon, S. H., and Ellen, S. D. (1988). "Rainfall that resulted in abundant debris-flow activity during the storm of January 3-5, 1982, in the San Francisco Bay region, California." *Rep. No. 1434*, United States Geological Survey.

Cannon, S. H., and Gartner, J. E. (2005). "Wildfire-related debris flow from a hazards perspective." 363-385.

Cannon, S. H., Gartner, J. E., Wilson, R. C., Bowers, J. C., and Laber, J. L. (2008). "Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California." *Geomorphology*, 96(3-4).

Capparelli, G., Biondi, D., De Luca, D. L., and Versace, P. (2009). "Hydrological and complete models for forecasting landslides triggered by rainfalls." Naples, Italy, 162-173.

Capparelli, G., and Tiranti, D. (2010). "Application of the MoniFLaIR early warning system for rainfall-induced landslides in Piedmont region (Italy)." *Landslides*.

Capparelli, G., and Versace, P. (2010). "FLaIR and SUSHI: two mathematical models for early warning of landslides induced by rainfall." *Landslides*.

Cepeda, J., and Devoli, G. Rainfall thresholds for landslide triggering following volcanic ash eruptions and earthquakes. European Geosciences Union General Assembly 2008 . 2008. Vienna, Austria.

Ref Type: Generic

Cepeda, J., Hoeg, K., and Nadim, F. (2010). "Landslide-triggering rainfall thresholds: a conceptual framework." *Quarterly Journal of Engineering Geology and Hydrogeology*, 43(1), 69-84.

Cepeda, J., Díaz, M. R., Nadim, F., Høeg, K., and Elverhøi, A. An empirical threshold model for rainfall-induced landslides: application to the Metropolitan Area of San Salvador, El Salvador. 2009.

Ref Type: Generic

Ceriani, M., Lauzi, S., and Padovan, N. (1992). "Rainfall and landslides in the Alpine area of Lombardia Region, central Alps, Italy." *Interpraevent International Symposium*, 2, 9-20.

Chang, K. T., Chiang, S. H., and Hsu, M. L. (2007). "Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression." *Geomorphology*, 89(3-4).

Chen, M. S., Han, J. W., and Yu, P. S. (1996). "Data mining: An overview from a database perspective." *Ieee Transactions on Knowledge and Data Engineering*, 8(6), 866-883.

Chien-Yuan, C., Tien-Chien, C., Fan-Chieh, Y., Wen-Hui, Y., and Chun-Chieh, T. (2005). "Rainfall duration and debris-flow initiated studies for real-time monitoring." *Environmental Geology*, 47(5).

Clarizia, M., Gullà, G., and Sorbino, G. (1996). "Sui meccanismi di innesco dei soil slip." L'Artistica Savigliano pub, Alba, 585-597.

Congdon, P. (2006). "Bayesian statistical modelling." John Wiley & Sons, Chichester, England ; Hoboken, NJ.

Corominas, J. (2000). "Landslides and climate." Balkema, 1-33.

Corominas, J., Ayala, F. J., Cendrero, A., Chacón, J., Díaz de Terán, J. R., González, J., Moya, J., and Vilaplana, J. M. (2005). "Impacts on natural hazards of climatic origin: slopes instability risk." *A Preliminary General Assessment of the Impacts in Spain Due to the Effects of Climate Change*, J. M. Moreno Rodríguez, ed., Ministerio de Medio Ambiente, Ciudad Real, Castilla-La Mancha, Spain, 529-558.

Cortes, C., and Vapnik, V. (1995). "Support-Vector Networks." *Machine Learning*, 20(3), 273-297.

Crosta, G. (1998). "Regionalization of rainfall thresholds: an aid to landslide hazard evaluation." *Environmental Geology*, 35(2-3).

Crosta, G. B., and Frattini, P. (2001). "Rainfall thresholds for triggering soil slips and debris flow." *Proceedings 2nd EGS Plinius Conference on Mediterranean Storms*, 463-487.

Crozier, M. J. (1986). "Landslides causes, consequences and environment." Croom Helm, London.

Dhanya, C. T., and Kumar, D. N. (2009). "Data mining for evolution of association rules for droughts and floods in India using climate inputs." *J. Geophys. Res. -Atmos.*, 114.

Dillon, W. R., and Goldstein, M. (1984). "Multivariate analysis : methods and applications." Wiley, New York.

Dirksen, C., Kool, J. B., Koorevaar, P., and Van Genuchten, M. T. (1992). "HYSWASOR – simulation model of hysteric water and solute transport in the root zone, description of model." Wageningen Agricultural University, Wageningen, The Netherlands.

DUTI (Détection and Utilisation des Terrains Instables) (1986). "Le glissement de Cergnat - La Frasse, Projet EPFL.".

Edgers, L., and Nadim, F. Rainfall Induced Slides of Unsaturated Slopes. W.Lacerda, M.Ehrlich, S.A.B.Fontoura, and A.S.F.Sayao. Landslides:Evaluation and Stabilization.Ninth International Symposium on Landslides 2. 2004. Rio de Janeiro, Taylor & Francis.

Ref Type: Generic

EPFL-NCG, and Association technique Norbert De Cérenville Géotechnique + EPFL pour l'étude du glissement de La Frasse (2004). "Glissement de La Frasse, modélisation et étude de faisabilité.".

Fawcett, T. (2006). "An introduction to ROC analysis." *Pattern Recogn Lett*, 27(8).

Fayyad, U., PiatetskyShapiro, G., and Smyth, P. (1996). "From data mining to knowledge discovery in databases." *Ai Magazine*, 17(3), 37-54.

Flageollet, J. C., Maquaire, O., Martin, B., and Weber, D. (1999). "Landslides and climatic conditions in the Bracelonnette and Vars basins (Southern French Alps, France)." *Geomorphology*, 30(1-2), 65-78.

Floris, M., Mari, M., Romeo, R. W., and Gori, U. (2004). "Modelling of Landslide-Triggering Factors - A Case Study in the Northern Apennines, Italy." 745-753.

Frattini, P., Crosta, G., Carrara, A., and Agliardi, F. (2004). "Assessment of rockfall susceptibility by integrating statistical and physically-based approaches." Elsevier Science Bv, Florence, ITALY.

Frattini, P., Crosta, G., and Sosio, R. (2009). "Approaches for defining thresholds and return periods for rainfall-triggered shallow landslides." *Hydrol. Process.*, 23(10).

Frattini, P., Crosta, G. B., and Carrara, A. (2008). "Techniques for the evaluation and comparison landslide susceptibility models." Vienna.

Furseth, A. (2006). "Slide Accidents in Norway (in Norwegian)." Tun Forlag, Oslo.

Giannecchini, R. (2005). "Rainfall triggering soil slips in the southern Apuan Alps (Tuscany, Italy)." *Adv. Geosci.*, 2, 21-24.

Godt, J. W., Baum, R. L., and Chleborad, A. F. (2006). "Rainfall characteristics for shallow landsliding in Seattle, Washington, USA." *Earth Surface Processes and Landforms*, 31(1).

Goh, A. T. C. (1994). "Seismic Liquefaction Potential Assessed by Neural Networks." *J. Geotech. Eng. -ASCE*, 120(9), 1467-1480.

Guadagno, F. M. (1991). "Debris flows in the Campanian volcaniclastic soils." Thomas Telford Services Ltd, London, 125-130.

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P. (2007). "Rainfall thresholds for the initiation of landslides in central and southern Europe." *Meteorology and Atmospheric Physics*, 98(3-4).

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P. (2008). "The rainfall intensity-duration control of shallow landslides and debris flows: an update." *Landslides*, 5(1).

Heneker, T. M., Lambert, M. F., and Kuczera, G. (2001). "A point rainfall model for risk-based design." *J. Hydrol.*, 247(1-2), 54-71.

Heyerdahl, H., Harbitz, C. B., Domaas, U., Sandersen, F., Tronstad, K., Nowacki, F., Engen, A., Kjekstad, O., Devoli, G., Buezo, S. G., Diaz, M. R., and Hernandez, W. Rainfall Induced Lahars in Volcanic Debris in Nicaragua and El Salvador: Practical Mitigation. International Conference on Fast Slope Movements – Prediction and Prevention for risk Mitigation, IC-FSM2003 . 2003. Naples, Patron Pub.

Ref Type: Generic

Hong, Y., Hiura, H., Shino, K., Sassa, K., Suemine, A., Fukuoka, H., and Wang, G. (2005). "The influence of intense rainfall on the activity of large-scale crystalline schist landslides in Shikoku Island, Japan." *Landslides*, 2(2), 97-105.

Iiritano, G., Versace, P., and Sirangelo, B. (1998). "Real-time estimation of hazard for landslides triggered by rainfall." *Environmental Geology*, 35(2-3).

Innes, J. L. (1983). "Debris Flows." *Prog. Phys. Geogr.*, 7(4).

Jaedicke, C., and Kleven, A. (2008). "Long-term precipitation and slide activity in south-eastern Norway, autumn 2000." *Hydrol. Process.*, 22(4).

Jain, A. K., Duin, R. P. W., and Mao, J. C. (2000). "Statistical pattern recognition: A review." *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4-37.

Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). "Data clustering: A review." *Acm Computing Surveys*, 31(3), 264-323.

Jakob, M., and Weatherly, H. (2003). "A hydroclimatic threshold for landslide initiation on the North Shore Mountains of Vancouver, British Columbia." *Geomorphology*, 54(3-4).

Jakob, M. (2009). "A real time debris flow warning system for the North shore mountains of Vancouver." K. Sassa, and P. Canuti, eds., Springer-Verlag, Berlin.

Jan, C. D., and Chen, C. L. (2005). "Debris flows caused by Typhoon Herb in Taiwan." 539-563.

Jansson, P. E., and Karlberg, L. (2001). "Coupled heat and mass transfer model for soil-plant atmosphere systems.", Department of Civil and Environmental Engineering, Royal Institute of Technology.

Jibson, R. W. (1989). "Debris flows in southern Puerto Rico." A. P. Schultz, and R. W. Jibson, eds., Geological Society of America.

Jørstad, F. A. (1968). "Clay slides in Norway (in Norwegian)." *NorskGeografisk Tidsskrift*, 22, 214-219.

Karlsrud, K., Aas, G., and Gregersen, O. (1985). "Can we predict landslide hazards in soft sensitive clays? Summary of Norwegian practice and experiences." Toronto, 107-130.

Karlsrud, K., and Lillevik, S. (1984). "Inventory of clay slides evaluated by NGI in the period 1969–1984." *Rep. No. NGI-report 50904-1*, Norwgian Geotechnical Institute, Oslo, Norway.

Keefer, D. K., Wilson, R. C., Mark, R. K., Brabb, E. E., Brown, W. M., Ellen, S. D., Harp, E. L., Wieczorek, G. F., Alger, C. S., and Zatkin, R. S. (1987). "Real-time landslide warning during heavy rainfall." *Science*, 238(4829).

Laloui, L., Tacher, L., Moreni, M., and Bonnard, C. (2004). "Hydro-mechanical modeling of crises of large landslides : application to the La Frasse Landslide." Balkema, 1103-1110.

Larsen, M. C., and Simon, A. (1993). "A rainfall intensity - duration threshold for landslides in a humid-tropical environment, Puerto-Rico." *Geogr. Ann. Ser. A-Phys. Geogr.*, 75(1-2).

LeCun, Y. (1985). "Une procédure d'apprentissage pour réseau a seuil asymmetrique (a Learning Scheme for Asymmetric Threshold Networks)." 599-604.

Lee, I. M., and Lee, J. H. (1996). "Prediction of pile bearing capacity using artificial neural networks." *Comput. Geotech.*, 18(3), 189-200.

Lippmann, R. (1987). "An introduction to computing with neural nets." *ASSP Magazine, IEEE*, 4(2), 4-22.

Ljung, L. (2002). "System Identification Toolbox User's guide for use with Matlab." The MathWorks, Inc..

Ljung, L., and Söderström, T. (1983). "Theory and practice of recursive identification." MIT Press, Cambridge, Mass.

Lugeon, M., Patschoud, E., and Rothpletz, F. (1922). "Rapport d'expertise sur le glissement des Frasses." Etat de Vaud, Département des Travaux Publics, Service des Routes.

Lumb, P. (1975). "Slope failures in Hong Kong." *Quarterly Journal of Engineering Geology and Hydrogeology*, 8(1).

Maquaire, O., Malet, J.-P., Remaƈtre, A., Locat, J., Klotz, S., and Guillon, J. (2003). "Instability conditions of marly hillslopes: towards landsliding or gullying? The case of the Barcelonnette Basin, South East France." *Engineering Geology*, 70(1-2), 109-130.

Marchi, L., Arattano, M., and Deganutti, A. M. (2002). "Ten years of debris-flow monitoring in the Moscardo Torrent (Italian Alps)." *Geomorphology*, 46(1-2), 1-17.

Mayoraz, F., and Vulliet, L. (2002). "Neural Networks for Slope Movement Prediction." *International Journal of Geomechanics*, 2(2), 153-173.

McCulloch, W., and Pitts, W. (1943). "A logical calculus of the ideas immanent in nervous activity." *Bulletin of Mathematical Biology*, 5(4), 115-133.

Montgomery, D. R., Schmidt, K. M., Greenberg, H. M., and Dietrich, W. E. (2000). "Forest clearing and regional landsliding." *Geology*, 28(4), 311-314.

Moody, J. A., and Martin, D. A. (2001). "Post-fire, rainfall intensity-peak discharge relations for three mountainous watersheds in the western USA." *Hydrol. Process.*, 15(15), 2981-2993.

Moser, M., and Hohensinn, F. (1983). "Geotechnical Aspects of Soil Slips in Alpine Regions." *Engineering Geology*, 19(3), 185-211.

Nadim, F., Cepeda, J., Sandersen, F., Jaedicke, C., and Heyerdahl, H. (2009). "Prediction of Rainfall-Induced Landslides through Empirical and Numerical Models." Naples, Italy.

Noverraz, F., and Bonnard, C. (1990). "Technical note on the visit of the La Frasse landslide." Balkema, 1549-1554.

Paronuzzi, P., Coccolo, A., and Garlatti, G. (1998). "Eventi meteorici critici e *debris flows* nei bacini montani del Friuli." *L'Acqua, Sezione I/Memorie*, 39-50.

Parriaux, A., Lutz, T., and Tissières, P. (1987). "Traage au gouffre du Chevrier (Pralpes vaudoises) et mthodes didentification de luranine faible concentration.".

Patel, N. R. Data Mining. 2003. 1-27-2010.

Ref Type: Online Source

Planchat, J., Bonnard, C., Peron, H., and Laloui, L. (2009). "After 10,000 years of movements, is the La Frasse landslide stabilized? numerical modeling of the crises of a large landslide and of its mitigation measures." CIMNE, Barcelona, 95-126.

Rahardjo, H., Li, X. W., Toll, D. G., and Leong, E. C. (2001). "The effect of antecedent rainfall on slope stability." *Geotechnical and Geological Engineering*, 19(3), 371-399.

Rahardjo, H., Ong, T. H., Rezaur, R. B., and Leong, E. C. (2007). "Factors controlling instability of homogeneous soil slopes under rainfall." *J. Geotech. Geoenviron. Eng.*, 133(12).

Remaître, A. (2006). "Morphologie et dynamique des laves torrentielles: Applications aux torrents des Terres Noires du bassin de Barcelonnette (Alpes du Sud).", University of Caen-Basse-Normandie.

Rességuier, S. (2006). "River bank stability in a changing climate." *Rep. No. 20051095-2*, Norwegian Geotechnical Institute.

Rodolfo, K. S., and Arguden, T. (1991). "Rain-lahar generation and sediment-delivery systems at Mayon volcano, Philippines." R. V. Fisher, and G. A. Smith, eds., SEPM (Society for Sedimentary Geology), Tulsa.

Sandersen, F., Bakkehøi, S., Hestnes, E., and Lied, K. (1996). "The influence of meteorological factors on the initiation of debris flows, rockfalls, rockslides and rockmass stability." Trondheim, Norway, 97-114.

Sirangelo, B., and Versace, P. (1992). "Modelli stocastici di precipitazione e soglie pluviometriche di innesco dei movimenti franosi." Florence, Italy, D361-D373.

Sirangelo, B., and Versace, P. (1996). "A real time forecasting model for landslides triggered by rainfall." *Meccanica*, 31(1), 73-85.

Sirangelo, B., Versace, P., and Capparelli, G. (2003). "Forewarning model for landslides triggered by rainfall based on the analysis of historical data file." 289-304.

Sirangelo, B., Versace, P., and De Luca, D. L. (2007). "Rainfall nowcasting by at site stochastic model PRAISE." *Hydrol. Earth Syst. Sci.*, 11(4), 1341-1351.

Starkel, L. (1979). "The role of extreme meteorological events in the shaping of mountain relief." *Geographia Polonica*, 41, 13-20.

Tacher, L., Bonnard, C., Laloui, L., and Parriaux, A. (2005). "Modelling the behaviour of a large landslide with respect to hydrogeological and geomechanical parameter heterogeneity." *Landslides*, 2(1), 3-14.

Tatizana, C., Ogura, A. T., Cerri, L. E. S., and Rocha, M. C. M. Análise de correlação entre chuvas e escorregamentos – Serra do Mar, Município de Cubatão. 5 Congresso Brasileiro de Geologia de Engenharia 2. 1987. São Paulo, Associação Brasileira de Geologia de Engenharia e Ambiental.

Ref Type: Generic

Terlien, M. T. J. (1998). "The determination of statistical and deterministic hydrological landslide-triggering thresholds." *Environmental Geology*, 35(2-3).

Tuñgol, N. M., and Regalado, Ma. T. M. (1996). "Rainfall, Acoustic Flow Monitor Records, and Observed Lahars of the Sacobia River in 1992." *Fire and mud: eruptions and lahars of Mount Pinatubo, Philippines*.

Van Westen, C. J., and Daag, A. S. (2005). "Analysing the relation between rainfall characteristics and lahar activity at Mount Pinatubo, Philippines." *Earth Surface Processes and Landforms*, 30(13).

Versace, P., and Capparelli, G. (2008). "Empirical Hydrological Models for Early Warning of Landslides Induced by Rainfall." 627-630.

Versace, P., Sirangelo, B., and Capparelli, G. (2003). "Forewarning model of landslides triggered by rainfall." 1233-1244.

Vulliet, L., and Hutter, K. (1988). "Continuum Model for Natural Slopes in Slow Movement." *Geotechnique*, 38(2), 199-217.

White, I. D., Mottershead, D. N., and Harrison, S. J. (1992). "Environmental systems: an introductory text." Chapman & Hall, London.

Wieczorek, G. F., Morgan, B. A., and Campbell, R. H. (2000). "Debris-flow hazards in the Blue Ridge of Central Virginia." *Environmental & Engineering Geoscience*, 6(1).

Wieczorek, G. F., and Sarmiento, J. (1988). "Rainfall, piezometric levels, and debris flows near La Honda, California in storms between 1975 and 1983." *Rep. No. 1434*, United States Geological Survey.

Wieczorek, G. F. (1987). "Effect of rainfall intensity and duration on debris flows in central Santa Cruz Mountains, California." *Reviews in Engineering Geology*, VII, 93-104.

Wieczorek, G. F. (1996). "Landslide triggering mechanisms." *Landslides: investigation and mitigation*, A. K. Turner, and R. L. Schuster, eds., National Academy Press, Washington D.C., 76-90.

Wieczorek, G. F., and Glade, T. (2005). "Climatic factors influencing occurrence of debris flows." 325-362.

Wilson, R. C., and Jayko, A. S. (1997). "Preliminary Maps Showing Rainfall Thresholds for Debris-Flow Activity, San Francisco Bay Region, California." United States Geological Survey, Menlo Park.

Woodworth, G. G. (2005). "Biostatistics : a Bayesian introduction." Wiley-Interscience, Hoboken, N.J.

Zêzere, J. L., Trigo, R. M., and Trigo, I. F. (2005). "Shallow and deep landslides induced by rainfall in the Lisbon region (Portugal): assessment of relationships with the North Atlantic Oscillation." *Nat. Hazards Earth Syst. Sci.*, 5(3).

Zimmermann, M., Mani, P., Gamma, P., Gsteiger, P., Heiniger, O., and Hunziker G. (1997). "Murganggefahr und Climaänderung - ein GIS-basierter Ansatz." Zürich.

8 APPENDIX A: BAYESIAN PROBABILISTIC NETWORKS (BPN) AND NATURAL HAZARDS

8.1 RISK MANAGEMENT OF LARGE SCALE NATURAL HAZARDS

Natural hazards comprise an important contribution to risk in most countries of the world; their relevant types and intensities as well as the associated risks depend strongly on the specific location of occurrence. In some parts of the world, natural hazards are associated with risks endangering the existence of the societies located there. It is hence a great challenge for the engineering profession to provide methods and tools enhancing decision making for the purpose of efficient management of natural hazards.

In principle, risk management may be seen relative to the time frame of occurrence of events of natural hazards – i.e. risk management in the situations before, during and after the event of a natural hazard. This is because the possible decision alternatives or boundary conditions for decision making change depending on the corresponding time frame. Before a hazard occurs the issue of concern is to optimize investments into safeguarding or preventive measures such as protecting societal assets, adequately designing and strengthening societal infrastructure as well as developing preparedness and emergency strategies. During the event of a natural hazard the issue is to limit damages by rescue, evacuation and aid actions. After an event the situation is to some degree comparable to the situation before the event; however, after the event resources might be very limited and the main concern might be to re-establish societal functionality as well as to safeguard in regard to the possible next event. In Figure 1, the different decision situations and the focus of risk management for natural hazards are illustrated for the case of management of earthquake risks in an urban area.

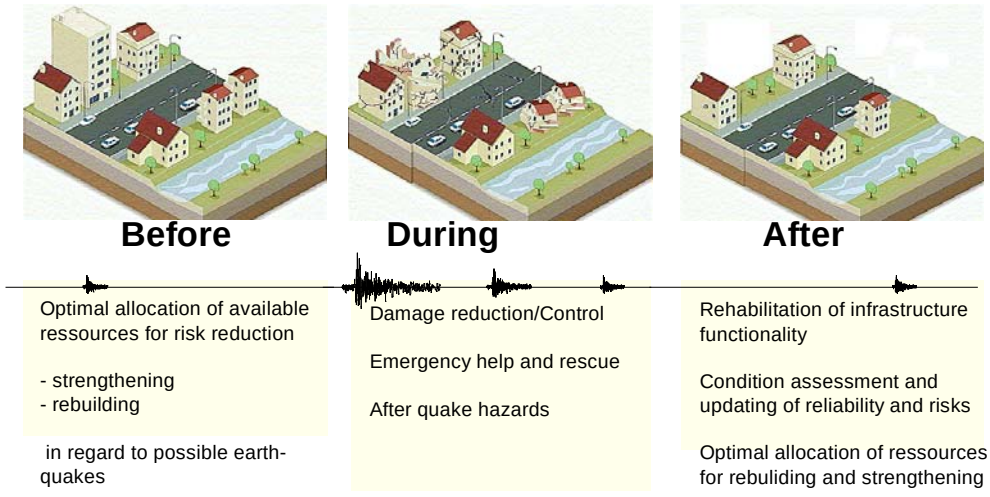


Figure 1: Decision situations for management of earthquake risks

The characteristics of natural hazards are very different depending on the individual exposure type. Gravitational hazards such as landslides, rockfall and avalanches are generally suddenly occurring events. The same applies for earthquakes, tsunamis and volcanic eruptions. On the other hand, floods, fire storms, climatic changes and droughts are generally more slowly evolving. In a risk management context, the probabilistic description required for the characterization of the natural hazards must take into account these differences in order to facilitate a realistic assessment of the possible consequences resulting from these events as well as to allow for the identification of possible relevant measures of risk reduction.

The consequences which potentially may be caused by such different exposures are manifold and generally depend strongly on the specific characteristics of the hazard as well as the location where it occurs. The immediate or direct consequences comprise loss of lives, damages to societal infrastructure and life lines as well as damages to the qualities of the environment. Follow-up or indirect consequences may include additional loss of lives caused by the outbreak of epidemics or hunger, loss of livelihoods, damages to the local and/or global economy as well as sociological and political effects.

8.2 THE JCSS FRAMEWORK FOR RISK ASSESSMENT

A guideline document (JCSS, 2008) describing the framework and principles for risk based engineering decision making has been developed by the Joint Committee on Structural Safety (JCSS). Further details can also be found in Faber et al. (2007a). The risk assessment for a given system is facilitated by considering a generic representation of the system as shown in Figure 2. Hazards are represented as exposures to the system and include any event having the potential to damage the constituents of the system. The constituents of the system can be considered as the first defence of the system against the exposures. The damages to the constituents are considered to be associated with direct consequences. Direct consequences may include monetary losses, loss of lives, damages to the qualities of the environment or changed characteristics of the constituents.

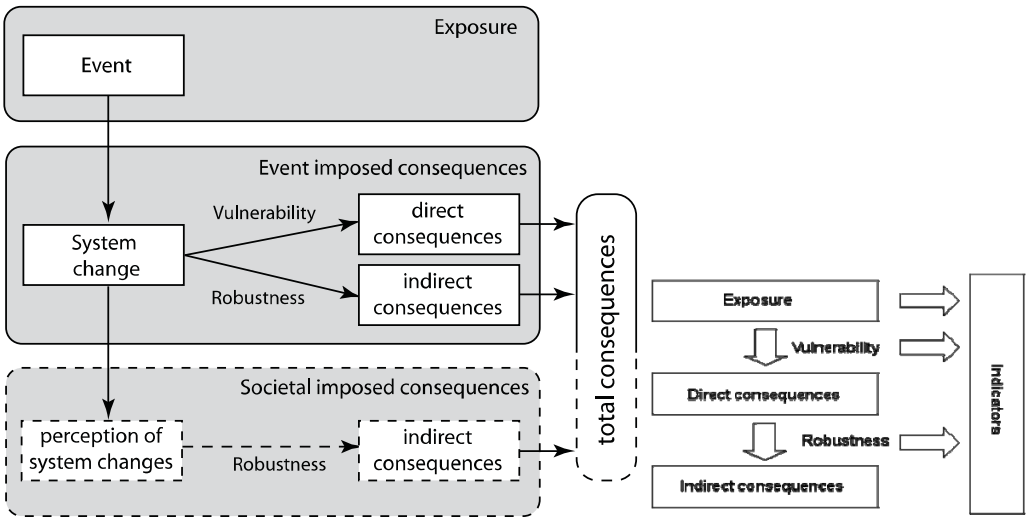


Figure 2: Generic representation used for the risk assessment of a system.

Based on the combination of events of constituent failures and the corresponding consequences, indirect consequences may occur. Indirect consequences may be caused by e.g. the sum of monetary losses associated with the constituent failures and the physical changes of the asset as a whole caused by the combined effect of constituent failures. They may occur due to physical system changes and/or due to the societal or public perception of such changes. The reason for this differentiation is to indicate how risk management may efficiently be supported by risk communication. The better and more targeted risk communication undertaken before, during and after events of natural hazards, the smaller the consequences caused by perception will be.

Based on this system representation, two system characteristics are defined – the vulnerability and the robustness of the system. The vulnerability of a given system characterizes the risk associated with the direct consequences whereas the robustness characterizes the degree to which the total risk is increased beyond the direct consequences. The description and quantification of the three characteristics – exposure, vulnerability and robustness is provided by the use of risk indicators. Risk indicators can be understood as any observable or measurable characteristic of the considered system containing information about its risk. In the context of landslide and rockfall hazards, the exposure can be related to the triggering factors for the landslides or rockfall, the vulnerability represents the physical process of the landslide/rockfall, damages to infrastructure and loss of lives and the robustness is associated with the follow-up consequences and the socio-economic or political impact.

The risk management of large scale natural hazards requires a systematic and consistent representation and management of information for a typically complex system with a large number of constituents or sub-systems. Such representation must enable a rational

treatment and quantification of the various uncertainties associated with the constituents as well as the system. The consistent handling of new knowledge about the system and its constituents as and when it becomes available and its use in the risk assessment and decision making process is also essential. Further, the numerous dependencies and linkages that exist between different constituents of the system need to be systematically considered. The above requirements and considerations necessitate the use of generic risk models for the assessment and management of risks due to natural hazards. The use of Bayesian Probabilistic Networks (BPNs) has proven to be efficient in such risk assessment applications (Graf et al., 2009; Faber et al., 2007b; Nishijima and Faber, 2007; Bayraktarli et al., 2006; Bayraktarli et al., 2005; Faber et al. 2005; Schubert et al, 2005 and Straub, 2005). The following section provides an overview of the principles and use of Bayesian Probabilistic Networks; details can be found in Jensen (2001).

8.3 BAYESIAN PROBABILISTIC NETWORKS – AN OVERVIEW

Bayesian probabilistic networks (BPNs) or Bayesian belief networks (BBNs) have been developed during the last two decades, as decision support tools originally targeted for purposes of artificial intelligence engineering. In contrast to rule based decision support systems, BPNs can be considered as normative expert systems meaning that:

- instead of modelling the expert they model the domain of uncertainty
- instead of using a non-coherent probability calculus tailored for rules, they are based on classical probability calculus and decision theory
- instead of replacing the expert they support her/him.

BPNs are constructed on the principles of causality and interrelationships between considered events and variables of interest. Graphical representations of such causally interrelated events are called causal networks. More formally, a causal network is a set of variables and a set of directed links or edges between the variables representing uncertain events and is also referred to as a directed graph; an illustration is shown in Figure 3. The relationship between different variables is expressed in terms of family relations; for example, when the link goes from a variable *A* to another variable *B*, the variable *A* is a parent of *B* and *B* is a child of *A*. The variables can, in principle, have any number of discrete states or a continuous sample space but can only attain one realisation at a time.

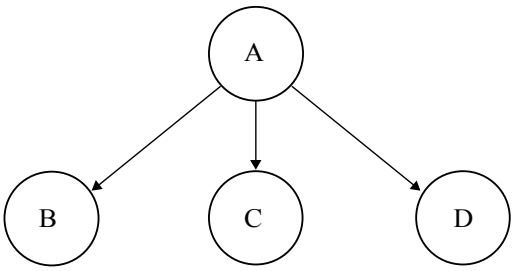


Figure 3: Illustration of a causal network.

A BPN is composed of a set of variables and a set of directed edges (or connections) between the variables. Each variable may have a countable or uncountable set of mutually exclusive states. The variables together with the directed edges form a directed acyclic graph. A conditional probability structure $P(A|B,C,D)$ is assigned to each variable *A* with parents *B*, *C* and *D*.

It is possible to work with BPNs containing variables with continuous states as well as discrete states. However, for mathematical reasons, it often becomes necessary to discretise continuous state variables into discrete state variables. Consider that n variables $A_1, A_2, ..A_n$ of a BPN are collected in the vector $\mathbf{U} = (A_1, A_2, ..A_n)^T$, also called the universe. In general it is of interest to be able to assess:

i) the joint probability distribution of the universe i.e. $P(\mathbf{U}) = P(A_1, A_2, ..A_n)$

ii) any marginalized set of the universe $P(A_i)$, and

   a. any conditional probability distributions subject to evidence e in regard to the states of individual variables, e.g. $P(A_i|e)$.

A BPN can be considered to be a special representation of such probability distributions. Using the chain rule of probability calculus, it is possible to write the probability distribution function $P(\mathbf{U})$ in the following form:

$$P(\mathbf{U}) = \prod_i P(A_i | pa(A_i)) \qquad (1)$$

where $pa(A_i)$ is the parent set of the variable A_i. The probability distribution function for elements of $\mathbf{U}$, e.g. for A_j can be achieved by marginalization i.e.

$$P(A_j) = \sum_{\mathbf{U}\setminus A_j} P(\mathbf{U}) = \sum_{\mathbf{U}\setminus A_j} \prod_i P(A_i | pa(A_i)) \qquad (2)$$

Evidence can now be introduced through the application of the chain rule. If it is assumed that evidence e is available in terms of statements such as: “the variable *A* with *n* possible states for some reason can only attain realizations in state *i* or *j* with probabilities a_i or a_j”,

then the joint probability distribution is given as $P(A,e)=(0,0,..,0,a_i,.0.,a_j,..0,0)$. It is seen that this probability distribution is achieved simply through the multiplication of $P(A)$ with the vector $(0,0,..,0,1,0,..,1,0,..,0,0)$. Such vectors (or tables) are also called as findings $\underline{e}$.
In general terms, this can be represented as:

$$P(\mathbf{U},e)=P(\mathbf{U})\underline{e} \tag{3}$$

Using Equation (1) and assuming that evidence represented in terms of *m* findings is available, Equation (3) can be written as:

$$P(\mathbf{U},e)=\prod_i P(A_i\,|\,pa(A_i))\prod_{j=1}^{m}\underline{e}_j \tag{4}$$

Finally the conditional probability distribution functions $P(A_j\,|e)$ can be derived through:

$$\begin{aligned} P(A_j\,|e) &= \frac{\sum_{\mathbf{U}\setminus A_j}\prod_i P(A_i\,|\,pa(A_i))\prod_{j=1}^{m}\underline{e}_j}{P(e)} \\ &= \frac{\sum_{\mathbf{U}\setminus A_j}\prod_i P(A_i\,|\,pa(A_i))\prod_{j=1}^{m}\underline{e}_j}{\sum_{\mathbf{U}} P(U,e)} \\ &= \frac{\sum_{\mathbf{U}\setminus A_j}\prod_i P(A_i\,|\,pa(A_i))\prod_{j=1}^{m}\underline{e}_j}{\sum_{\mathbf{U}} (P(U)\underline{e})} \end{aligned} \tag{5}$$

In summary, the steps involved in the construction of a BPN are:

- Formulation of causal interrelations of events leading to the events of interest (consequences), in terms of nodes (variables) connected by edges (arrows).
- Assigning to each variable a number of discrete mutually exclusive states.
- Assigning probability structures (tables) for the states of each of the variables (conditional probabilities in case that the variables are child nodes).
- Assigning consequences corresponding to the states represented by the network.

BPNs can be used at any stage of a risk analysis and assessment procedure and can readily substitute the traditionally used fault trees and event trees in logical tree analysis. Whereas common cause or dependency phenomenon poses significant complications in classical fault tree analysis, this is not the case with BPNs. These networks are basically designed to facilitate the modelling of such dependencies. Finally BPNs provide an enormously strong tool for decision analysis, including prior analysis, posterior analysis and pre-posterior analysis.

In the following sections, the application of the described risk assessment framework implemented with the use of BPNs is illustrated for different natural hazards.

8.4 EARTHQUAKE RISK MANAGEMENT USING BPNS

The management of earthquake risks for a large urban area is considered in this section. Details can be found in Faber et al. (2007b), Bayraktarli et al. (2006) and Bayraktarli et al. (2005).

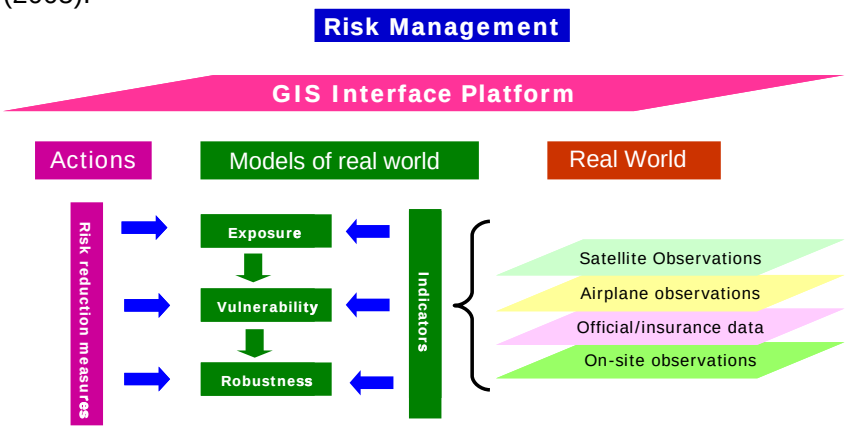


Figure 4: Components of a GIS based risk management system.

Generally, the exposures relating to natural hazards as well as the possible ensuing consequences can be considered to depend strongly on the specific geographical location of the occurrence of the event. For this reason, the use of Geographical Information Systems (GIS) needs to be considered in the context of natural hazards risk management. In Figure 4, the structure and components of a GIS based natural hazards risk management system are illustrated. The indicators of relevance for the characterization of exposures and consequences may be related to the models of the real world which form the basis for the risk assessment, considering the exposures, vulnerability and the robustness of the considered system. The risk assessed from these models and related to the real world through the indicators can then be managed by means of various risk reduction actions.

For earthquake exposure, the indicators of interest include the characteristics of soil, types of building and building materials, design codes applied for the design of buildings, occupancy of buildings, and the time of the earthquake as well as emergency preparedness. The GIS platform serves as a database for storing and managing the information required for the risk management process and strategy optimization. The data stored in the different layers of a GIS data base may directly be utilized in the modelling of the risks for the system.

The idea behind the application of generic BPN risk models is to identify categories of assets such as categories of buildings for which the risk assessment model has the principally same

structure. BPNs are then formulated for each category but with incorporation of the indicators characterizing exposures, vulnerability and robustness. In this way the individual generic risk models can be made specific for a given asset (e.g. building) by relating the risk model to the asset through the information of the indicators stored in the GIS data base. In Figure 5, an illustration is provided showing how a generic earthquake risk model in terms of a BPN has been formulated in terms of risk indicators. The BPN also includes decision nodes to facilitate optimal decision making in regard to possible strengthening of the existing structures.

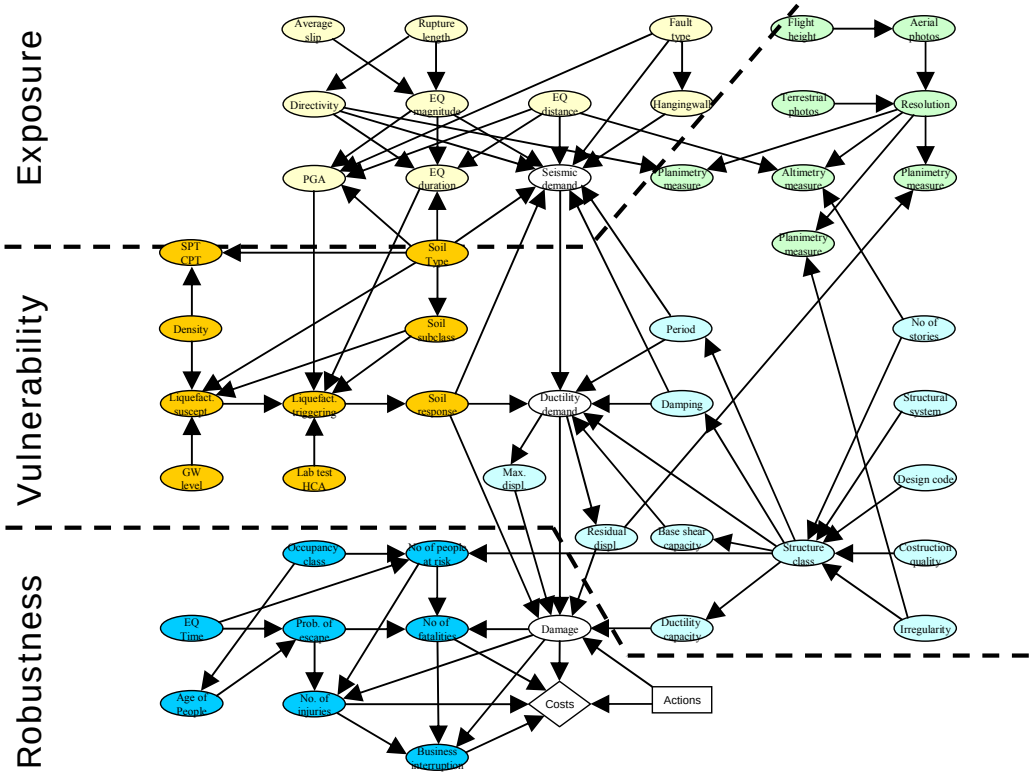


Figure 5: Generic indicator based BPN for the assessment of earthquake risks for one building class.

In Figure 6, the results of a generic risk assessment performed using BPNs integrated in a GIS database are shown. The risk assessment results illustrated in Figure 6 strongly depend on the use of generic BPN based risk models and the efficient management of relevant data in the GIS data base. The number of assessed building structures is in the order of thousands, making the development and analysis of risks models for each individual building practically not possible. Besides providing a very efficient means for risk assessment, the use of BPNs in large scale risk assessment also facilitates a consistent modelling of the relevant dependencies between the parameters which influence the risk. In the case of earthquake risk assessment, this concerns, in particular, the dependency in earthquake excitation of different buildings, the dependency between damages of different buildings, dependency between liquefaction failures at different locations due to dependencies in soil properties. It is very important to include such dependencies in the risk modelling process

owing to their possibly significant influence on the results of the risk assessment. The use of BPNs readily allows for the inclusion of dependencies between the risk associated with different assets at different locations by use of common variables or nodes linking the individual networks together.

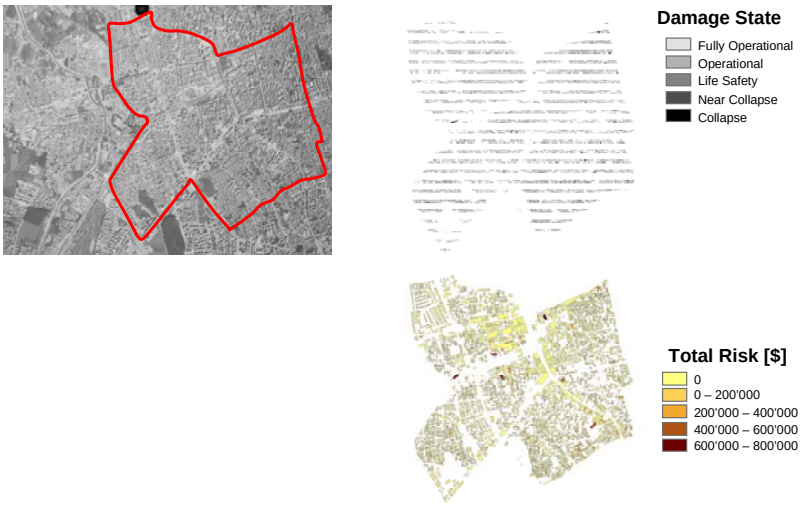


Figure 6: Results of a risk assessment utilizing generic BPN risk models and GIS data bases.

8.5 APPLICATION OF BPNS FOR TYPHOON RISK MANAGEMENT

The application of BPNs for typhoon risk management in the North West Pacific region is illustrated in this section. Further details can be found in Faber et al. (2007b) and Graf et al. (2009).

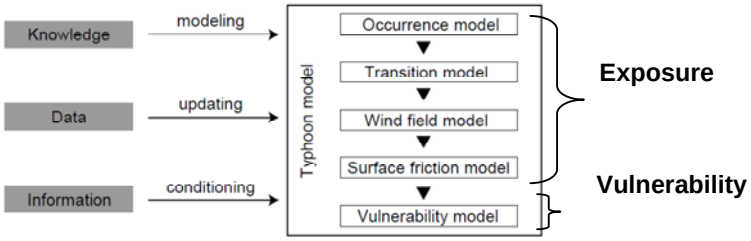


Figure 7: Bayesian modelling approach for typhoon risk management

A probabilistic typhoon model is first established; the model consists of two parts – an exposure model and a vulnerability model. The exposure model describes the probabilistic nature of the entire lifetime of typhoons and associated wind fields from their occurrence to dissipation. The exposure model is composed of four sub-models – the occurrence model, transition model, wind field model and surface friction model. The vulnerability model

represents the probability distribution of the loss of individual exposures as a function of the wind speed. The modelling structure is illustrated in Figure 7.

The occurrence model describes the probabilistic characteristics of the occurrence of typhoons as a function of location and season. The transition model describes the probabilistic characteristics of the movement of typhoons and the change of the intensity of typhoons. The wind field model describes the wind fields as a function of the relevant variables of the state of a typhoon. The surface friction model developed describes the relation between the wind speeds at gradient height and the wind speeds at nominal height on surface. The vulnerability model represents the probability distribution of the loss of individual exposures as a function of the wind speed.

Figure 7 also illustrates the proposed Bayesian perspective for the risk management of typhoons. The Bayesian framework provides a viable platform on which the available as well as incoming information can be utilized to condition and update the models in a consistent and systematic manner. The BPNs developed for some of the constituent models are shown and briefly discussed in the following.

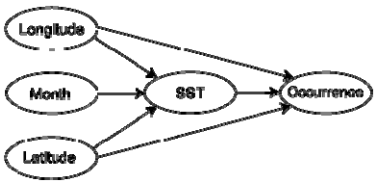


Figure 8: BPN for the occurrence of typhoons.

Figure 8 shows the BPN developed for the occurrence model. In the development of this model, the threshold for initiation of typhoon events is assumed to be defined by the moment at which the central pressure of each tropical storm becomes less than 1000 [hPa] for the first time in the lifetime of the tropical storm. Here, the probabilistic characteristics of the random variable ‘Occurrence’ are defined as a function of the states of the random variables ‘SST’, ‘Latitude’ and ‘Longitude’, the probabilistic characteristics of the random variable ‘SST’ is defined as a function of the states of the variables ‘Latitude’, ‘Longitude’ and ‘Month’. The quantitative probabilistic dependencies of all the variables are estimated using historical data.

The BPNs developed for the transition model and the wind field model are shown in Figure 9. The maximum wind speed during a typhoon event is taken as the indicator associated with the exposure. The transition model provides the conditional probabilities of the future states (6 hours later) of the central pressure and translation velocity, given the current states of the typhoon. The wind field model provides the probability of the wind speed at a target location using the states of the typhoons, which are obtained from the transition model.

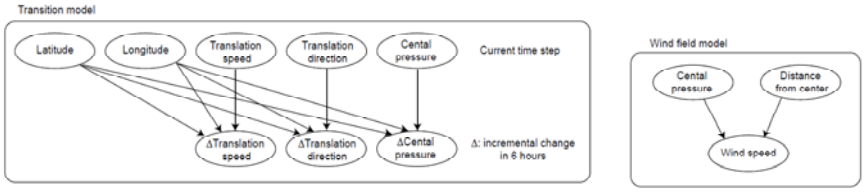


Figure 9: BPNs for transition model and wind field model.

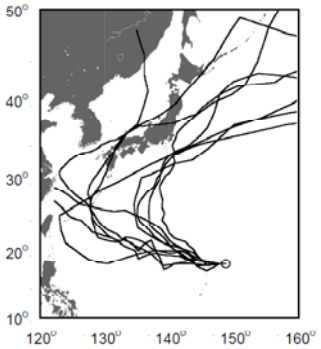


Figure 10: Illustration of simulated typhoon tracks.

Using these two models, the distribution of the maximum wind speed during one typhoon event is estimated. Some simulation results of the typhoon tracks are shown in Figure 10.

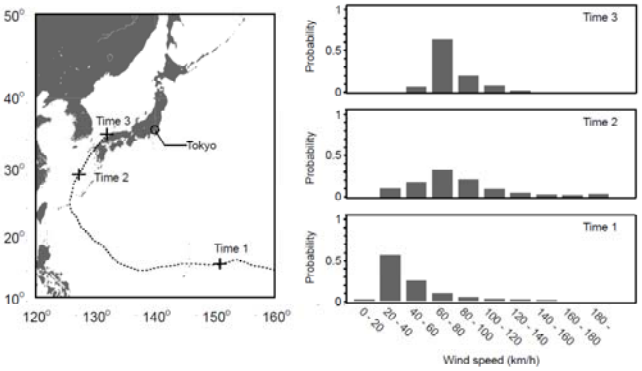


Figure 11: Conditional probabilities of maximum wind speed.

For further analysis, information about the central pressure, location and translation velocity of the typhoon is assumed to be available. Figure 11 (left) shows the time points at which this information is assumed to be obtained. The dotted line represents the realized typhoon track. Figure 11 (right) shows the calculated conditional probabilities of the maximum wind speed at Tokyo when the information becomes available at the three time points 1, 2 and 3 respectively.

The conditional probability of the maximum wind speed is seen to change according to the information available from time to time. As the typhoon approaches closer to the target

region, the precision of the indicator associated with exposure (represented by the maximum wind speed) becomes higher. In the context of risk management, the optimal decision time can then be obtained as a trade-off between this degree of precision and the time left for the mitigation actions.

8.6 ROCKFALL RISK ASSESSMENT USING BPNS

This section summarises the results from a study on rockfall risk assessment using BPNs. Further details can be obtained from Straub (2005) and Schubert et al. (2005).

Several rock-fall hazard classification systems are currently used for the assessment of risks from rockfall on roads sections. These procedures are generally based on a series of indicators determined for specific sections of the considered road. As a function of these indicators, a rating is obtained; this becomes the basis for decisions on further actions. These classification systems provide a quick overview of the risk from rockfall on an entire road network. However, the basic concepts underlying the application of such classification systems are generally inconsistent with risk analysis because of their simplified format. For this reason, the use of BPNs is illustrated in order to consistently represent the influence of the indicators on the evolving risk.

As an example, the classification system proposed in Budetta (2004), which is a modified version of the Rockfall Hazard Rating System (RHRS) developed at the Oregon State Highway Division, is considered. The following nine indicators or categories are applied in the classification process:

- Slope height
- Ditch effectiveness
- Average vehicle risk (the traffic volume)
- Decision sight distance
- Roadway width
- Slope mass ratio (A description of the geological character)
- Block size / volume of rockfall per event
- Annual rainfall and freezing periods
- Observed rockfall frequency

Each indicator is divided into four intervals and a score of 3, 9, 27 or 81 points is assigned to each of the indicators depending on its value. The total score, which represents the risk from a particular road section, is then obtained by adding the points for all the indicators.

The above listed indicators are now implemented in a BPN, which serves to reflect the various causal relations between the indicators and the considered processes. The

developed BPN is of principally illustrative character designed to demonstrate the capabilities of BPN approach. Hence the modelled relations between the various variables are not exhaustive and the resulting probabilistic model may not be completely realistic.

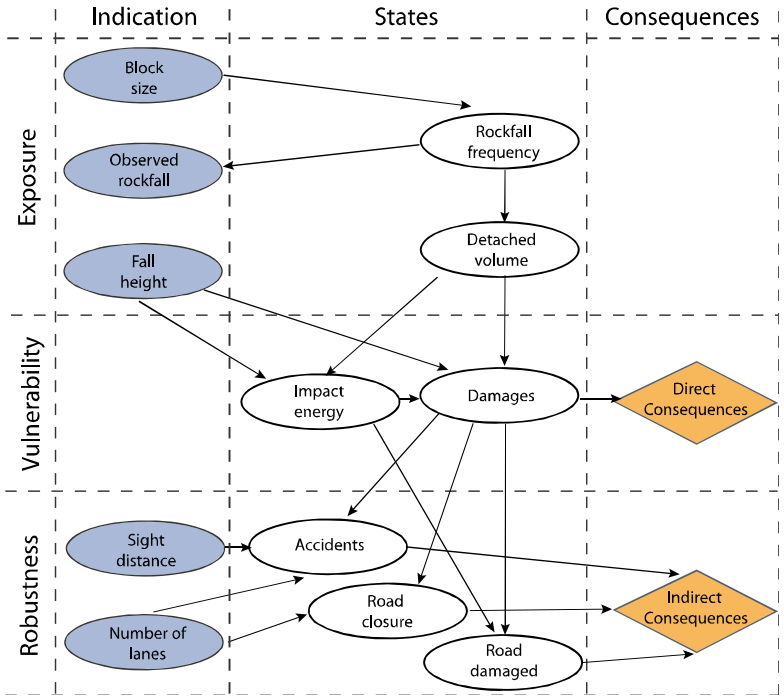


Figure 12: BPN for classification of rockfall hazard.

The developed BPN is shown in Figure 12. The BPN is arranged in accordance with the generic system representation characterised by exposure, resistance and robustness. The dark coloured oval nodes represent the indicator variables, the white oval nodes are variables introduced to represent the causal relations in the system and the diamond shaped nodes are the utility functions, characterizing the consequences associated with the system. The utility nodes define the expected value of the costs as a function of the number of fatalities and injuries, the physical damage on the road, the duration of road closure and the normal average traffic volume. The utilities are expressed in monetary terms. The BPN model is utilized by entering the state of the indicators as evidence in the respective nodes. The resulting score, which corresponds to the expected utility (and consequently the risk), is then obtained by evaluating the updated probabilities of all nodes in the network.

The developed BPN enables a detailed evaluation of the joint influence of the different indicators and the associated dependencies on the risk. It provides results which are consistent with the mathematical concept of risk and can hence be directly used for optimization purposes. The network can be further extended when additional investigations are performed and new information becomes available or when there is unavailability of information corresponding to certain indicator variables.

8.7 REFERENCES

Bayraktarli, Y.Y., Ulfkjaer, J.P., Yazgan, U. and Faber, M.H. (2005). On the Application of Bayesian Probabilistic Networks for Earthquake Risk Management. In: Augusti et al., eds. Safety and Reliability of Engineering Systems and Structures. Proceedings of the ICOSSAR 2005, Rome. Rotterdam: Millpress, 3505-3512.

Bayraktarli, Y.Y., Yazgan, U., Dazio, A. and Faber, M.H. (2006). Capabilities of the Bayesian probabilistic networks approach for earthquake risk management. In: Proceedings of ECEEs First European Conference on Earthquake Engineering and Seismology, September 3-8, Geneva, Switzerland.

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

Faber, M.H., Ranamukhaarachchi, S.L. and Rego, L. (2005). Sustainable Management of Natural Hazards in the Region of South-East Asia and South Asia. In: Proceedings of the International Symposium on Disaster Reduction on Coasts. November 14-16, 2005, Melbourne, Australia.

Faber, M.H., Maes, M., Baker, J.W., Vrouwenvelder T. and Takada, T. (2007a). Principles of risk assessment of engineering systems. In Kanda et al., eds. Applications of Statistics and Probability in Civil Engineering. Proceedings of ICASP10, Tokyo. The Netherlands: Taylor & Francis, 33-34.

Faber, M.H., Bayraktarli, Y.Y. and Nishijima, K. (2007b). Recent developments in the management of risks due to large scale natural hazards. In: Proceedings of SMIS XVI. Mexican National Conference on Earthquake Engineering. November 1-4, 2007, Ixtapa, Guerrero, Mexico.

Graf, M. et al., Nishijima, K. and Faber, M.H. (2009). A probabilistic typhoon model for the Northwest Pacific Region. In: Proceedings of the 7th Asia Pacific conference on Wind Engineering. November 8-12, Taiwan.

JCSS (2008). Risk Assessment in Engineering : Principles, System Representation & Risk Criteria. Internet publication : www.jcss.ethz.ch. ISBN 978-3-909386-78-9.

Jensen F.V. (2001). Bayesian Networks and Decision Graphs. Springer, New York.

Nishijima, K. and Faber, M.H. (2007). A Bayesian framework for typhoon risk management. In: Proceedings of the 12th International Conference on Wind Engineering. July 1-6 2007, Cairns, Australia. pp. 1959-1966.

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

Straub, D. (2005). Natural hazards risk assessment using Bayesian networks. In: Augusti et al., eds. Safety and Reliability of Engineering Systems and Structures. Proceedings of the ICOSSAR 2005, Rome. Rotterdam: Millpress, 2535-2542.

9 APPENDIX B: WATER SUPPLY THRESHOLDS CRITICAL FOR DEBRIS FLOW INITIATION IN NORWAY DERIVED IN THE CONTEXT OF THE INFRARISK PROJECT

9.1 SUMMARY

In Norway, debris flows cause considerable damage on the national transport infrastructure each year. The extremely rapid and water-saturated mass movements are frequently triggered by extreme precipitation and/or rapid snow melt – the latter as trigger mechanism often excluded in existing analyses. In this study, the data on the actual water supply, provided by the Norwegian Water and Energy Directorate (NVE) and the Norwegian Meteorological Institute (met.no), were used to assess the threshold values. Compared to rainfall data, these data define the hydro-meteorological threshold conditions more accurately throughout the year - i.e. the debris flow triggering conditions due to snow accumulation in autumn and winter and snow melt in spring and summer. Two sets of intensity-duration (ID) threshold curves were derived by analyzing the data on 502 past debris flows distributed over mainland Norway. Each set comprises minimum, medium and maximum ID curves. One set consists of absolute threshold values, which are uniform over the entire country; for the other data set normalization was accomplished by using the local precipitation day normal (PDN) to account for regional differences in climate. While the absolute threshold curve consists of fixed intensity-duration values which are uniform over the entire country, the normalized thresholds show a strong spatial variability. Absolute threshold values range between 15 and 107 mm/day, depending on duration and threshold level. In comparison, the normalized minimum threshold ranges between 6 and 63 mm/day, the medium threshold between 7 and 131 mm/day and the maximum threshold between 12 and 250 mm/day for different locations and durations. Statistical analyses reveal that the set of normalized threshold values shows a better performance than the absolute thresholds.

9.2 INTRODUCTION

In Norway, road and railway closures, due to natural hazards induced by extreme weather, cause high economical losses each year (e.g., Bråthen et al. 2008). Occupants of vehicles passing through the affected routes at that time may also be at risk. The Norwegian Public Roads Administration (*Statens Vegvesen)* and the Norwegian National Rail Administration (*Jernbaneverket)* are therefore highly interested in assessing the risk of natural hazards affecting their transportation networks.

The research project InfraRisk (http://www.ngi.no/en/prosjektnett/infrarisk) is concerned with this topic, focusing on the relationship between extreme weather events, natural hazards and transport infrastructure vulnerability in Norway. Rapid mass movements such as snow avalanches, rock falls and debris flows are often triggered by extreme weather events and cause considerable damage on the transport infrastructure. Thus, it is of central concern for the project to gather more knowledge about specific triggering conditions related to these events. Previous studies introduced thresholds that take practical advantage of the correlations between hazard initiation and meteorological triggering variable(s) for the purpose of early

warning. For snow avalanches Jaedicke and Bakkehøi (2007) introduced a framework which combines the amount of new snow with certain wind speeds over different durations. Furthermore, Sandersen et al. (1996) made a comparative study between debris flows and rock falls, and precipitation events related to their initiation. They concluded that debris flows show a closer relationship and a shorter response time regarding weather impacts than rock falls. However, the thresholds were merely based on the triggering variable precipitation and dismissed temperature as a factor thus excluding processes such as snow melt or freeze-thaw cycles. The measured amount of precipitation (accumulated as snow) can differ significantly from the actual amount of water that is released to the ground, and debris flows in spring often occur during periods of rapid snow melt.

This study is based on a large sample of debris flow events documented over the past 50 years. The goal is to define hydro-meteorological threshold conditions for debris flow initiation that include also the water supply from snow melt.

9.3 DATASET

All topographical analyses within this study are based on a Digital Elevation Model (DEM) with a 10 m x 10 m resolution, which is provided by the Norwegian Mapping Authority (*Statens Kartverk)*. The actual threshold definition uses a comprehensive database with over 600 debris flow events which recently became available as well as gridded data providing spatially continuous information on water supply for the entire country. Both data sets in combination give the unique opportunity to estimate critical initiation conditions.

9.3.1 Gridded hydro-meteorological data

The climate grids used in this study are provided by the Norwegian Meteorological Institute (met.no) and the Norwegian Water and Energy Directorate (NVE) and available on the internet portal www.SeNorge.com. They have a resolution of 1 km x 1 km and are produced on a daily basis for the time period since 1957 until today.

The hydro-meteorological information related to debris flow events is extracted from a water supply grid which yields information about the actual release of water to the ground (including storage effects of snow fall and snow melt). The grid is computed from a combination of temperature and precipitation data, both recorded by national meteorological stations and corrected for elevation. Water supply information is the amount of available water, aggregated on one day and specified in millimetres [mm].

In addition, the precipitation day normal (PDN) calculated for the time period 1961-1990 is used for normalization issues. The PDN is the mean annual precipitation (MAP) divided by the average number of precipitation days per year (APD). PDN indicates the intensity of water supply throughout a year.

9.3.2 Slide database

For Norway a slide database with almost 33 000 entries was compiled from slide records documented by the national road and railway authorities, as well as governmental and consulting organizations. The database is open to public and available on www.skrednett.no. Each slide event is stored as a point located at the area of deposition and contains information about the time of occurrence. Within the database approximately 700 events are specified as debris flows (*flomskred*) from which 648 events occurred in the time period 1976-2008 and have an event time accuracy of 30 minutes or less. However, 146 from the 648 recorded debris flows had to be discarded due to missing water supply information or low data quality. Hence, in total 502 debris flow events were finally available for further analyses.

Since we were interested in the hydro-meteorological conditions that led to the initiation of debris flows, the documented points had to be relocated from the area of deposition to a potential source area. This was done by a simple GIS procedure illustrated in Fig. 1 (for further information see Meyer et al. submitted). The relocated events were used to extract water supply information for the day of the occurrence of the event until four weeks in advance of the debris flow. These data served as input for the threshold model.

9.4 THRESHOLD MODEL

Threshold models based on the analysis of conditions that triggered debris flows in the past usually define critical intensity-duration (ID) relationships for debris flow initiation. Power laws are most common to describe critical ID relationships as threshold curves which usually have the following form:

(1) $I = a*D^b$

where I is the rain or water supply intensity, D is the duration, and a and b are parameters (Guzzetti et al. 2007).

To account for the spatial variability in local climate the intensity is often normalized by the mean annual precipitation (MAP) or the precipitation day normal (PDN) (Cannon and Ellen 1985, Wilson and Jayko 1997).

To define a valid intensity-duration ID threshold the duration of the rainfall and/or snow melt event that triggered the documented debris flow has to be defined. From the water supply data extracted for four weeks in advance of each debris flow event, the maximum cumulative average intensity [mm/day] was chosen to represent the beginning of the triggering event (Table 1). The ID relationships used for threshold definition can thus be defined as

(2) $I = max(WS_C/D) \sim D$

where WS_C is the cumulative water supply in advance of the day of occurrence and D is the corresponding duration in days.

Approximately 70% of the data consist of ID relationships with durations of one day and 90% feature water supply durations of less than seven days. In accordance to that, thresholds were established for water supply durations of one to seven days and should therefore also be applied within this time frame.

For this study three different threshold levels - minimum, medium and maximum – were defined in order to offer three levels of preparedness for implementation in an early warning system.

So-called minimum ID thresholds are based on the assumption that debris flows can occur if a specific minimum rainfall/ water supply intensity is exceeded over a given duration. It thus indicates the lowest threshold level below which a process is very unlikely to occur (Guzzetti et al. 2007, Baum and Godt 2010). In addition, medium and maximum thresholds can be often defined simultaneously. They indicate the levels above which debris flows become very likely or almost always occur, respectively.

9.5 CALIBRATION OF THRESHOLD

Since there were several sources of uncertainty related to the data (for more detailed information see Meyer et al. submitted) and no data was available on non-triggering events it was decided to aim for a minimum threshold that captures 90% of the obtained ID relationships. In accordance to that the medium threshold was designed to capture 50%, and the maximum threshold 10% of the data, respectively. Minimum, medium and maximum threshold curves were drawn by executing quantile regressions along the 0.1, 0.5 and 0.9 percentiles.

9.5.1 Absolute ID threshold

The obtained absolute ID thresholds can be described with the following equations:

(1) $I_{MIN} = 23.3 \times D_{day}^{-0.24}$
(2) $I_{MED} = 60.1 \times D_{day}^{-0.52}$
(3) $I_{MAX} = 107.2 \times D_{day}^{-0.57}$

where I_{MIN} is the minimum water supply intensity above which debris flow initiation becomes possible, I_{MED} is the medium threshold level above which debris flows become very likely, I_{MAX} is the maximum threshold level above which triggering of debris flows is seen as almost certain and D_{day} is the corresponding duration in days. All the intensities are expressed in mm/day.

Fig. 2 shows the three threshold levels graphically. The negative exponent in all functions indicates that with increasing duration of water supply, the intensity needed to trigger debris flows decreases.

9.5.2 Normalized ID threshold

The obtained normalized ID thresholds represented in Fig. 3 can be described with the following equations:

(1) $I_{MIN} = (2.18 \times D_{day}^{-0.24}) * PDN$
(2) $I_{MED} = (4.51 \times D_{day}^{-0.51}) * PDN$
(3) $I_{MAX} = (8.66 \times D_{day}^{-0.55}) * PDN$

where I_{MIN} is the lowest threshold level of water supply intensity, I_{MED} is the intermediate threshold level, I_{MAX} is the highest threshold level, PDN is the precipitation day normal and D_{day} is the corresponding duration in days. All the intensities are expressed in mm/day.
As the PDN changes locally also the threshold intensities show a spatial variation (Fig 4). Accordingly, intensities needed for debris flow initiation have to be considerably higher at the wet Norwegian west coast than those in the drier east of the country. Depending on location and considered duration, minimum threshold intensities range between 6 to 63 mm/day, medium threshold intensities between 7 to 131 mm/day and maximum threshold intensities between 12 to 250 mm/day. However, in areas with high PDN the variation of the thresholds is larger as in areas with a lower PDN. Thus, the local PDN has a stronger impact on threshold conditions in the former areas than in the latter. Compared to the absolute ID threshold introduced in this study the normalized threshold intensities are slightly lower in drier areas and remarkably higher in wet areas.

9.6 COMPARISON WITH OTHER THRESHOLDS

The minimum absolute threshold levels obtained in this study are in the range of other minimum ID thresholds found in literature (Fig. 5). Our absolute minimum threshold is lower than most of the other regional thresholds but compares well with thresholds obtained for temperate regions that are characterized by a high annual precipitation, e.g. western Canada and the north-western U.S. (Montgomery et al. 2000, Jakob and Weatherly 2003, Baum et al. 2005). The level of our absolute minimum threshold is further supported by a local threshold obtained by Cepeda (2010) for a small catchment at the Norwegian west coast where 17 mm/day is defined as the critical amount of rainfall (see Fig. 5a, triangle symbol). Norway shows similarities to the climate of the above regions; it indicates that the critical amount of supplied water needed for debris flow initiation is in fact dependent on the common regional climate. Procedures to normalize for the local climate are thus justified.
The minimum ID threshold normalized with the PDN is considerably lower than the two thresholds introduced by Guzzetti et al. (2007) for areas in central and southern Europe. This indicates that despite normalization significant differences remain between regions.
However, a statistical sensitivity test of the spatial distribution of debris flow triggering events revealed that the normalized ID thresholds show a better performance than the absolute ID thresholds, in terms of a local reduction of the number of False and Missed Alarms, and a local increase of the number of Correct Alarms (for further information see Meyer et al. submitted).

9.7 CONCLUDING REMARKS

The study defines two sets of intensity-duration (ID) thresholds for debris flow initiation in Norway. The first set is based on absolute water supply intensities while in the second set these intensities are normalized by the precipitation day normal (PDN) to account for the local climate. Beside minimum ID thresholds also medium and maximum thresholds are defined. The main conclusions from this study are:

- All thresholds obtained in this study decrease with time. Thus, over longer durations lower water supply intensities are able to trigger debris flows than over shorter durations.
- The defined absolute minimum thresholds are in the range of other minimum ID thresholds obtained for regions with a climate comparable to Norway.
- The normalized minimum thresholds obtained in this study show a strong spatial variation over the country and are highest on the Norwegian west coast.
- Statistical sensitivity analyses reveal that the normalized thresholds give better results than the absolute threshold and thus are recommended for further use.

9.8 TABLES

Table 1: Individual series of water supply data prior to a debris flow event that occurred near Vang (central Norway) on 04.05.2008: The cumulative water supply [mm] and the cumulative average intensities [mm/day] are calculated for four weeks in advance of an event. The maximal cumulative average intensity (WS_C/D) and the corresponding duration (D) (printed bold) form the ID relationship that is used as input for the threshold definition.

Duration [day] (D)	**Water supply [mm] (WS)**	**Cumulative water supply [mm] (WS_C)**	**Cumulative average intensity [mm/day] (WS_C/D)**
1_{Event}	13	13	13
2	23	36	18
3	33	69	23
4	38	107	**27**
5	21	128	26
6	23	151	25
7	20	171	24
8	8	179	22
9	0	179	20
10	0	179	18
[…]	[…]	[…]	[…]
28	38	107	27

9.9 FIGURES

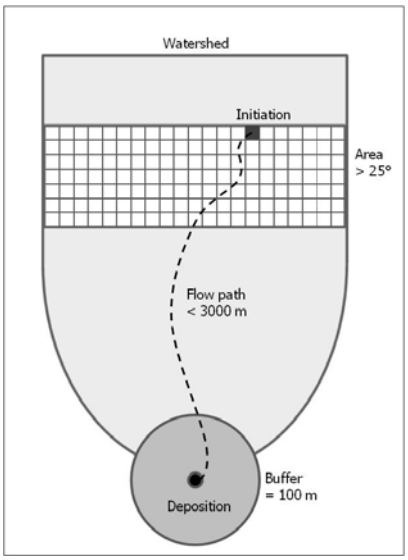


Figure 1: Schematic illustration of the parameters used for the relocation of recorded debris flows from the deposition to a likely initiation area.

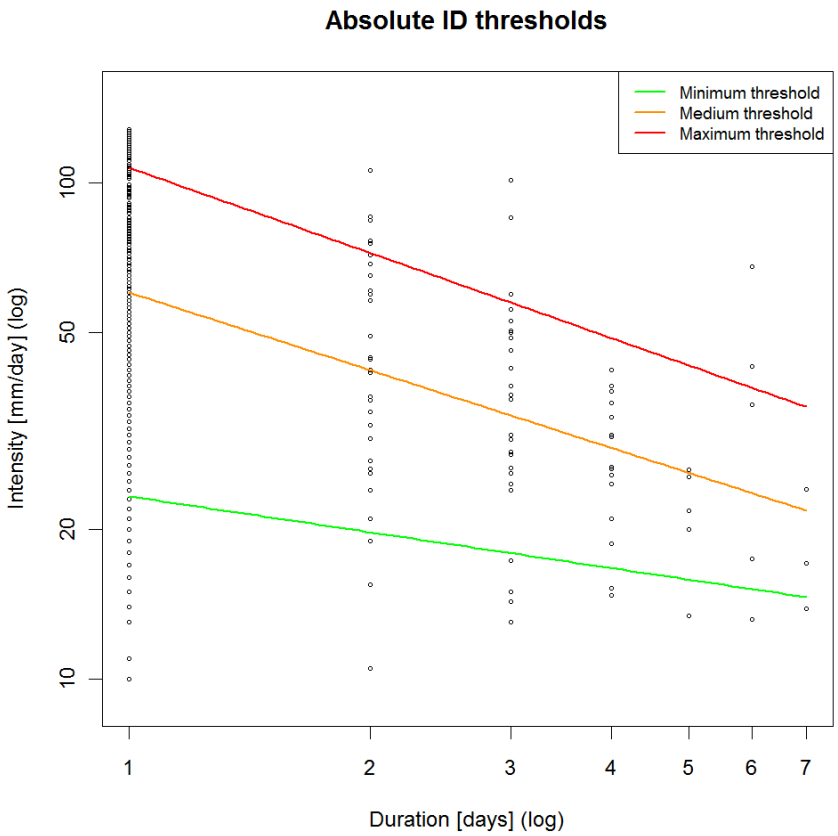


Figure 2: Absolute ID threshold curves for debris flow initiation obtained from the 3-day moving average drawn along the 0.1 percentile (minimum threshold), the 0.5 percentile (medium threshold) and the 0.9 percentile (maximum threshold) (n= 502).

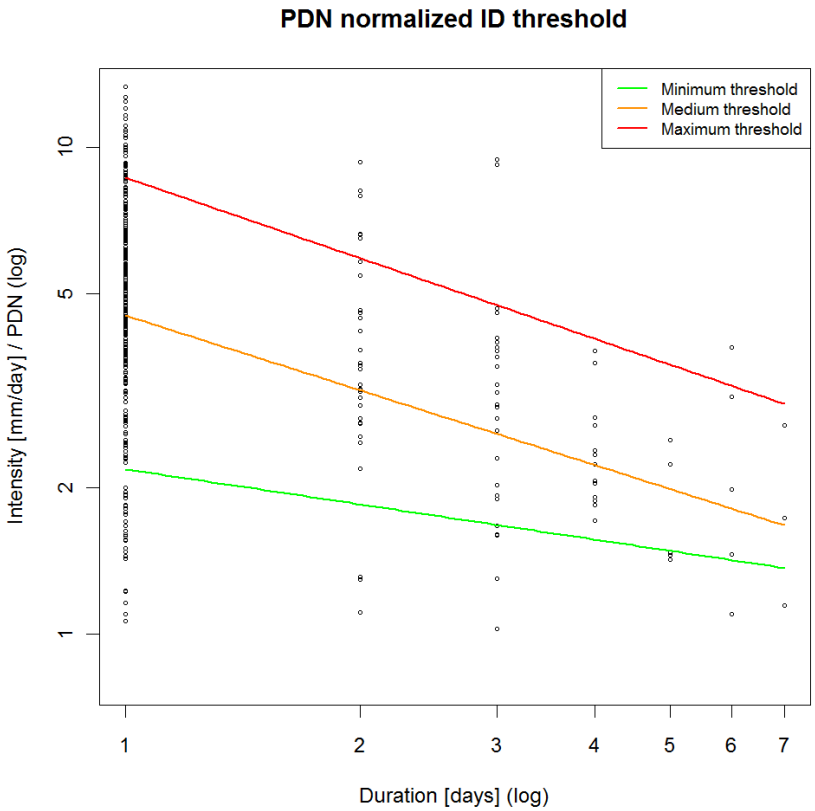


Figure 3: ID threshold curves for debris flow initiation normalized with the precipitation day normal (PDN). The curves were drawn along the 3-day moving average of the 0.1 percentile (minimum threshold), the 0.5 percentile (medium threshold) and the 0.9 percentile (maximum threshold) (n= 502).

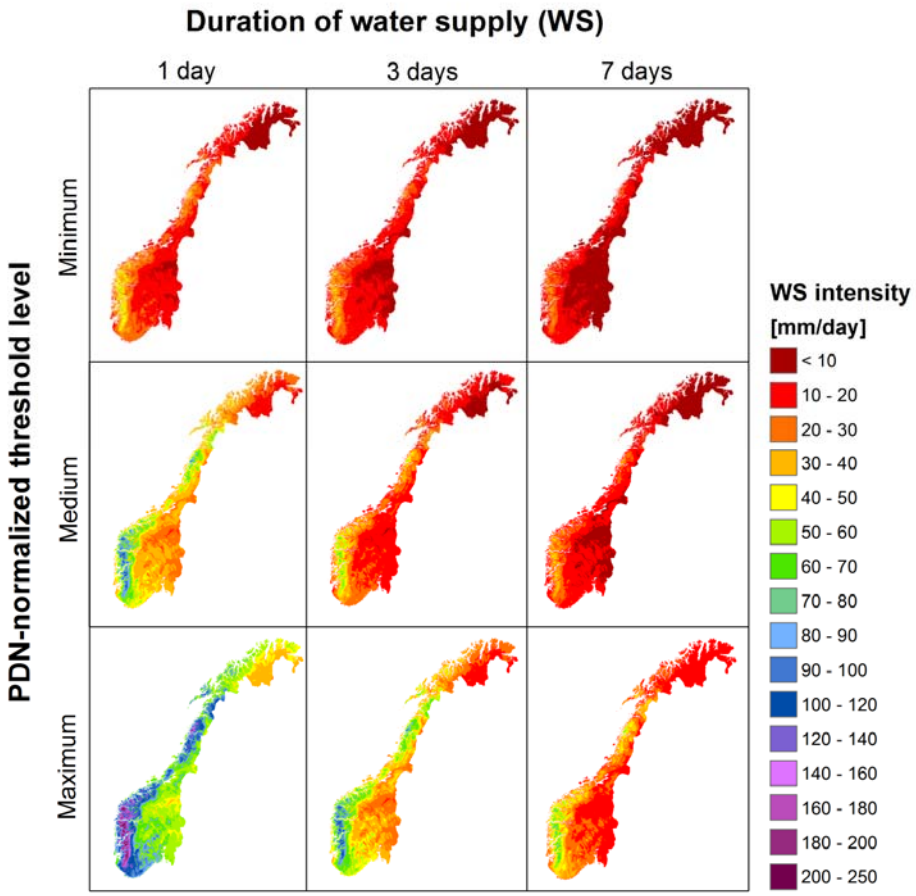


Figure 4: Spatial distribution of the debris flow triggering intensities after 1 day, 3 days and 7 days of water supply according to the normalized minimum, medium and maximum ID threshold.

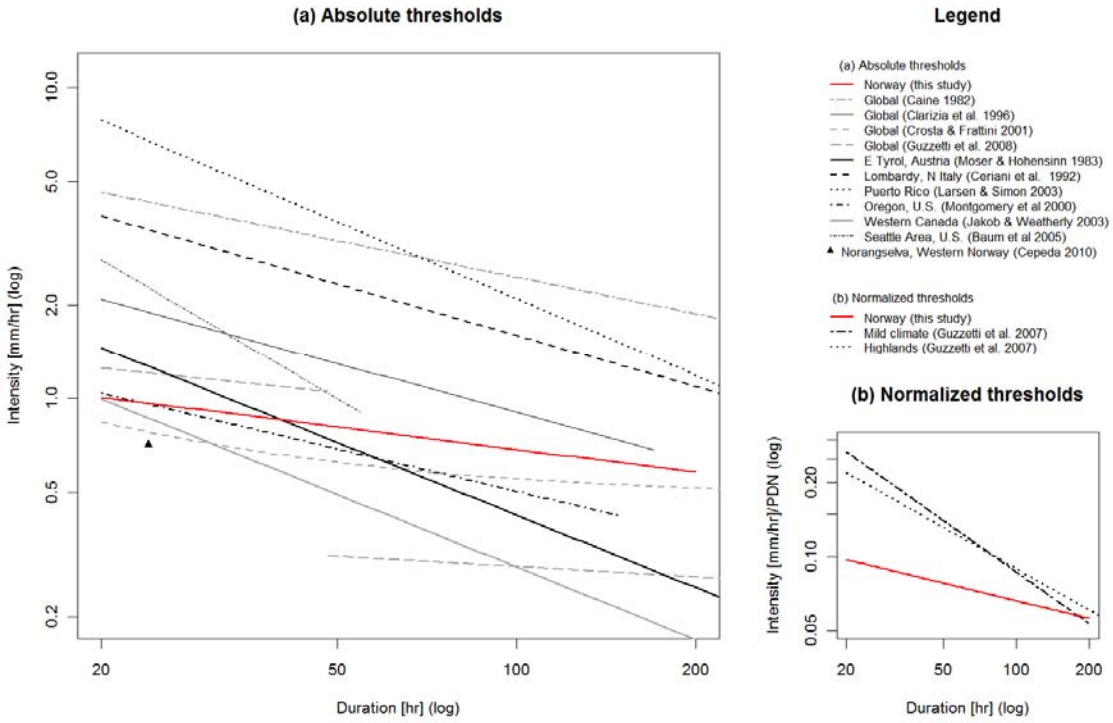


Figure 5: The obtained absolute minimum threshold (a) and the obtained normalized minimum threshold (b) in comparison to a selection of other global, regional and local minimum ID thresholds. Note that durations are given in hours [hr].

9.10 ACKNOWLEDGEMENTS

The contents of this appendix are a summary of a paper entitled "Hydro-meteorological threshold conditions for debris flow initiation in Norway" which was submitted to the Journal *Natural Hazards and Earth System Sciences* in January 2012. The authors of this article wish to thank Egil Syre (NGI) for his help with GIS tasks and José Mauricio Cepeda (NGI) for his comments on the threshold issue. Funding for this study was provided by the International Centre for Geohazards (ICG) and the InfraRisk project (http://infrarisk.ngi.no) financed by the Norwegian Research Council, the Norwegian Public Roads Administration and the Norwegian National Rail Administration.

9.11 REFERENCES

Baum, R.L., Godt, J.W., Harp, E.L. and McKenna, J.P.: Early warning of landslides for rail traffic between Seattle and Everett, Washington. Proceedings 2005 International Conference on Landslide Risk Management, 731-740, 2005.

Baum, R.L. and Godt, J.W.: Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides 7, 259-272, 2010.

Bråthen, S., Husda, J., Rekdal, J. 2008. Samfunnsøkonomisk verdi av rassikring. Noen beregninger knyttet til verdi av å unggå stengte veier. Møreforsking Molde AS, Report 0801: 59 pp. (ISSN 0806-0789/ ISBN 978-82-7830-124-1)

Cannon, S.H. and Ellen, S.: Rainfall conditions for abundant debris avalanches in the San Francisco Bay region, California. California Geology 38 (12), 267-272, 1985.

Cepeda, J.: Threshold values for precipitation-induced landslides in Norway. Proceedings of the Geoteknikkdagen 2010. Chapter 36, 2010.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C.P.: Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorol. Atmos. Phys. 98, 239-267, 2007.

Jaedicke, C. and Bakkehøi, S.: Climate database for avalanche consulting and warning in Norway. Cold Reg. Sci. Technol. 47, 171-179, 2007.

Jakob, M. and Weatherly, H.: A hydroclimatic threshold for landslide initiation on the North Shore Mountains of Vancouver, British Columbia. Geomorphology 54, 137-156, 2003.

Meyer, N.K., Dyrrdal, A.V., Frauenfelder, R., Etzelmüller, B. and Nadim, F.: Hydro-meteorological threshold conditions for debris flow initiation in Norway. NHESS, submitted in January 2012.

Montgomery, D.R., Schmidt, K.M. and Greenberg, H.M.: Forest clearing and regional landsliding. Geology 28 (4), 311-314, 2000.

Sandersen, F., Bakkehøi, S., Hestnes, E. and Lied, K.: The influence of meteorological factors on the initiation of debris flows, rockfalls, rockslides and rockmass stability. In: Senneset, K. (Ed.): Landslides. A.A. Balkema Publishers, Rotterdam, 97-114, 1996.

Wilson, R.C. and Jayko, A.S.: Preliminary maps showing rainfall thresholds for debris flow activity, San Francisco Bay region, California. Open File Report. U.S. Geological Survey, 1997.

10 APPENDIX C: PROBABILISTIC CLASSIFICATION OF THRESHOLDS FOR RAPID SOIL SLIDES IN NORWAY.

10.1 INTRODUCTION

This appendix summarizes the results of a research project presented in NGI (2012). The aim of this project was the probabilistic estimation of thresholds for triggering of rapid soil slides in Norway. A database of 281 landslide days for the period 2000-2011 was prepared with contributions from the Norwegian Water Resources and Energy Directorate - NVE, the Norwegian Public Roads Administration - SVV, Jernbaneverket and NGI reports. The database included also 15224 no-landslide days. Each event in the database was characterised by 18 different explanatory variables of meteorological and hydrological nature. Univariate and bivariate models were tested using the following classification methods: classification tree, linear and quadratic discriminant analyses, and Naïve-Bayes method. The performance of the models was tested using cross-validation and Receiver Operating Characteristic (ROC) analysis. The top ranked models were combinations of water supply and ground saturation.

10.2 DATASETS

The database comprised 281 landslide day-events and 15 924 no-landslide day-events in the period 2000-2011. Incidents in man-made slopes or not induced by water were excluded from the analyses.

Considering the wide variation in meteorological, geomorphological, and lithological conditions over Norway, the following subsets were analysed:

- Precipitation regions: regions 2, 6 & 8 (Østlandet and Vestlandet) See Fig. 1, and
- New zonation proposed by NVE based on precipitation regions and geomorphology: regions 1a, 1b & 3 (Østlandet and Vestlandet) See Fig. 2.

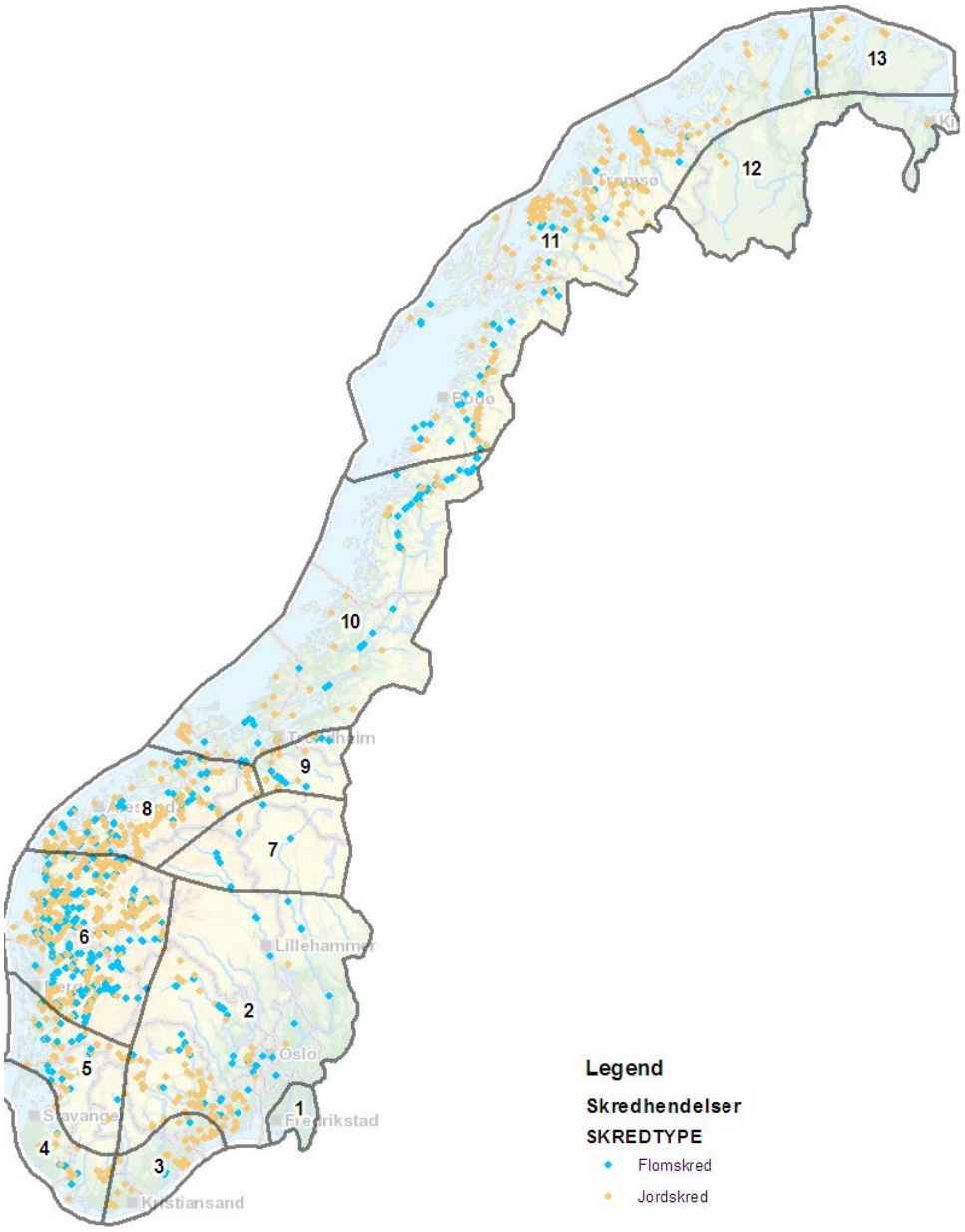


Figure 1. Precipitation regions and complete inventory of soil slides and debris flows in Norway. Blue points (710 events) are debris flows and brown points (2904 events) are soil slides. More than 90% of these events are mapped on the points were the landslide paths intersect the affected roads and railways. Considering the long runout and high height difference of most of these events, this poses a challenge for using the database for characterisation of release areas.

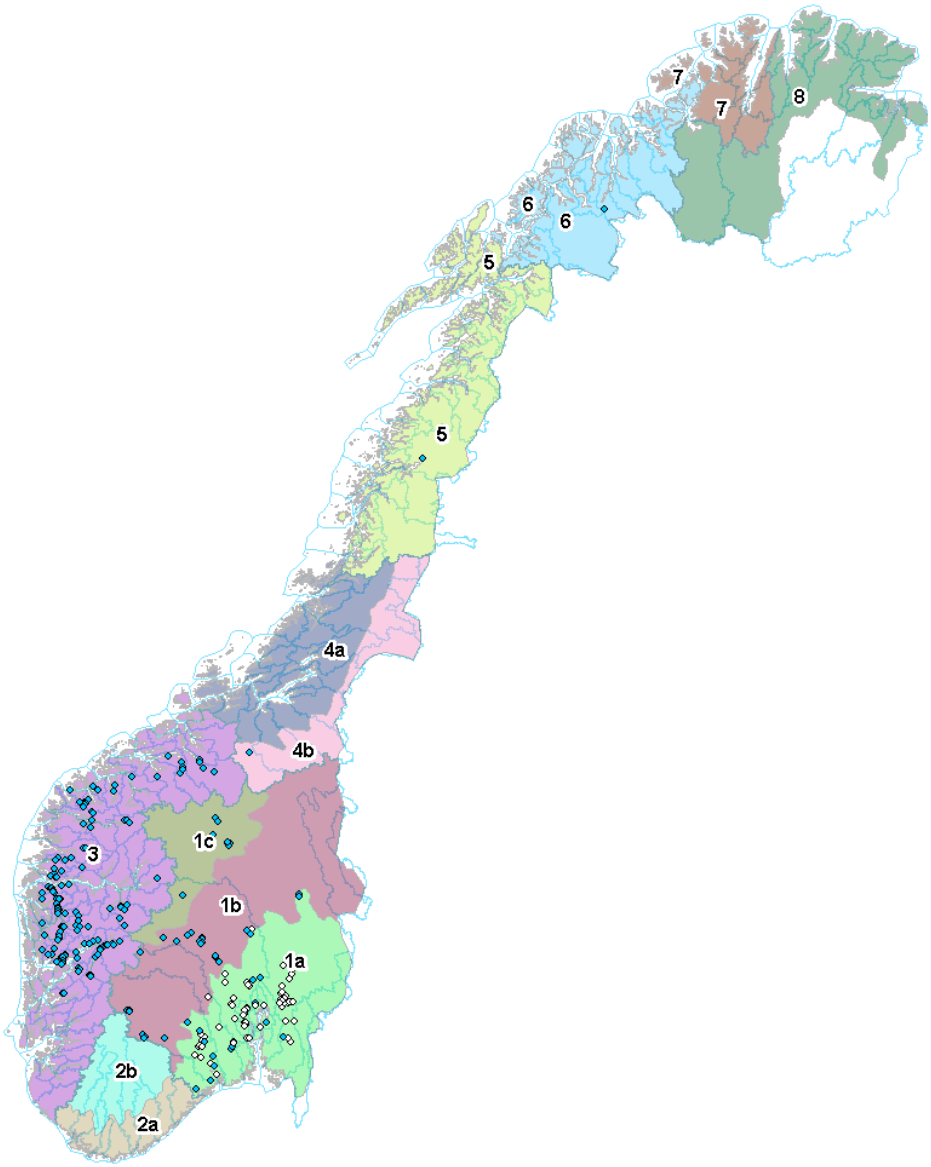


Figure 2. Location of the 281 landslide events used for estimation of thresholds. Coloured regions are new zones proposed by NVE based on precipitation and geomorphology. Blue points: landslides triggered by short-duration precipitation/water supply. White points: landslides triggered by long-duration episodes of precipitation/water supply.

Each event in the database was characterised by 18 different explanatory variables of meteorological and hydrological nature grouped in the following categories (see Table 1):

- Water supply (precipitation and snow melt)
- Runoff
- Groundwater
- Ground saturation

Table 1. Explanatory variables tested in the models. The values are extracted from a 1-km grid from SeNorge and from the output of NVE's hydrological models .

Category	**Variables**
Water supply (precipitation and snow melt)	• 1-day precipitation • 1-day water supply (precip. + snow melt) • 3-day water supply • Normalised by annual precipitation, max. 30-year value, and 200-year return period
Runoff	• 1-day runoff • 1-day runoff normalised by 30-year value • Daily change of runoff
Groundwater	• Groundwater volume • Normalised by max. 30-year value • Daily change of groundwater
Ground saturation	• Ground saturation • Ground water capacity
Other variables	• Daily air temperature • Frost depth

10.3 THRESHOLD MODELS

Two types of models were considered for the 18 different explanatory variables and the 6 subdatasets:

- Univariate (i.e., a threshold controlled by a single variable)
- Bivariate
  - Linear combination $\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 = 0$
  - Quadratic combination $\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_1 X_2 + \alpha_4 X_1^2 + \alpha_5 X_2^2 = 0$

The classification methods used were:

- Classification tree
- Linear Discriminant Analysis (LDA)
- Quadratic Discriminant Analysis (QDA)
- Naive-Bayes classification

The performance of the threshold models was evaluated by:

- A 10-fold cross-validation error (i.e., 10 random training and validation subsets)
- Area under the ROC curve

Considering the amount of analysed data (8 subsets, 18 univariate models, 153 bivariate combinations, and 4 classification methods – resulting in a total of 5040 models), it was considered neccessary to further lump the performance of the models in order to facilitate comparison for the selection of the critical explanatory variables. A performance score (*PS*) for lumping the results of all models for a single combination of explanatory variables, such that:

$$PS_i = \sum_{j=1}^{n} w_{ij}\left(1 - mer_{ij}\right)AUC_{ij}$$

Where, for the *i*-th set (combination of explanatory variables):

PS_i : performance score

For the *j*-th model:

w_{ij} : priority weight ($0 \le w_{ij} \le 1$), $\sum_{j=1}^{n} w_{ij} = 1$

mer_{ij} : cross-validation misclassification error rate

AUC_{ij} : area under the ROC curve

PS is within [0,1], 1 being the best performance.

10.4 RESULTS

The top 3 performances for each type of thresholds were:

- Univariate thresholds:
  - 1-day water supply normalised by the annual precipitation
  - 1-day water supply
  - 1-day water supply normalised by the 200-year return period event
- Bivariate thresholds: 1-day water supply normalised by the annual precipitation vs:
  - Ground water capacity
  - Ground saturation
  - Groundwater volume normalised by the max. 30-year value

The type of classification that yielded the best performance was the Naïve-Bayes classification. The method with the worst performance was the classification tree method.

As an example of the results, Figure 3 presents a bivariate threshold with 1-day water supply normalised by the annual precipitation vs ground saturation using a Naïve-Bayes classification on the subdataset on region 3.

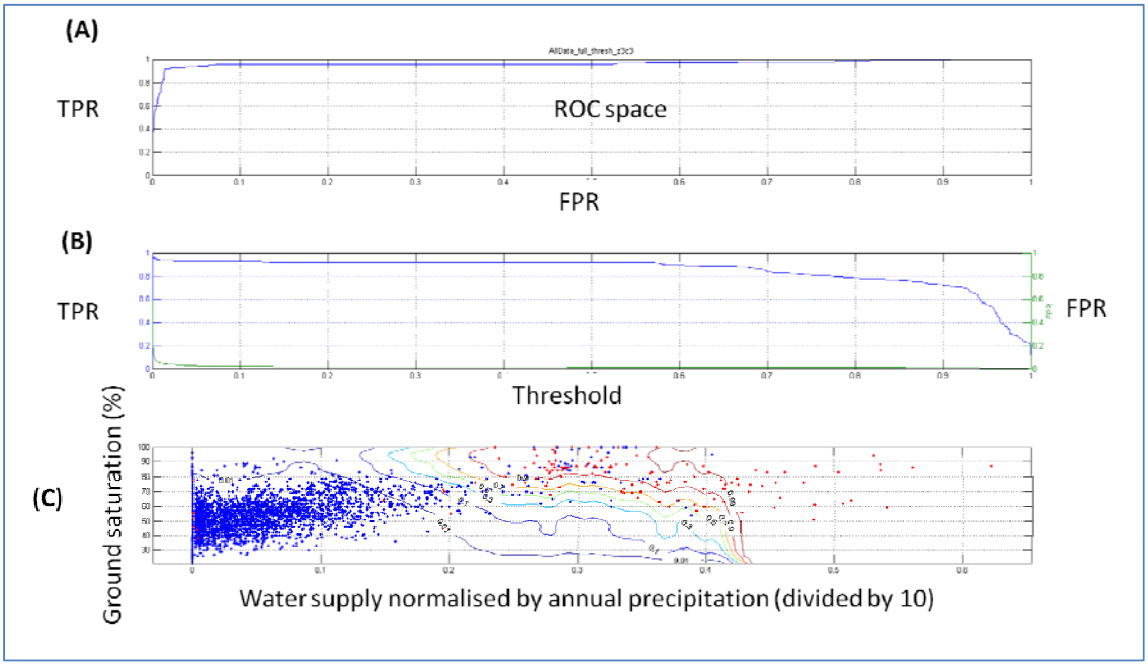


Figure 3. Results of Naïve-Bayes classification on the subdataset on region 3 (see Figure 2). (A) ROC curve. (B) Mapping between TPR, FPR and the Threshold values. (C) Dataset for Water supply vs. Ground saturation (blue – no landslides, red – landslides). Curves are constant threshold curves from 0.01 (left) to 0.99 (right).

10.5 CONCLUDING REMARKS

A database of 281 landslide day-events and 15 924 no-landslide day-events characterised by 18 hydrometeorological variables was analysed using 5040 univariate and bivariate threshold models. The best performance is for 1-day water supply combined with ground saturation parameters using the Naive-Bayes classification. The selection of threshold levels corresponding to warning levels can be based on the ROC curve of the model performance.

10.6 ACKNOWLEDGMENTS

The project presented in this appendix was financed by the Norwegian Research Council, Norwegian Water Resources and Energy Directorate - NVE, the Norwegian Public Roads Administration - SVV, and Jernbaneverket. Their support is acknowledged with thanks. Most of the datasets were also provided by NVE, SVV and Jernbaneverket.

10.7 REFERENCES

NGI (2012) Probabilistic estimation of thresholds for rapid soil-slides and -flows in Norway. NGI Document No. 20110253-00-4-R. *In preparation*.