Nat. Hazards Earth Syst. Sci. Discuss., 2, 4987–5036, 2014
www.nat-hazards-earth-syst-sci-discuss.net/2/4987/2014/
doi:10.5194/nhessd-2-4987-2014

Natural Hazards and Earth System Sciences Discussions
Open Access

Statistical modeling of rainfall-induced shallow landsliding using static predictors and numerical weather predictions: preliminary results

V. Capecchi[1,2], M. Perna[1,2], and A. Crisci[3]

[1]Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via Madonna del piano 10, Sesto Fiorentino, 50019 Florence, Italy
[2]Consorzio LaMMA, Via Madonna del piano 10, Sesto Fiorentino, 50019 Florence, Italy
[3]Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via Caproni 8, 50145 Florence, Italy

Received: 30 June 2014 – Accepted: 27 July 2014 – Published: 4 August 2014

Correspondence to: V. Capecchi (capecchi@lamma.rete.toscana.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We present a quantitative indirect statistical modeling for predicting rainfall-induced shallow landsliding. We consider as input layers both static thematic predictors, such as geomorphological, geological, climatological information, and numerical weather model's forecast. Two different statistical techniques are used to combine together the above mentioned predictors: a Generalized Linear Model and Breiman's Random Forests. We tested these two techniques for two rainfall events that occurred in 2011 and 2013 in Tuscany region (central Italy). Model's evaluation is measured by means of sensitivity-specificity ROC analysis. In the 2011 rainfall event, the Random Forests technique performs slightly better, whereas in the 2013 rainfall event the Generalized Linear Model provides more accurate predictions. This study seeks also to establish whether the rainfall-induced shallow landsliding prediction might substantially benefit from the information provided by the numerical weather model's outputs. Using the variable importance parameter provided by the Random Forests algorithm, we asses the added value carried by numerical weather forecast, in particular in the rainfall event characterized by deep atmospheric convection and heavy precipitations.

1 Introduction

In the last years, in the north-western part of Tuscany region and nearby areas, noticeable heavy rainfall events occurred (Parodi et al., 2012; Sacchi, 2012; Avanzi et al., 2013; Rebora et al., 2013; Fiori et al., 2014; Buzzi et al., 2014). During these events, the rainfall amounts and intensities triggered a great number of shallow landslides, causing damages, injuries and human losses. Steep slopes and deep valleys induced a persistently high relief of energy and a shallow landsliding susceptibility.

In this work, we considered two heavy rainfall events occurred in 2011 and 2013, that affected Lunigiana and Garfagnana in the north-western part of Tuscany region (central Italy). We carried out an analysis including a statistical modeling of spatial landslide occurrence by using two models: the Generalized Linear Model (hereafter GLM, McCullagh, 1984; McCullagh and Nelder, 1989) and the Random Forests classifier (hereafter RF, Breiman, 2001).

For both statistical models, we used, as predictors, static geographical layers (referred also as instability or predisposing factors, e.g. digital elevation model, slope, land use, see Sect. 2.2 for further details) characterizing the areas affected by the heavy precipitations from a geomorphological, geological and climatological point of view. Moreover, since recently (Schmidt et al., 2008; Segoni et al., 2009; Mercogliano et al., 2013), the Numerical Weather Prediction's (NWP) outputs are arising as a promising tool for the prediction of shallow landslides triggered by precipitation, in the statistical models we considered, as dynamical predictors, the forecast achieved by running the Weather and Research Forecasting (WRF) model (Skamarock et al., 2005; Skamarock and Klemp, 2008) for the selected dates.

For both rainfall events, we used, as ground truth, the landslides inventory maps, created via field surveys of the expert personnel of Civil Protection Office (2011 event, Lunigiana area) and of the Genio Civile Office (2013 event, Garfagnana area) a few days after the heavy precipitations.

The approach we adopted, is an attempt to conjugate and integrate the added information carried by a regional numerical weather model which operates at the meso-γ scale ($\simeq$ 2–20 km of spatial resolution), with the micro-γ scale ($\leq$ 20 m of spatial resolution, according to Orlanski, 1975) which is an average value of the mapping unit of landslide size occurring at the basin scale (Guzzetti et al., 1999, 2005). This goal is achieved without performing any downscaling of the NWP data (3 km of horizontal resolution) to a finer resolution. In this way we preserve the original information content provided by the numerical model.

Results obtained show how both statistical models (GLM and RF) perform adequately (i.e. we obtain similar results as found in previous studies) in predicting the shallow landsliding occurrence. In the 2011 rainfall event, the model based on the RF classifier performs slightly better than that one based on the GLM model, whereas in the 2013 rainfall event the GLM model gives more accurate predictions. The evaluation of the results is performed through the analysis of the Receiver Operating Characteristics Curve (ROC) in terms of the underlying area (AUC), a threshold-independent index widely used (Frattini et al., 2010).

In the discussion, we assess the relative importance of the added value provided by the numerical weather predictions in particular in the event occurred in 2011, where deep atmospheric convection, yielding high rainfall intensities ($mm\,h^{-1}$), characterized the precipitation type. Using the RF's variable importance parameter, we point out the fact that NWP data are relevant for landslide hazard mapping not only because of predictions on precipitations amounts, but also because of predictions on precipitations rate ($mm\,h^{-1}$) and on soil water content at different levels below ground.

The positive impact of mesoscale NWP's outputs, supports the reliability of numerical forecast and further confirms (Schmidt et al., 2008; Segoni et al., 2009; Mercogliano et al., 2013) its use for the setting-up of a real-time forecasting chain for the prediction of the occurrence of rainfall-induced shallow landslides over large areas (basin catchment scale).

The paper is organized as follows: in Sect. 2.1 we describe the two rainfall events, with particular regard to the meteorological and atmospheric features in terms of precipitation type and rainfall intensities. In Sect. 2.2 we describe the geographical static predictors used in the statistical models, stressing their importance for landsliding as reported in previous works. Section 2.3 provides details about the numerical weather model used to produce the meteorological dynamical predictors, namely rainfall data and soil moisture estimates, that feed the statistical models. The design and details of the statistical modeling framework based on the GLM model and on the RF classifier are described in Sect. 2.4. The preliminary results for the selected study cases are shown in Sect. 3 and discussed in Sect. 4.

NHESSD
2, 4987–5036, 2014

2 Materials and methods

2.1 Description of the study cases and of the areas of interest

As stated in the introduction, we developed the statistical modeling of shallow landsliding induced by precipitation, focusing our attention on two heavy rainfall events that occurred in the north-western part of Tuscany region (central Italy) on 25 October 2011 (Lunigiana) and on 18 March 2013 (Garfagnana). In the following two sub-sections, we describe the rainfall events from a meteorological point of view and we give a brief description of the area of interest considering geographical and geomorphological features.

2.1.1 Study case 25 October 2011

The first rainfall event, hereafter 25OCT2011, occurred on 25 October 2011 and involved the Lunigiana area belonging to the administrative province of Massa-Carrara (see inset figure in the left side of Fig. 1). The area is located along the Appennine chain and is mainly mountainous (highest peaks reaches almost 2000 m). It is very close to the Ligurian Sea gulf from which it is only a few kilometers away. Due to its orography and geographical position, the area represents a natural barrier for the Atlantic humid air masses and frequently the precipitation amounts reaches or exceeds 3000 mm per year, making this area one of the more rainy in Italy. From a hydrological point of view, it is characterized by the presence of one main river basin (Magra basin) having an area of about 992 km^2 (in the administrative province of Massa-Carrara). A detailed study on the critical thresholds able to trigger shallow landslides in this area was carried out by Giannecchini (2006). Avanzi et al. (2013) studied the fragility of the territory by analyzing the damages occurred in two heavy rainfall events in 2009 and 2010 (in this latter paper the study area was slightly larger than that one here under exam).

NHESSD
2, 4987–5036, 2014

From a geological point of view, following Di Naccio et al. (2013), Northern Apennines are a NW–SE-trending belt formed by NE-verging tectonic units stacked since the late Oligocene after the collision of the Corsica-Sardinia and Adria continental blocks. Main tectonic units are (i) the Liguride allochthon, (ii) the Subligurian unit, and (iii) the Tuscan unit (for a comprehensive synthesis and review see Argnani et al., 2003 and references therein).

From a meteorological point of view, the 25OCT2011 rainfall event was deeply investigated by Buzzi et al. (2014) using a numerical weather model and by Rebora et al. (2013) using the measurements available from a large number of sensors, both ground based and space-borne. In this latter paper, the authors concluded that the large scale features of the event and the complex geographical characteristics of the area, determined the conditions for the persistence of heavy precipitation systems over the same region, i.e. organized and self-regenerating mesoscale convective system (MCS). In the area of interest, rainfall amounts were registered by the remotely automated weather station network operated by the National Civil Protection Department. The maximum cumulative rainfall was recorded at the Pontremoli rain-gauge (Magra river valley) with maximum rainfall rates of 374 mm $(24\,\mathrm{h})^{-1}$, 317 mm $(12\,\mathrm{h})^{-1}$, 243 mm $(6\,\mathrm{h})^{-1}$, 158 mm $(3\,\mathrm{h})^{-1}$ and 67 mm $(1\,\mathrm{h})^{-1}$. As stated in Rebora et al. (2013) this rainfall event has the key atmospheric conditions for heavy precipitations and severe flood events over complex orography, i.e.: (i) conditionally or potentially unstable air masses, (ii) moist low-level winds, (iii) steep orography that helps to release the conditional instability associated with the low-level jet, and (iv) a slowly evolving synoptic pattern that slows the advance of the heavy precipitation system, hence increasing their persistence.

The landslide inventory map for this event was created by the expert personnel of the Regional Civil Protection Office. The map reported 243 shallow landslides in an area of about 212 km^2 (see the minimum bounding rectangle in the left side of Fig. 1) while the convex hull where landslides were observed has an area of about 123 km^2.

NHESSD
2, 4987–5036, 2014

2.1.2 Study case 18 March 2013

The second rainfall event, hereafter 18MAR2013, occurred on 18 March 2013 and involved the Garfagnana area belonging to the administrative province of Lucca (see inset figure in the right side of Fig. 1). This is mainly a mountainous area (the average elevation of the main catchment is 717 m) and, as Lunigiana, also this area is very close to the Ligurian Sea gulf. Long-time series of precipitation data recorded by local rain-gauges report yearly average about 2000–2300 mm (Avanzi et al., 2013). Hydrologically, it is characterized by the presence of one main river basin (Serchio basin) with an area of about 1565 km^2 plus several other minor rivers.

Geological features of the area are very similar to the ones described in Sect. 2.1.1 for Lunigiana. For an extensive and deeper analysis see Di Naccio et al. (2013) and references therein.

The 18MAR2013 rainfall event occurred during the month of March 2013, which recorded the monthly highest precipitation amounts over the last 30 years (Regione Toscana, 2013), for what concerns the north-western part of Tuscany and the Serchio and Magra basins in particular. During the period 5–19 March 2013, the rain-gauges belonging to the administrative province of Lucca and to the Serchio river basin, registered about 310 mm of precipitation against an average monthly value of about 80 mm (climatology is based on the period 1983–2012). This relevant amount of precipitation was the result of two major rain-storms that affected the area of interest: the first one occurring in the period 11–12 March 2013, the second one occurring on 18 March 2013 (the one under exam here). Due to the high degree of saturation of the soils and due to the surface runoff, on 18 March 2013, several regional hydrometers exceeded the warning levels and flooding alerts were issued by the local Civil Protection Office (Regione Toscana, 2013) for 5 rivers (Ombrone Pistoiese, Bisenzio, Serchio, Magra, Cecina). As can be argued from synoptic analysis, the 18MAR2013 rainfall event was determined firstly by a warm front over the northern Tyrrhenian Sea and Ligurian Sea, driven by a deep low over Great Britain (988 hPa at 06:00 UTC).

NHESSD
2, 4987–5036, 2014

Then in the second part of the day the cold front hit the Tuscany region, while the precipitations ended by the late evening/night. The regional rain-gauges network registered hourly precipitation intensities up to $31\,\mathrm{mm\,h^{-1}}$ (rain-gauge located near the Monte Macine peak at 1480 m a.s.l.), whereas the average hourly precipitation intensity among the available pluviometers was about $9\,\mathrm{mm\,h^{-1}}$. See Table 3 for summary statistics on observed and modelled 1 h precipitation intensities.

The landslide inventory map for this event was created by the expert personnel of the local Genio Civile Office. The map reported 127 shallow landslides in an area of about $2038\,\mathrm{km^2}$ (see the minimum bounding rectangle in the right side of Fig. 1), while the convex hull where landslides were observed has an area of about $1416\,\mathrm{km^2}$.

2.2 Description of the geographical static predictors

In the following, we list the geographical static predictors considered in the statistical modeling of landslide hazard. We divided them in four groups: geomorphology, hydrology, geology and climate related predictors. The layers are raster datasets and were produced using GIS technologies. The pixel resolution of each layer is 30 m, if not otherwise specified.

An extensive and exhaustive discussion about the choice of the input parameters (typology and number of predictors) in susceptibility assessment studies can be found in Catani et al. (2013) and we used this work as a main reference for the choice of the predictors. Here we recall that the usefulness of some predictors is still debated and can depend on the methodology adopted or the area of investigation and its landslide features. Moreover the number of predictors taken into account is also debated and it has been also found that increasing the number of predisposing factors could lead to a worsening of the prediction accuracy (Floris et al., 2011). For this reason, in landslide susceptibility assessment, it is important to implement an automated procedure for the selection of the meaningful variables. As discussed in more detail in Sect. 2.4, we chose two suitable methods: the logistic regression with an AIC selection (the GLM model) and the RF algorithm, since it naturally estimates the variable's importance for predictive classification.

All the predictors here described are schematically listed in Table 4.

Geomorphology-related predictors:

- Elevation (DEM): this dataset is a hydrologically corrected 30 m Digital Elevation Model, resampled from an original database produced at 10 m of resolution. Elevation is a very common parameter often taken into account in landslide susceptibility assessments (Catani et al., 2013; Felicísimo et al., 2013), since it is related to several predisposing factors such as average precipitation, vegetation, etc...
- Altitude above channel network (AaCN): the algorithm for producing the altitude above channel network uses the channel network for streams. It measures the altitude for each grid cell of the DEM to the nearest channel network elevation. A splines interpolation surface is created, called Channel Network Base Level, then this value is subtracted from the DEM to obtain the Altitude Above Channel Network. This parameter has been used in recent works of landslide susceptibility assessment by Marjanovic et al. (2011) and Mărgărint et al. (2013)
- Aspect (ASP): it represents the orientation of each cell with respect to the adjacent cells. It influences the landslide susceptibility because it determines how the terrain is exposed to rainfall and solar radiation (Guzzetti et al., 1999) and thus to soil water content
- LS factor (LSF): it represents the topographic factor (length-slope factor) from the Revised Universal Soil Loss Equation (RUSLE) according to Moore and Wilson (1992). Despite the fact that the RUSLE equation is commonly used to predict soil erosion on an cell-by-cell basis, recently a high correlation has been found between (R)USLE-based soil erosion map and landslide locations (Pradhan et al., 2012)

- Planar curvature (PLAC): basically it is the second derivative of DEM and corresponds to the concavity/convexity of the land surface measured perpendicular to aspect, i.e. parallel to the contour. Catani et al. (2013) used this parameter (and its standard deviation) in their landslide susceptibility study based on RF model

- Profile curvature (PRFC): it is a common morphological layer derived from the digital elevation model. It describes the shape of the relief in the direction of the steepest slope. It corresponds to the concavity/convexity of the land surface measured parallel to aspect, i.e. perpendicular to the contour. It is known to affects the flow velocity of water and influences erosion and deposition. It has been used in several landslide assessment studies among which we recall Catani et al. (2013) who used this parameter (and its standard deviation) in their landslide susceptibility study based on RF model

Hydrology-related predictors:

- Convergence index (COVI): this index represents the convergence/divergence with respect to overland flow. It is similar to plan or horizontal curvature, but gives much smoother results. The calculation uses the aspects of surrounding cells and looks to which degree the surrounding cells point to the center cell. The result is given as percentages, negative values correspond to convergent flow conditions, positive to divergent ones. This predictor has been recently used in landsliding susceptibility maps by Nefeslioglu et al. (2011)

- Time of concentration (ToC): it measures the response of a watershed to a rainfall event. It measures the time (in hour) needed by a rainfall drop to reach the closure of a watershed from the farther point of it. It is a function of the topography, geology, and land use within the watershed. It is considered as one of the most critical factor for the estimation of the duration of the triggering rainfall (D'Odorico and Fagherazzi, 2003)

NHESSD
2, 4987–5036, 2014

- Topographic Wetness Index (TWI): it is commonly used to quantify topographic control on hydrological processes. It is calculated by using the formula:

$$\mathrm{TWI} = \frac{a}{\tan\beta} \quad (1)$$

where a is the local upslope contributing area and β is the local slope angle. This index is related to the soil moisture (Nefeslioglu et al., 2008; Yilmaz, 2010). The main limitation of the above formula is that it assumes a steady-states conditions and uniform soil properties. However researchers denote that the formula is applicable in a wide range of cases and it has been used in assessment in landslide susceptibility mapping (Nefeslioglu et al., 2012).

- Distance from drainage channel network (DfCN): this is the euclidean distance from rivers network. The distances from rivers have been evaluated by computing the minimum distance between cells and the nearest watercourse. This layer has been considered in similar works as a predisposing factor, because it takes into account possible activating mechanism related to erosion along the slope foot (Mancini et al., 2010). Recently it has been used by several authors as a predictor in landslide susceptibility mapping (Floris et al., 2011; Catani et al., 2013; Demir et al., 2013; Devkota et al., 2013).

Geology-related predictors: this group of predictors includes data from two regional databases produced by the Tuscany administration: the Regional Geological Continuum (scale is 1 : 10 000) and the regional pedological database (scale is 1 : 50 000). The Regional Geological Continuum is the joint effort of several local institutions (universities, research institutes, private entities, coordinated and leaded by the regional administration) and was recently updated with extensive field campaigns covering about 70 % of the territory. This database is freely available through web facilities. The regional pedological database (level 2) has been revised during the period 2009–2012. It was derived using data collected over sample areas of the territory. On average,

NHESSD
2, 4987–5036, 2014

the sample areas extent were about 15–25 km^2 and 20 to 40 observations were performed with the standard of 2 to 4 vertical profiles. The controls consisted of soil stratigraphic profiles, described, sampled and analyzed from wells or exploratory drillings. In a second stage of the work, an unsupervised classification of the whole territory was performed and further corrected by expert personell.

- Distance from main tectonic features (DfTF): this is the euclidean distance from main tectonic features. This layer has been used by Costanzo et al. (2012) for landslide susceptibility modelling on large scale, resulting as an effective factor for translational slides.
- Bedrock litho-technical map (BLT): it comprehends 15 different classes of bedrock based on litho-technical properties derived from bibliography. This layer is time-invariant and it has been considered as a relevant causal factor in predictive hazard models assuming that future landslide are likely to occur in the past and present instability sites (Guzzetti et al., 1999). Catani et al. (2005) acknowledged the bedrock lithology as a strong controlling factor on landslide occurrence in their study for the Arno river basin (Tuscany region)
- Landslides main scarps (LMS): this layer represents the exposed portions of the surface of rupture. These features are obtained with automated procedures from landslide crowns and DEM
- Soil permeability (SKST): this predictor is derived from the regional pedological database and has been determined using HYRES pedo-transfer function (PDTf). The term "permeability" as used in soil surveys, indicates saturated hydraulic conductivity (Ksat). In other words, it indicates the rate of water movement, centimeters per hour, when the soil is saturated.
- Landslides and superficial deposits (LaSD): this layer takes into account the presence of landslide bodies, or areas where superficial formations (debris cones,

NHESSD
2, 4987–5036, 2014

talus, colluvial and eluvial deposits, etc...) outcrop. The use of this layer in susceptibility assessment studies is justified by the hypothesis that future landslides will be likely to occur under the same conditions that led to past landslide events (Varnes et al., 1984; Carrara et al., 1991)

– Slope Structural Setting (SSS): it represents the relation between the structural setting and the slope aspect (Cruden and Hu, 1998). This factor is rarely considered in large scale susceptibility analysis due to the difficulty of data acquisition and its expression in a continuous surface (Atkinson and Massari, 1998; Guzzetti et al., 1999; Donati and Turrini, 2002). In this study the SSS factor was obtained by the spatialization of the attitude data available in the regional database, taking into account all the elements that lead to the rupture of the geological substrate continuity. The continuous surface realized was then combined with the slope aspect and slope gradient to obtain information about the relation between landslides and different combinations of slope structural setting.

Climate-related predictors: recently rainfall climatology has been considered into landslide susceptibility models as a predisposing factor instead as a triggering factor (Schicker and Moon, 2012; Catani et al., 2013). In fact the average precipitation values describe the attitude of the territory to be hit by a storm of a given type. In the following, we included a set of variables accounting for the precipitation amount (expressed in mm) of a rainfall event occurring in a defined time interval (expressed in hours) and having a defined returning period (expressed in years). In this, our predictor is slightly different from that one considered by Catani et al. (2013) who evaluated the returning period of a defined precipitation amount occurring in a defined time interval.

– Rainfall 12, 24, 48, 96 h duration and 100 years return period (R12, R24, R48, R96): this dataset is the result of rainfall frequency analysis (Baldi et al., 2014). It estimates the amount of rainfall falling at a given point for a specific duration and returning period. In the present study, the durations considered are 12, 24, 48 and 96 h and the returning period is 100 years. It was derived from statistical analysis

NHESSD
2, 4987–5036, 2014

of rainfall time series of regional rain-gauges network (30 years minimum), that were interpolated over the area of interest.

Besides the above mentioned layers, we included in the static predictors two additional thematic maps:

- Corine land cover (COR): land cover provides information on vegetation and takes into account human activity on hills slope. It is considered a predisposing factor and has been used for landslides probability of occurrence mapping (Varnes et al., 1984; Costanzo et al., 2012; Catani et al., 2013). In the present study we derived a raster layer of land cover with 45 classes starting from the original Corine dataset (Bossard et al., 2000) at scale 1 : 10 000.

- Vegetation Index (EVI): the EVI vegetation index is derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board of the Terra and Aqua satellite (Huete et al., 2002). The index is designed for providing accurate measurements of regional to global scale vegetation dynamics (phenology). Conceptually it is complementary to the well known Normalized Difference Vegetation Index (NDVI) from which differs because it is more responsive to canopy structural variations, including leaf area index (LAI), canopy type, plant physiognomy, and canopy architecture (Gao et al., 2000). Formally it is a difference of Near Infrared, red and blue atmosphere-corrected surface reflectances. In the present paper we used a layer derived from the temporal climatology of the index, using the available satellite imagery for the time series starting from February 2000 and ending in December 2013. Time series data are aggregated to 16 days to minimize cloud contamination. The spatial resolution of the layer is 250 m. Vegetation status, density and health is considered a predisposing factor for shallow landslide and debris flows because it is basically a proxy for wetness. It reflects the variation in subsurface water and because deep-rooted vegetation bind colluvium to bedrock. Vegetation index (namely NDVI and in particular its radiometric signature), has been used as an aid to the visual detection of landslides and for the semi-automatic classification of satellite images into stable or unstable slopes (Borghuis et al., 2007; Mondini et al., 2011; Guzzetti et al., 2012).

2.3 Description of the NWP model and of the numerical weather predictors

The limited-area numerical model used in this study is the Weather and Research Forecasting (WRF) model (Skamarock et al., 2005; Skamarock and Klemp, 2008). It is the result of the joint efforts of US governmental agencies and university. It is a fully compressible, Eulerian, non-hydrostatic mesoscale model, specifically designed to provide accurate numerical weather forecast both for research activities, with the dynamical core Advanced Research WRF (ARW), and for operations, with the dynamical core Non-hydrostatic Mesoscale Model (NMM). In the present work we used the WRF-ARW core updated at version 3.5 (April 2013). The model dynamics, equations and numerical schemes implemented in the WRF-ARW core are fully described in Skamarock et al. (2005), Klemp et al. (2007) and Skamarock and Klemp (2008). The model physics, including the different options available, is described in Chen and Dudhia (2000).

A summary of the model's settings chosen for the present study is shown in Table 1, while the geographical extent of the simulation area is depicted in Fig. 2. Here we briefly recall that the horizontal spatial resolution adopted (3 km) is known to be adequate to resolve explicitly the convective processes (Kain et al., 2008; Bryan and Morrison, 2012).

Initial and lateral boundary conditions were obtained from the ECMWF-IFS (European Centre for Medium-Range Weather Forecasts-Integrated Forecasting System) global model. The spectral resolution of the global model is T1279, which roughly corresponds to 16 km of horizontal resolution; vertical levels are 91. Since one of the main purposes of this work is to investigate the potential ability of the regional numerical model to predict in advance possible landslides triggered by heavy rainfall, as boundary conditions, we used the forecast (not analysis) provided by the ECMWF-IFS model. The analysis time is 00:00 UTC 24 October 2011 for the 25OCT2011 event and

00:00 UTC 17 March 2013 for the 18MAR2013 event. The length of the simulations is 48 h for both events.

The choice of the geographical static instability factors has been discussed in Sect. 2.2. For what concerns the NWP predictors, in this preliminary stage of the investigation, we subjectively decided to include a minimal set of explanatory variables, namely: precipitation amounts cumulated over the rainfall event, mean and maximum hourly precipitation intensity, mean and maximum soil moisture in four layers below ground. Soil moisture is evaluated in the following four layers: 0–10, 10–40, 40–100 and 100–200 cm below ground. This is the partition of soil implemented in the Noah land surface model (Chen et al., 1996) and the WRF model incorporates and runs this model for what concerns the the physical processes occurring in the interface between land and the near surface atmosphere. A summary of the meteorological predictors is reported in Table 4 (bottom part of the table)

2.4 Description of the statistical modeling of landslide hazard

For the two rainfall events under exam, we developed a landslide hazard modeling based on a quantitative indirect statistical model (Carrara et al., 1991; Guzzetti et al., 1999). In other words, using separately the GLM model and the RF classifier, we construct a statistical functional relationship between instability factors (such as geological, geomorphological, climatological thematic layers and NWP outputs) with the distribution of landslides as obtained from the event inventory maps. A consequence of this approach is the mapping unit which is forced to be grid-cells. It is important to underline again that no statistical downscaling is performed to nudge the NWP outputs (3 km of horizontal resolution) to the resolution of the static instability factors (30 m of horizontal resolution). The final result of the modeling is a map showing the classification of the area of interest into domains of different hazard degree ranging between 0 (stable slopes) to 1 (unstable slopes). In bibliography this type of map is also referred as landslide hazard map. Schematically a flow chart of the forecasting chain of the statistical modeling is sketched in Fig. 3. Summarizing, we developed, implemented and tested

NHESSD
2, 4987–5036, 2014

two different statistical models: one is based on the GLM model, the other is based on Breiman's RF. The GLM was chosen because it is widely known and acknowledged in landslide susceptibility mapping (see below for references). The RF classifier was chosen because it is very flexible, recently used in landslide susceptibility mapping and has interesting and useful diagnostics (see below). No interactions were implemented between the two models.

The GLM model (McCullagh, 1984; McCullagh and Nelder, 1989) is a statistical technique used to model the relation between a response variable L and a set of explanatory variables $\{X_i\}$, $i = 1, \ldots, n$. In the present case, L is the presence/absence of landslide, while the $\{X_i\}$ variables are the static parameters detailed in Sect. 2.2 and the NWP outputs detailed in Sect. 2.3. The GLM with a logit link function is one of the most frequently used techniques in landslide susceptibility modeling and it has been largely and successfully applied; see for example the review paper by Brenning (2005). In the present work, logistic regression is performed after applying an automatic stepwise backward variable selection based on the Akaike Information Criterion (AIC).

The RF classifier (Breiman, 2001) belongs to the family of machine learning algorithms. It is based on classification trees (Breiman et al., 1984) and on the idea of bagging (i.e. bootstrap-aggregation) predictors (Breiman, 1996). A RF is an ensemble of classification trees, where each tree is constructed from a random subset of the observations (i.e. the dependent variable) and at each node of the tree only a random subset of the predictors (i.e. the independent variables) is used. The data not chosen to construct the tree ("out-of-bag") is used to asses the predictive skill of the tree. The most common classification among all the tree is the prediction of the RF.

Schematically some features and advantages of the RF technique are: (a) it handles both continuous and categorical predictors naturally, (b) no formal distributions of variable's predictors is assumed, (c) it has an automatic variable selection and handles missing values, (d) it does not need a cross-validation of the results but has a built-in estimate of model's accuracy, (e) there is little need to fine-tune parameters to achieve excellent performances, (f) it incorporates highly non-linear interactions among predictors

NHESSD
2, 4987–5036, 2014

and (g) it is designed to work with “wide” data, i.e. when the cardinality of the predictors is much larger than the cardinality of the variable to be predicted (Díaz-Uriarte and De Andres, 2006). To our purposes, an important feature of the RF algorithm is the variable importance, which is a natural output of the procedure and measures the deterioration of the predictive ability of the model when each predictor is replaced in turn by random noise.

This method has been extensively used in bibliography in a variety of applications ranging from bioinformatics (Díaz-Uriarte and De Andres, 2006) to remote sensing (Pal, 2005; Gislason et al., 2006; Ghimire et al., 2010) and ecology (Cutler et al., 2007; Peters et al., 2007; Moriondo et al., 2008) just to mention an, incomplete, list of application fields. For what concerns mapping of landslides, this method was used by Stumpf and Kerle (2011a, b), who used the RF technique to implement an automatic landslide inventory mapping on the basis of very high resolution remote sensing imagery. Vorpahl et al. (2012) used the RF classifier (and several other statistical methods) to analyze the driving factors of natural landslides. They took into account, as predictors, the terrain attributes derived from a digital elevation model and trained the RF model on a set of five historical landslide inventories. Brenning (2005) applied the RF classifier to produce susceptivity maps in Ecuador using geomorphometric attributes and information on land-use. Recently, Catani et al. (2013) applied the RF algorithm to produce landslide susceptibility maps for the Arno river basin (about 9100 km^2) at different mapping unit ranging from 10 to 500 m. They considered a variety of predisposing factors mainly related to the lithology, the land use, the geomorphology, the structural and anthropogenic constrains.

Apart from the literature survey in RF area, we stress the fact that we applied the RF method in a “black box” approach and further work is needed to properly use this powerful and easy-to-use tool. This especially for what concerns the choice of the predictors and the choice of the number of input variable tried at each split of the classification tree. Nevertheless, as it is shown in Sect. 3.2 and discussed in Sect. 4, this method provided interesting results and gave hints for meaningful discussions.

NHESSD
2, 4987–5036, 2014

In the present work, we used the R translation (Liaw and Wiener, 2002) of the original RF code developed by L. Breiman and A. Cutler.

In both statistical methods, GLM and RF, we used, as predictors, the same set of layers detailed in Sects. 2.2 and 2.3 and summarized in Table 4.

3 Results

Since one of the crucial points to properly forecast the rainfall-triggered landslides is an accurate prediction of spatial patterns and temporal intensity of rainfall (Crozier, 1999), in Sect. 3.1 we briefly present the validation of the WRF predictions for the selected dates. In Sect. 3.2 we present the landslide hazard maps for the two events along with their accuracy in terms of ROC plots and the corresponding underlying area.

3.1 Evaluation of the forecasting skills of NWP outputs

Using the remotely automated weather station network operated by the National Civil Protection Department, we were able to evaluate the predictive skills of the WRF model in terms of quantitative precipitation forecast (QPF).

Considering the rainfall occurred during the 25OCT2011 event (from 00:00 UTC 25 October 2011 to 00:00 UTC 26 October 2011), it was possible to collect 20 rain-gauges recording precipitation every hour (see the locations of the rain-gauges in Fig. 4). For a visual comparison between the rainfall data simulated by the WRF forecast with the observed rainfall data collected in the 20 rain-gauges see Fig. 5a, for WRF data and Fig. 5b for observations. The ability of the model to simulate the precipitation's amount was analyzed using the contingency tables (Wilks, 2011) for selected rainfall's thresholds. For each rain-gauge locations, we extracted the predicted values of the numerical simulation and compared them with the observed rainfall amounts (24 h accumulated precipitation). In Fig. 6, we show the False Alarm Rate and Probability of Detection for selected rainfall thresholds. In Table 2, we show the descriptive statistics (average

NHESSD
2, 4987–5036, 2014

values and percentiles) of the observed and modelled rainfall data for the selected rain-gauge locations. From the analysis of the previous plot and table, it is quite clear how the model largely underestimates the rainfall amounts for high thresholds (greater than 50 mm) and overestimates precipitation for low thresholds (roughly precipitations below 50 mm or similarly below the 1 quantile of observed data). Nevertheless if we extract the modelled data in the larger area affected by the 25OCT2011 event (see the inset rectangle in the picture in the left side of Fig. 1), we can see from Table 2 the shift of the predicted data towards higher values. This can be addressed to a lack of the model in predicting the exact locations of the deep convection processes. On the other hand, we can state that the model is able to capture the characteristics of heavy rainfall event in the area of interest. In particular the maximum value obtained in the model data (218 mm 24 h^{-1}) is similar to that one obtained by Buzzi et al. (2014), who analyzed the same rainfall event with the ISAC convection-permitting MOLOCH model (Buzzi et al., 2004). In their paper, the authors found a maximum rainfall amount of 286 mm 24 h^{-1}. It has to be noticed, nevertheless, that they used the ECMWF-IFS analysis at 12:00 UTC 24 October 2011 (instead of 00:00 UTC 24 October 2011 as done here) and that their model horizontal resolution is 1.5 km (instead of 3 km as setup here).

For the 25OCT2011 event, a special regard was devoted to the verification of the rainfall intensities (mm h^{-1}) predicted by the model. In the rain-gauges dataset we can observe very high rainfall intensities (up to 67 mm h^{-1} in the Pontremoli rain-gauge station). On average, over the 20 rain-gauges considered, the Root Mean Square Error (RMSE) of modelled data is around 6 mm h^{-1} (summary statistics of both observed and modelled 1 h rainfall amounts are shown in Table 3).

For the 18MAR2013 event it was possible to collect 60 rain-gauges recording precipitation every hour (see the locations of the rain-gauges in Fig. 4). As shown for the 25OCT2011 case, for a visual comparison between the rainfall data simulated by the WRF forecast with the observed rainfall data collected in the 60 rain-gauges see Fig. 5c for WRF data and Fig. 5d for observations.Beside the visual comparison, here we show the results obtained when validating the modelled data, output of the WRF simulation, with respect to these observed data. Considering the rainfall registered by the rain-gauges during the event (from 00:00 UTC 18 March 2013 to 00:00 UTC 19 March 2013), the distribution, in terms of percentiles, of the model data and of the observed data is summarized in Table 2. The RMSE is 40 mm 24 h^{-1} and the correlation coefficient is 0.63. The False Alarm Rate and Probability of Detection plots are shown in Fig. 7 where FAR and POD skill scores are computed and plotted against selected thresholds ranging from the rounded minimum (10 mm) to the rounded maximum (130 mm) of the modelled data.

3.2 Evaluation of landslide hazard maps

Considering the extent of the areas of interest, the resolution of both static predictors and NWP outputs and considering the statistical models adopted, landslide hazard maps were produced requiring less than 10 min of CPU time on a HPC multi-core Linux server.

The results for the study case 25OCT2011 are shown in Fig. 8a for the GLM model and 8b for the RF model. While the results for the study case 18MAR2013 are shown in Fig. 9a for the GLM model and 9b for the RF model. In the maps, we subjectively masked out the pixels where slope is below 6 %. Corresponding ROC plots are shown with the values of the AUC for each curve. In the maps, values range from 0 (green color) indicating pixels classified as stable slopes to 1 (light gray color) indicating pixels classified as unstable slopes. For the study case 25OCT2011 AUC value is 0.909 for the GLM model, whereas it is 0.968 for the RF classifier. For the 18MAR2013 study case, AUC values are lower: 0.833 for GLM, 0.764 for RF.

As outlined previously in Sect. 2.4, one of the main features of the RF algorithm is the variable importance output. It measures the importance of each variable to perform a correct classification in the tree when the model is validated on the OOB (out-of-bag) data. Alternatively its measure is based on the decrease of classification accuracy when the variables in a node of a tree are permuted randomly (Breiman, 2001). Among the four different measures of variable importance implemented in the code, we chose the default one: increased node purity (IncNodePurity), see RF's manual (Breiman, 2002) for details.

Figure 10 shows the variables importance for the two study cases: 25OCT2011 (top) and 18MAR2013 (bottom). For 25OCT2011 the total number of important (e.g effective) variables is 23. Ten of them are related to the dynamical NWP predictors, namely: mean and maximum hourly rainfall intensity and mean and maximum soil moisture in the four layers 0–10, 10–40, 40–100 and 100–200 cm below ground. Four of the variables are geomorphology-related predictors, namely: elevation, topography factor, slope, concavity/convexity of the land parallel to aspect. Also geology-related predictors are four, namely: geological continuum, soil permeability, distance from main tectonic features and landslides main scarps. Three of the variables are climate-related predictors, namely: the precipitation amounts with a returning period of 100 years and occurring in 12, 48 and 96 h. Finally, two are hydrology-related predictors, namely: time of concentration and Topography Wetness Index. All these layers are highlighted in Table 4 with the * symbol. For 18MAR2013, nine predictors are classified as important by the RF algorithm. Two are dynamical NWP predictors, namely mean and maximum soil moisture in the layer 0–10 cm below ground. Two are geomorphological predictors (e.g. elevation and the altitude above channel network), one is the climatological precipitation amounts of a rainfall event occurring in 24 h and having a returning period of 100 year, one is the time of concentration (hydrology-related predictor), one is the euclidean distance from the main tectonic features and the last two are the vegetation index and the land cover. All these layers are highlighted in Table 4 with the • symbol.

4 Discussions

We developed a statistical framework for the modeling and prediction of shallow landslides triggered by heavy precipitations. We considered as input predictors both static thematic layers such as geomorphology, hydrology, geology and climate related predisposing factors and dynamical information provided by NWP short term forecast, namely precipitation amounts and soil water content. We combined these predictors together by means of two different statistical models: the well known and widely used Generalized Linear Model and Breiman's Random Forest. In the procedure implemented, the two models have no interactions between each other and are tested separately (see Fig. 3 for a schematic flowchart of the modeling). We chose the GLM model because it has be proven (see references cited in Sect. 2.4) to provide reliable results in landslides susceptibility assessment, while the RF algorithm has been recently and successfully applied in landslide research and applications (Stumpf and Kerle, 2011b; Catani et al., 2013). We tested the whole procedure for two study cases occurred in Tuscany region (central Italy) in the recent past. The two events were characterized by heavy precipitations that induced several shallow landslide and debris flows as reported by after-event field surveys and inventory maps (see Fig. 1). The statistical modeling produced landslide hazard maps, i.e. a partition of the territory in different degrees of landslide susceptibility ranging from stable slopes to high unstable slopes. More specifically, the study is aimed at testing and evaluating the information content provided by numerical weather predictions in landslide assessment. This aim is justified by the fact that recently NWP data are arising as a promising and reliable source of information for real-time forecasting chain of rainfall induced shallow landslides (Schmidt et al., 2008; Segoni et al., 2009; Mercogliano et al., 2013). The use of NWP data is justified in the setting up of landslides warning systems over large areas (i.e. basin scale) and when both a spatial and a temporal forecasting of shallow landslide occurrence is desirable. In the above cited bibliography, NWP data were basically considered only for what concerns precipitation amounts. Schmidt et al. (2008) used a regional model assimilating NWP data to feed a physically-based model to simulate basin hydrology on the basis of rainfall forecast forcing. Segoni et al. (2009) combined both rainfall fields observed from meteorological ground-based radar and high resolution rainfall forecast, statistically downscaled, to feed hydro-geological models to yield a factor of safety for the area of interest. Mercogliano et al. (2013) presented a similar forecasting chain composed by a downscaling procedure of NWP rainfall data to feed a geo-technical model to gain a factor of safety on a pixel-by-pixel basis.

Our study differs from the above mentioned because: (i) we consider not only NWP rainfall amounts but also NWP hourly rainfall intensities and NWP soil moisture in four layers below ground, (ii) we do not perform any statistical downscaling of the NWP data towards the mapping unit and (iii) we combine both static predictors and NWP data into a "black-box" statistical model.

As described in Sects. 2.1.1 and 2.1.2, the study cases selected show different features: 25OCT2011 was characterized by deep atmospheric convection and heavy precipitations cumulated over a short time interval and in a, relatively, small area (hundred of km^2), while 18MAR2013 was characterized by a weather storm with precipitation spread over a 24 h period and affecting a larger area (thousand of km^2).

Considering the area under the Receiver Operating Characteristics curve (AUC area) as a good representative index of the accuracy of landslide probabilistic forecast (Frattini et al., 2010), results here achieved are similar to those find in recent bibliography. For the 25OCT2011 study case, the AUC values obtained are quite high both for the GLM model (AUC = 0.909) and, even better, for the RF classifier (AUC = 0.968). For the 18MAR2013 case AUC values are lower: 0.833 for the GLM model and 0.761 for the RF classifier. Rossi et al. (2010) in an area with an extension about 100 km^2 applied four different statistical methods for landslide susceptibility zonation in central Italy. In the validation set (i.e. observed landslides were divided into a training set to tune the models and into a validation set to validate the models themselves) the authors obtained AUC values around 0.74, whereas in the training set AUC values were higher (from 0.84 for the linear discriminant model to 0.99 for the neural network model). In this latter study the mapping unit was "slope unit" which is not comparable to that one adopted in the present study. Frattini et al. (2010) applied five debris-flow susceptibility models (both statistical and physically based) in an area extending for about 300 km^2 in the north of Italy and with a grid cells resolution of 10 m. They obtained AUC values ranging from 0.64 for a physically based model to 0.84 for a discriminant model. For what concerns studies regarding small areas of interest (below basin scale with an extension around 10 km^2), Mondini et al. (2011) in their study regarding a semi-automatic mapping of rainfall induced shallow landslides using optical satellite images, obtained AUC values around 0.87 in the training area and around 0.82 in the validation area. They used four different statistical methods (linear discriminant analysis, quadratic discriminant analysis, logistic regression and a combined regression model) with a quite high mapping resolution (0.6 m) and their study area was located in south Italy (Sicily). More recently, Catani et al. (2013) in their study investigating the landslide susceptibility in Tuscany by using the RF technique, found AUC values ranging from 0.74 to 0.97 when increasing the number of samples required to calibrate a model. The resolution of their study was 50 m, which is comparable to that one here adopted (30 m).

In the susceptivity maps produced (see Fig. 8a and b in particular), the pixelated shape of some areas is due to the fact that no downscaling is performed to the NWP outputs (produced at 3 km of horizontal resolution) towards the resolution of the static thematic layers (30 m of horizontal resolution). Nevertheless this coarse approximation is not affecting dramatically the results as demonstrated by the AUC values achieved.

Summarizing, from the comparison with recent state-of-art studies we can conclude that our findings are encouraging and justify the statistical models adopted for the area under exam.

However, as stated in the introduction, the aim of this work was also to establish the, possible, positive impact of NWP predictions on landslide susceptibility assessment. Following the variable importance provided by the RF technique, from Fig. 10 we state that NWP information is crucial for the 25OCT2011 event and marginal, even if not null, for the 18MAR2013 event. In fact for 25OCT2011 10 out of 23 relevant predisposing factors were derived from NWP contents, see Table 5 for a list of such predictors in decreasing order of importance. Not surprisingly the total amount of precipitation over the rainfall event is not occurring in the list, whereas the mean and the maximum hourly rainfall intensity are ranked in the top list, since, as described in Sect. 2.1.1 and in references cited, the 25OCT2011 was characterized by deep convection activity with high precipitation rates. For what concerns the 18MAR2013, the impact of NWP factors are marginal. The soil moisture content in the first 10 cm below ground is the only variable occurring in the importance ranking (see Fig. 10 in the bottom). This result confirms the findings of recent studies investigating the strong linkage between soil moisture and landslide occurrence (Ray and Jacobs, 2007; Ray et al., 2010). In particular Ponziani et al. (2012) in analyzing the role of antecedent soil moisture conditions at regional scale, assessed the soil moisture content to be as important as rainfall intensity for the triggering of landslides. Our study presents the opportunity to integrate NWP soil moisture forecast information content into a statistical method to take account of the relationship between rainfall, soil moisture and landslide movement. The rainfall forecast here produced are validated in Sect. 3.1, for what concerns the ability of a numerical regional model to estimate soil moisture content, the reader is referred to Schneider et al. (2014).

5 Final considerations and future developments

As stated in the previous section, the statistical framework developed and the predisposing factors considered were able to model the shallow landslide occurrence triggered by heavy precipitations at basin scale (at least for the two study cases considered). Results obtained are encouraging and are similar to those find in recent studies. We assessed the relative importance of NWP content information (not downscaled statistically) and we concluded that the benefits might be relevant in particular in rainfall events characterized by high precipitation rates (25OCT2011 in this study). We therefore tried to bridge the gap between the micro-γ scale (≤ 20 m), which is the characteristic scale of landslide occurring at basin scale, with the meso-γ, which is the typical scale of the NWP forecast. Results here shown, demonstrate that this gap could be filled also thanks to the help of black-box statistical modelling. In spite of the simplicity of such statistical approach, the drawback of the method proposed is that, being data-driven, a model built up for one region and for one particular event, cannot be applied, without any re-calibration, to a neighboring area or for a similar rainfall event.

Moreover, it has to be kept in mind that these results were achieved without any algorithm or model calibrations. A possible tuning of the whole forecasting system may rely on the improvement of NWP performances. For example, since antecedent soil condition is one of the crucial factor for determining landslide-triggering rainfall thresholds (Glade et al., 2000), the NWP's initial soil moisture could be better estimated by means of the assimilation of remote sensed data at regional scale (Schneider et al., 2014). On the other hand, nowadays, global models have sophisticated assimilation algorithms to ingest observed soil water content estimates and to produce "warm" analysis of soil conditions (Dharssi et al., 2011; de Rosnay et al., 2013). As pointed out by Segoni et al. (2009), in order to gain a more accurate temporal and spatial knowledge of the triggering rainfall, EPS (Ensemble Prediction Systems) or RUC (Rapid Updated Cycle) numerical forecasting chains could be adopted. This latter methodology could be further improved by means of the assimilation of radar data (Segoni et al., 2009) to keep antecedent soil moisture conditions as close to reality as possible.

Before the setting-up of the forecasting chain here developed into a fully operative warning system, further well-documented severe rainfall events need to be addressed and the results have to be validated with ground observations.

Acknowledgements. The authors wish to thank the personnel of the Civil Protection and Genio Civile offices involved in the collection of the data for preparing the landslide inventory maps for the study cases 25OCT2011 and 18MAR2013.

The authors are also grateful to the institutes involved in the development of the WRF model for the user's support provided.

This work has been partially financed by the Tuscany Regional Administration and its support is gratefully acknowledged.

References

Argnani, A., Barbacini, G., Bernini, M., Camurri, F., Ghielmi, M., Papani, G., Rizzini, F., Rogledi, S., and Torelli, L.: Gravity tectonics driven by Quaternary uplift in the Northern Apennines: insights from the La Spezia-Reggio Emilia geo-transect, Quatern. Int., 101, 13–26, 2003. 4992

Atkinson, P. and Massari, R.: Generalised linear modelling of susceptibility to landsliding in the Central Apennines, Italy, Comput. Geosci., 24, 373–385, 1998. 4999

Avanzi, G. D., Galanti, Y., Giannecchini, R., and Puccinelli, A.: Fragility of territory and infrastructures resulting from rainstorms in Northern Tuscany (Italy), in: Landslide Science and Practice, Springer, 239–246, 2013. 4988, 4991, 4993

Baldi, M., Ciardini, V., Dalu, J. D., De Filippis, T., Maracchi, G., and Dalu, G.: Hail occurrence in Italy: towards a national database and climatology, Atmos. Res., 138, 268–277, 2014. 4999

Borghuis, A., Chang, K., and Lee, H.: Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery, Int. J. Remote Sens., 28, 1843–1856, 2007. 5001

Bossard, M., Feranec, J., and Otahel, J.: CORINE land cover technical guide: Addendum 2000, European Environment Agency, Copenhagen, Denmark, 2000. 5000

Breiman, L.: Bagging predictors, Mach. Learn., 24, 123–140, 1996. 5003

Breiman, L.: Random forests, Mach. Learn., 45, 5–32, 2001. 4989, 5003, 5007

Breiman, L.: Manual on setting up, using, and understanding random forests v3. 1, Statistics Department University of California Berkeley, CA, USA, http://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf (last access: 2 August 2014), 2002. 5008

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A.: Classification and regression trees, Wadsworth and Brooks, Monterey, California, USA, 1984. 5003

Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, Nat. Hazards Earth Syst. Sci., 5, 853–862, doi:10.5194/nhess-5-853-2005, 2005. 5003, 5004

Bryan, G. H. and Morrison, H.: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics, Mon. Weather Rev., 140, 202–225, 2012. 5001

Buzzi, A., Davolio, S., D'Isidoro, M., and Malguzzi, P.: The impact of resolution and of MAP reanalysis on the simulations of heavy precipitation during MAP cases, Meteorol. Z., 13, 91–97, 2004. 5006

Buzzi, A., Davolio, S., Malguzzi, P., Drofa, O., and Mastrangelo, D.: Heavy rainfall episodes over Liguria in autumn 2011: numerical forecasting experiments, Nat. Hazards Earth Syst. Sci., 14, 1325–1340, doi:10.5194/nhess-14-1325-2014, 2014. 4988, 4992, 5006

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., and Reichenbach, P.: GIS techniques and statistical models in evaluating landslide hazard, Earth Surf. Proc. Land., 16, 427–445, 1991. 4999, 5002

Catani, F., Casagli, N., Ermini, L., Righini, G., and Menduni, G.: Landslide hazard and risk mapping at catchment scale in the Arno River basin, Landslides, 2, 329–342, 2005. 4998

Catani, F., Lagomarsino, D., Segoni, S., and Tofani, V.: Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues, Nat. Hazards Earth Syst. Sci., 13, 2815–2831, doi:10.5194/nhess-13-2815-2013, 2013. 4994, 4995, 4996, 4997, 4999, 5000, 5004, 5009, 5011

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., Duan, Q. Y., Ek, M., and Betts, A.: Modeling of land surface evaporation by four schemes and comparison with FIFE observations, J. Geophys. Res., 101, 7251–7268, 1996. 5002, 5022

Chen, S. and Dudhia, J.: Annual report: WRF physics, Air Force Weather Agency, 2000. 5001

Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, J. D., and Chacón, J.: Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: application to the river Beiro basin (Spain), Nat. Hazards Earth Syst. Sci., 12, 327–340, doi:10.5194/nhess-12-327-2012, 2012. 4998, 5000

Crozier, M. J.: Prediction of rainfall-triggered landslides: A test of the antecedent water status model, Earth Surf. Proc. Land., 24, 825–833, 1999. 5005

Cruden, D. and Hu, X.: Landslides in the Rocky Mountains of Canada, in: Geomorphological Hazards in High Mountain Areas, 133–148, Springer, 1998. 4999

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J.: Random forests for classification in ecology, Ecology, 88, 2783–2792, 2007. 5004

Demir, G., Aytekin, M., Akgün, A., İkizler, S. B., and Tatar, O.: A comparison of landslide susceptibility mapping of the eastern part of the North Anatolian Fault Zone (Turkey) by likelihood-frequency ratio and analytic hierarchy process methods, Nat. Hazards, 65, 1481–1506, 2013. 4997

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, L.: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF, Q. J. Roy. Meteorol. Soc., 139, 1199–1213, 2013. 5013

Devkota, K. C., Regmi, A. D., Pourghasemi, H. R., Yoshida, K., Pradhan, B., Ryu, I. C., Dhital, M. R., and Althuwaynee, O. F.: Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya, Nat. Hazards, 65, 135–165, 2013. 4997

Dharssi, I., Bovis, K. J., Macpherson, B., and Jones, C. P.: Operational assimilation of ASCAT surface soil wetness at the Met Office, Hydrol. Earth Syst. Sci., 15, 2729–2746, doi:10.5194/hess-15-2729-2011, 2011. 5013

Díaz-Uriarte, R. and De Andres, S. A.: Gene selection and classification of microarray data using random forest, BMC Bioinformatics, 7, 1, 2006. 5004

Di Naccio, D., Boncio, P., Brozzetti, F., Pazzaglia, F. J., and Lavecchia, G.: Morphotectonic analysis of the Lunigiana and Garfagnana grabens (northern Apennines, Italy): Implications for active normal faulting, Geomorphology, 201, 293–311, 2013. 4992, 4993

D'Odorico, P. and Fagherazzi, S.: A probabilistic model of rainfall-triggered shallow landslides in hollows: A long-term analysis, Water Resour. Res., 39, 1262, doi:10.1029/2002WR001595, 2003. 4996

Donati, L. and Turrini, M.: An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy), Eng. Geol., 63, 277–289, 2002. 4999

Felicísimo, Á. M., Cuartero, A., Remondo, J., and Quirós, E.: Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study, Landslides, 10, 175–189, 2013. 4995

Fiori, E., Comellas, A., Molini, L., Rebora, N., Siccardi, F., Gochis, D., Tanelli, S., and Parodi, A.: Analysis and hindcast simulations of an extreme rainfall event in the Mediterranean area: The Genoa 2011 case, Atmos. Res., 138, 13–29, 2014. 4988

Floris, M., Iafelice, M., Squarzoni, C., Zorzi, L., De Agostini, A., and Genevois, R.: Using online databases for landslide susceptibility assessment: an example from the Veneto Region (northeastern Italy), Nat. Hazards Earth Syst. Sci., 11, 1915–1925, doi:10.5194/nhess-11-1915-2011, 2011. 4994, 4997

Frattini, P., Crosta, G., and Carrara, A.: Techniques for evaluating the performance of landslide susceptibility models, Eng. Geol., 111, 62–72, 2010. 4990, 5010

NHESSD

2, 4987–5036, 2014

Statistical modeling of shallow landslides using static predictors and NWP outputs

V. Capecchi et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀| |▶

◀ ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Gao, X., Huete, A. R., Ni, W., and Miura, T.: Optical–biophysical relationships of vegetation spectra without background contamination, Remote Sens. Environ., 74, 609–620, 2000. 5000

Ghimire, B., Rogan, J., and Miller, J.: Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic, Remote Sens. Lett., 1, 45–54, 2010. 5004

Giannecchini, R.: Relationship between rainfall and shallow landslides in the southern Apuan Alps (Italy), Nat. Hazards Earth Syst. Sci., 6, 357–364, doi:10.5194/nhess-6-357-2006, 2006. 4991

Gislason, P. O., Benediktsson, J. A., and Sveinsson, J. R.: Random forests for land cover classification, Pattern Recogn. Lett., 27, 294–300, 2006. 5004

Glade, T., Crozier, M., and Smith, P.: Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "Antecedent Daily Rainfall Model", Pure Appl. Geophys., 157, 1059–1079, 2000. 5013

Guzzetti, F., Carrara, A., Cardinali, M., and Reichenbach, P.: Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy, Geomorphology, 31, 181–216, 1999. 4989, 4995, 4998, 4999, 5002

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., and Ardizzone, F.: Probabilistic landslide hazard assessment at the basin scale, Geomorphology, 72, 272–299, 2005. 4989

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., and Chang, K.-T.: Landslide inventory maps: new tools for an old problem, Earth-Sci. Rev., 112, 42–66, 2012. 5001

Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev., 134, 2318–2341, 2006. 5022

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sens. Environ., 83, 195–213, 2002. 5000

Kain, J. S., Weiss, S. J., Bright, D. R., Baldwin, M. E., Levit, J. J., Carbin, G. W., Schwartz, C. S., Weisman, M. L., Droegemeier, K. K., Weber, D. B., and Thomas, K. W.: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP, Weather Forecast., 23, 931–952, 2008. 5001

Klemp, J., Skamarock, W., and Dudhia, J.: Conservative split-explicit time integration methods for the compressible nonhydrostatic equations, Mon. Weather Rev., 135, 2897–2913, 2007. 5001

Liaw, A. and Wiener, M.: Classification and regression by randomForest, R News, 2, 18–22, 2002. 5005

Mancini, F., Ceppi, C., and Ritrovato, G.: GIS and statistical analysis for landslide susceptibility mapping in the Daunia area, Italy, Nat. Hazards Earth Syst. Sci., 10, 1851–1864, doi:10.5194/nhess-10-1851-2010, 2010. 4997

Mărgărint, M. C., Grozavu, A., and Patriche, C. V.: Assessing the spatial variability of weights of landslide causal factors in different regions from Romania using logistic regression, Nat. Hazards Earth Syst. Sci. Discuss., 1, 1749–1774, doi:10.5194/nhessd-1-1749-2013, 2013. 4995

Marjanovic, M., Kovacevic, M., Bajat, B., Mihalic, S., and Abolmasov, B.: Landslide assessment of the Starca basin (Croatia) using machine learning algorithms, Acta Geotech. Slov., 8, 45–55, 2011. 4995

McCullagh, P.: Generalized linear models, Eur. J. Oper. Res., 16, 285–292, 1984. 4989, 5003

McCullagh, P. and Nelder, J.: General linear models, Chapman and Halls, London, 1989. 4989, 5003

Mercogliano, P., Segoni, S., Rossi, G., Sikorsky, B., Tofani, V., Schiano, P., Catani, F., and Casagli, N.: Brief communication "A prototype forecasting chain for rainfall induced shallow landslides", Nat. Hazards Earth Syst. Sci., 13, 771–777, doi:10.5194/nhess-13-771-2013, 2013. 4989, 4990, 5009

Mondini, A., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., and Ardizzone, F.: Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images, Remote Sens. Environ., 115, 1743–1757, 2011. 5001, 5011

Moore, I. D. and Wilson, J. P.: Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation, J. Soil Water Conserv., 47, 423–428, 1992. 4995

Moriondo, M., Stefanini, F., and Bindi, M.: Reproduction of olive tree habitat suitability for global change impact assessment, Ecol. Model., 218, 95–109, 2008. 5004

Nefeslioglu, H. A., Duman, T. Y., and Durmaz, S.: Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey), Geomorphology, 94, 401–418, 2008. 4997

Nefeslioglu, H. A., Gokceoglu, C., Sonmez, H., and Gorum, T.: Medium-scale hazard mapping for shallow landslide initiation: the Buyukkoy catchment area (Cayeli, Rize, Turkey), Landslides, 8, 459–483, 2011. 4996

NHESSD
2, 4987–5036, 2014

Statistical modeling of shallow landslides using static predictors and NWP outputs

V. Capecchi et al.

Nefeslioglu, H. A., San, B., Gokceoglu, C., and Duman, T.: An assessment on the use of Terra ASTER L3A data in landslide susceptibility mapping, Int. J. Appl. Earth Obs., 14, 40–60, 2012. 4997
Orlanski, I.: A rational subdivision of scales for atmospheric processes, B. Am. Meteorol. Soc., 56, 527–530, 1975. 4989
Pal, M.: Random forest classifier for remote sensing classification, Int. J. Remote Sens., 26, 217–222, 2005. 5004
Parodi, A., Boni, G., Ferraris, L., Siccardi, F., Pagliara, P., Trovatore, E., Foufoula-Georgiou, E., and Kranzlmueller, D.: The "perfect storm": From across the Atlantic to the hills of Genoa, Eos Trans. Am. Geophys. Union, 93, 225–226, 2012. 4988
Peters, J., Baets, B. D., Verhoest, N. E., Samson, R., Degroeve, S., Becker, P. D., and Huybrechts, W.: Random forests as tool for ecohydrological distribution modelling, Ecol. Model., 207, 304–318, 2007. 5004
Ponziani, F., Pandolfo, C., Stelluti, M., Berni, N., Brocca, L., and Moramarco, T.: Assessment of rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in the Umbria region (central Italy), Landslides, 9, 229–237, 2012. 5012
Pradhan, B., Chaudhari, A., Adinarayana, J., and Buchroithner, M. F.: Soil erosion assessment and its correlation with landslide events using remote sensing data and GIS: a case study at Penang Island, Malaysia, Environ. Monit. Assess., 184, 715–727, 2012. 4995
Ray, R. L. and Jacobs, J. M.: Relationships among remotely sensed soil moisture, precipitation and landslide events, Nat. Hazards, 43, 211–222, 2007. 5012
Ray, R. L., Jacobs, J. M., and Cosh, M. H.: Landslide susceptibility mapping using downscaled AMSR-E soil moisture: a case study from Cleveland Corral, California, US, Remote Sens. Environ., 114, 2624–2636, 2010. 5012
Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, E., Pignone, F., Siccardi, F., Silvestro, F., Tanelli, S., and Parodi, A.: Extreme Rainfall in the Mediterranean: what can we learn from observations?, J. Hydrometeorol., 14, 906–922, 2013. 4988, 4992
Regione Toscana, S. I. R.: Report eventi meteo-idrologici occorsi nel periodo 5–19 Marzo 2013 sul territorio regionale, Tech. rep., Centro Funzionale della Regione Toscana, Pisa, Italy, 2013. 4993
Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., and Peruccacci, S.: Optimal landslide susceptibility zonation based on multiple forecasts, Geomorphology, 114, 129–142, 2010. 5010

Sacchi, J. A.: Analisi meteorologica degli eventi alluvionali eccezionali tra l'Alta Toscana e l'estremo Levante ligure dal 2009 al 2011, Atti della Societá Toscana di Scienze Naturali, Memorie serie A, 117–119, 75–88, 2012. 4988

Schicker, R. and Moon, V.: Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale, Geomorphology, 161, 40–57, 2012. 4999

Schmidt, J., Turek, G., Clark, M. P., Uddstrom, M., and Dymond, J. R.: Probabilistic forecasting of shallow, rainfall-triggered landslides using real-time numerical weather predictions, Nat. Hazards Earth Syst. Sci., 8, 349–357, doi:10.5194/nhess-8-349-2008, 2008. 4989, 4990, 5009

Schneider, S., Wang, Y., Wagner, W., and Mahfouf, J.-F.: Impact of ASCAT soil moisture assimilation on regional precipitation forecasts: a case study for Austria, Mon. Weather Rev., 142, 1525–1541, 2014. 5012, 5013

Segoni, S., Leoni, L., Benedetti, A. I., Catani, F., Righini, G., Falorni, G., Gabellani, S., Rudari, R., Silvestro, F., and Rebora, N.: Towards a definition of a real-time forecasting network for rainfall induced shallow landslides, Nat. Hazards Earth Syst. Sci., 9, 2119–2133, doi:10.5194/nhess-9-2119-2009, 2009. 4989, 4990, 5009, 5013

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227, 3465–3485, 2008. 4989, 5001

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 2, Tech. rep., DTIC Document, Mesoscale and Microscale Meteorology Division, NCAR, Boulder, Colorado, USA, 2005. 4989, 5001

Stumpf, A. and Kerle, N.: Combining Random Forests and object-oriented analysis for landslide mapping from very high resolution imagery, Proced. Environ. Sci., 3, 123–129, 2011a. 5004

Stumpf, A. and Kerle, N.: Object-oriented mapping of landslides using random forests, Remote Sens. Environ., 115, 2564–2577, 2011b 5004, 5009

Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme, Part II: Implementation of a new snow parameterization, Mon. Weather Rev., 136, 5095–5115, 2008. 5022

Varnes, D. J., the International Association Eng. Geol. Commission on Landslides, and on Slopes, O. M. M.: Landslide hazard zonation: a review of principles and practice, United Nations Educational, Paris, France, 1984. 4999, 5000

Vorpahl, P., Elsenbeer, H., Märker, M., and Schröder, B.: How can statistical models help to determine driving factors of landslides?, Ecol. Model., 239, 27–39, 2012. 5004

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, vol. 100, Academic Press, Oxford, UK, 2011. 5005

Yilmaz, I.: Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine, Environ. Earth Sci., 61, 821–836, 2010. 4997

Table 1. Basic options of the numerical WRF simulations.

Variable	Value
projection	Lambert
rows × columns	440 × 400
vertical levels	35
horizontal resolution	3 km
time step	18 s
cumulus convection	explicit (no parameterization)
micro-physics option	Thompson (Thompson et al., 2008)
boundary-layer option	Yonsei University (Hong et al., 2006)
land-surface option	Unified Noah model (Chen et al., 1996)

NHESSD
2, 4987–5036, 2014

Table 2. Statistics of the observed and modelled 24 h rainfall amounts for both study cases.

25OCT2011	Min	1st Qu	Median	Mean	3rd Qu	Max
Observed data	21.4	68.6	105.0	153.3	176.8	374.8
Modelled data	40.9	50.8	61.1	63.5	69.3	112.1
Modelled data (whole area)	18.2	43.8	59.3	70.4	90.7	218.9
18MAR2013	**Min**	**1st Qu**	**Median**	**Mean**	**3rd Qu**	**Max**
Observed data	0.0	49.9	64.6	73.0	85.0	284.8
Modelled data	13.6	21.0	38.0	43.7	59.6	127.6

NHESSD
2, 4987–5036, 2014

Statistical modeling of shallow landslides using static predictors and NWP outputs

V. Capecchi et al.

Table 3. Statistics of the observed and modelled 1 h rainfall amounts for both study cases.

25OCT2011	Min	1st Qu	Median	Mean	3rd Qu	Max
Observed data	6.4	26.0	28.6	31.7	33.2	67.2
Modelled data	10.7	16.4	18.9	21.7	26.5	41.8
18MAR2013	**Min**	**1st Qu**	**Median**	**Mean**	**3rd Qu**	**Max**
Observed data	0.2	6.6	8.7	9.3	10.5	30.6
Modelled data	2.1	3.9	6.2	6.9	8.1	20.7

NHESSD
2, 4987–5036, 2014

Table 4. List of the predictors taken into account in both GLM model and RF classifier for landslide hazard mapping. Using the RF's variable importance parameter, predictors relevant for the 25OCT2011 rainfall event have a ⋆ symbol, whereas those relevant for the 18MAR2013 event have a • symbol.

Variable/Code	Short description	Unit of measure
Geomorphology-related predictors		
dem_30m_topo/DEM ⋆•	Digital elevation model	m
h_channel_geo/AaCN•	Altitude above channel network	m
aspect_topo/ASP	Aspect	–
slope_topo/SLP⋆	Slope	–
ls_factor_geo/LSF⋆	Topography factor from RUSLE	–
plan_curv_geo/PLAC	concavity/convexity of the land perpendicular to aspect	$h\,m^{-1}$
prof_curv_geo/PRFC⋆	concavity/convexity of the land parallel to aspect	$h\,m^{-1}$
Hydrology-related predictors		
conv_index_geo/COVI	convergence/divergence to overland flow	–
tc_geo/ToC•	Time of concentration	h
TWI_geo/TWI⋆	Topographic Wetness Index	–
d_aste_fluvi_geo/DfCN	Euclidean distance from river network	m
Geology-related predictors		
d_lineamenti_tettonici_geo/DfTF⋆•	Euclidean distance from main tectonic features	m
litho_geo_int/BLT ⋆	Geological continuum of Tuscany Region	categorical/16 classes
a_distacco_geo/LMS ⋆	Landslides main scarps	boolean
pedopaesaggi_int/SKST ⋆	Soil permeability	categorical/7 classes
quaternario_frane_geo/LaSD	Landslides and superficial deposits	categorical/2 classes
assetto_geo/SSS	Slope Structural Setting	categorical/7 classes
Climate-related predictors		
p12h_100_clima/R12 ⋆	Rainfall 12 h duration and 100 years return period	mm
p24h_100_clima/R24 •	Rainfall 24 h duration and 100 years return period	mm
p48h_100_clima/R48 ⋆	Rainfall 48 h duration and 100 years return period	mm
p96h_100_clima/R96 ⋆	Rainfall 96 h duration and 100 years return period	mm
Other predictors		
corine_c_landscape/COR •	Corine Land Use	categorical
evi_media_land/EVI •	Vegetation Index	–
NWP predictors		
r_arw3km_SOILW0_10cm⋆•	Soil moisture 0–10 cm below ground layer	$m^3\,m^{-3}$
r_arw3km_SOILW10_40cm⋆	Soil moisture 10–40 cm below ground layer	$m^3\,m^{-3}$
r_arw3km_SOILW40_100cm⋆	Soil moisture 40–100 cm below ground layer	$m^3\,m^{-3}$
r_arw3km_SOILW100_200cm⋆	Soil moisture 100–200 cm below ground layer	$m^3\,m^{-3}$
r_arw3km_APCP	Precipitation amounts	$mm\,(24\,h)^{-1}$
r_arw3km_APCP.RI⋆	Precipitation intensity	$mm\,h^{-1}$

NHESSD

2, 4987–5036, 2014

Statistical modeling of shallow landslides using static predictors and NWP outputs

V. Capecchi et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Table 5. List of the ten NWP input predictors classified as "important" by the RF classifier for the 25OCT2011 event.

Coded name	Description	Importance (ranking)
r_arw3km_APCP.RI.mean	mean hourly rainfall intensity	1
r_arw3km_SOILW100_200cm.max	maximum soil moisture (100–200 cm below ground)	3
r_arw3km_SOILW40_100cm.max	maximum soil moisture (40–100 cm below ground)	6
r_arw3km_APCP.RI.max	max hourly rainfall intensity	11
r_arw3km_SOILW0_10cm.max	maximum soil moisture (0–10 cm below ground)	15
r_arw3km_SOILW10_40cm.max	maximum soil moisture (10–40 cm below ground)	16
r_arw3km_SOILW40_100cm.mean	mean soil moisture (40–100 cm below ground)	17
r_arw3km_SOILW10_40cm.mean	mean soil moisture (10–40 cm below ground)	19
r_arw3km_SOILW0_10cm.mean	mean soil moisture (0–10 cm below ground)	20
r_arw3km_SOILW100_200cm.mean	mean soil moisture (100–200 cm below ground)	22

NHESSD
2, 4987–5036, 2014

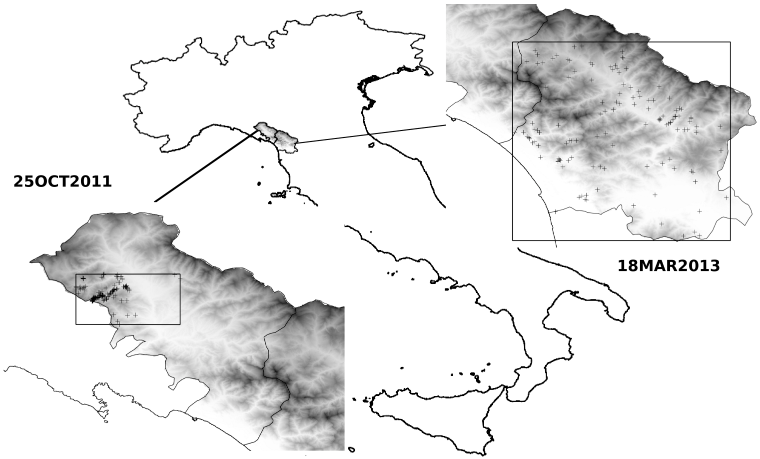

Figure 1. Italy domain with the two areas of interest: in the inset figure on the left side, it is shown the Lunigiana area belonging to the administrative province of Massa-Carrara. This area was interested by the event 25OCT2011. In the inset figure on the right, it is shown the Garfagnana area belonging to the administrative province of Lucca. This area was interested by the event 18MAR2013. In both inset figures it is depicted the extent (rectangular shaped area) where the statistical models were implemented and tested, while the cross signs represent the observed landslides.

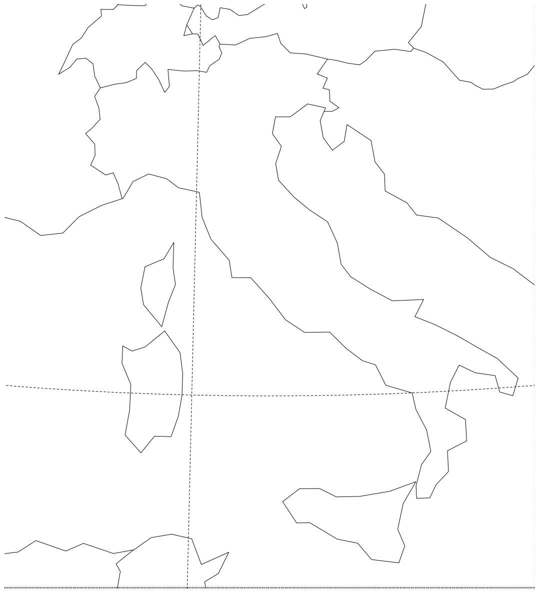

Figure 2. Area of the WRF simulations. The horizontal spatial resolution of the model's run is 3 km.

NHESSD
2, 4987–5036, 2014

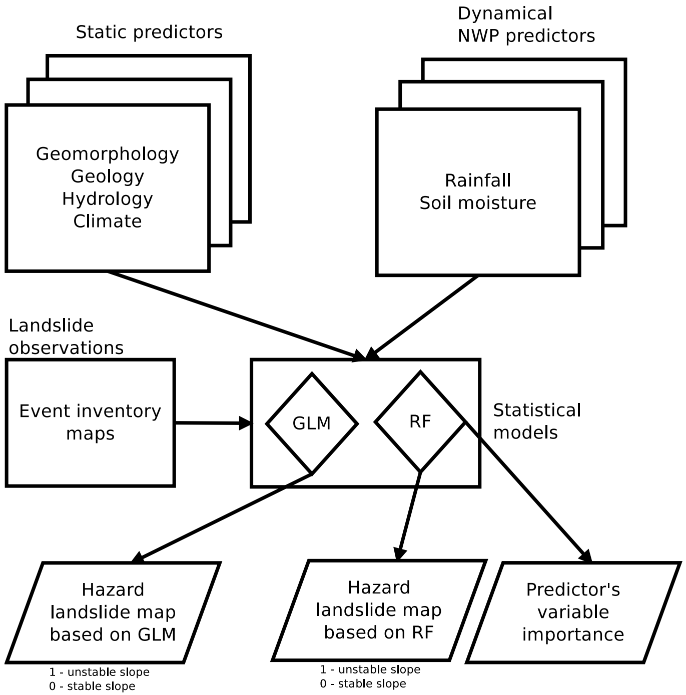

Figure 3. Flowchart of the statistical modeling of shallow landslide triggered by rainfall. Input data are indicated with the □ symbol, statistical models are indicated with the ◊ symbol and output data are indicated with the parallelogram symbol.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

NHESSD

2, 4987–5036, 2014

Statistical modeling of shallow landslides using static predictors and NWP outputs

V. Capecchi et al.

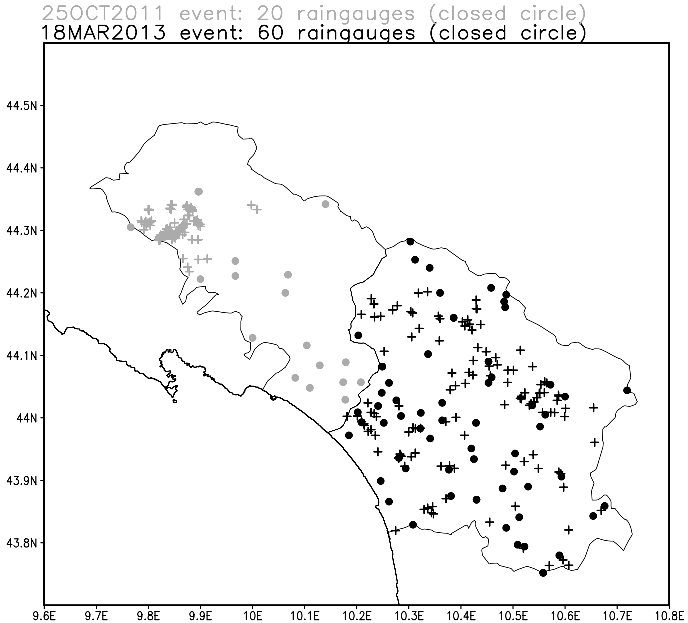

Figure 4. Locations of the rain-gauges (closed circles) for the two study cases: 25OCT2011 in gray, 18MAR2013 in black. Crosses indicate the location of the landslides for the two study cases: 25OCT2011 in gray, 18MAR2013 in black.

NHESSD
2, 4987–5036, 2014

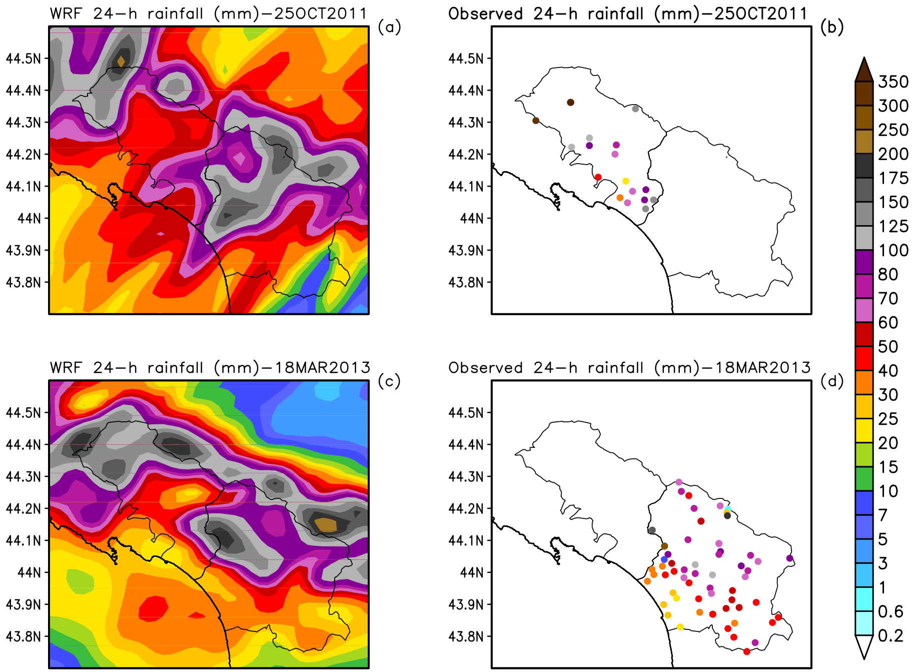

Figure 5. (a) and **(b)**: modelled and observed precipitation in mm accumulated in the 24 h period starting at 00:00 UTC of 25 October 2011. **(c)** and **(d)**: modelled and observed precipitation in mm accumulated in the 24 h period starting at 00:00 UTC of 18 March 2013.

NHESSD
2, 4987–5036, 2014

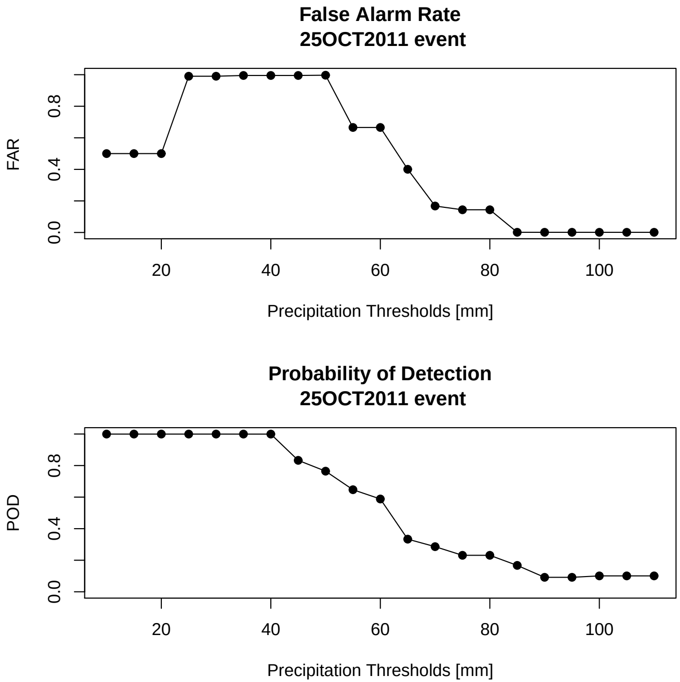

Figure 6. Study case 25OCT2011: False Alarm Rate (top panel) and Probability of Detection (bottom panel) skills obtained when validating the accumulated 24 h rainfall predicted by the WRF model with the rainfall data observed at rain-gauge locations displayed in Fig. 4.

NHESSD
2, 4987–5036, 2014

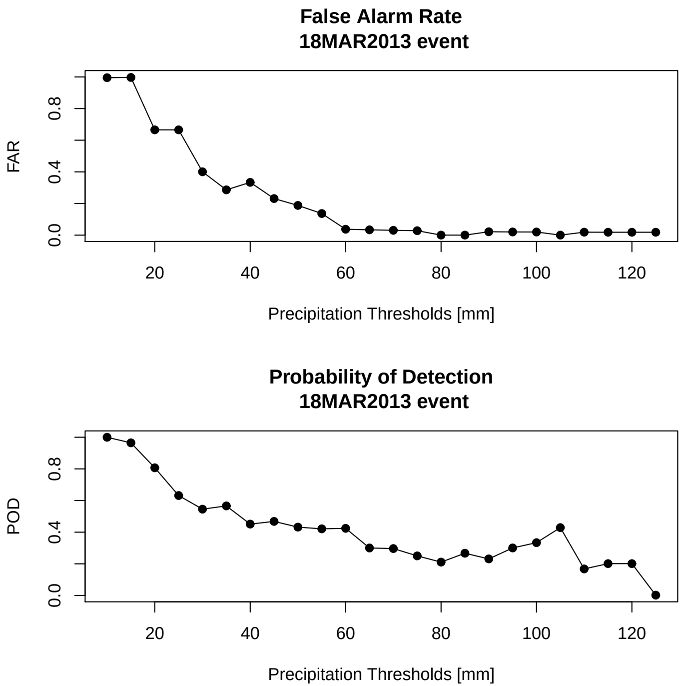

Figure 7. Study case 18MAR2013: False Alarm Rate (top panel) and Probability of Detection (bottom panel) skills obtained when validating the accumulated 24 h rainfall predicted by the WRF model with the rainfall data observed at rain-gauge locations displayed in Fig. 4.

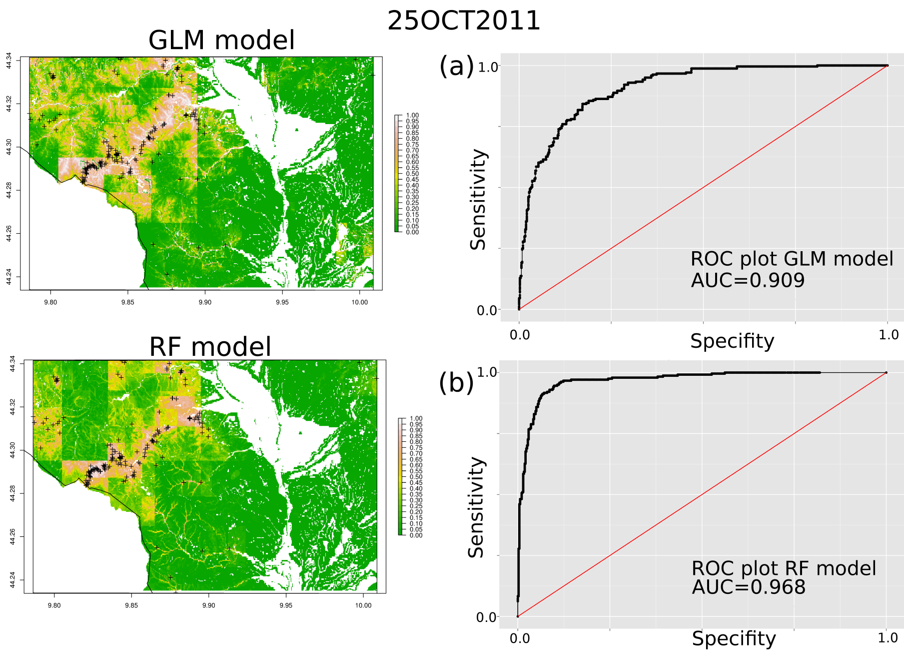

Figure 8. 25OCT2011 landslide hazard maps and corresponding Receiver Operating Characteristics Curves (ROC) with the values of the underlying area (AUC). Results from the GLM model **(a)** and from the RF classifier **(b)**. Crosses points are the event inventory maps produced by field surveys a few days after the rainfall event.

NHESSD
2, 4987–5036, 2014

Statistical modeling of shallow landslides using static predictors and NWP outputs

V. Capecchi et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

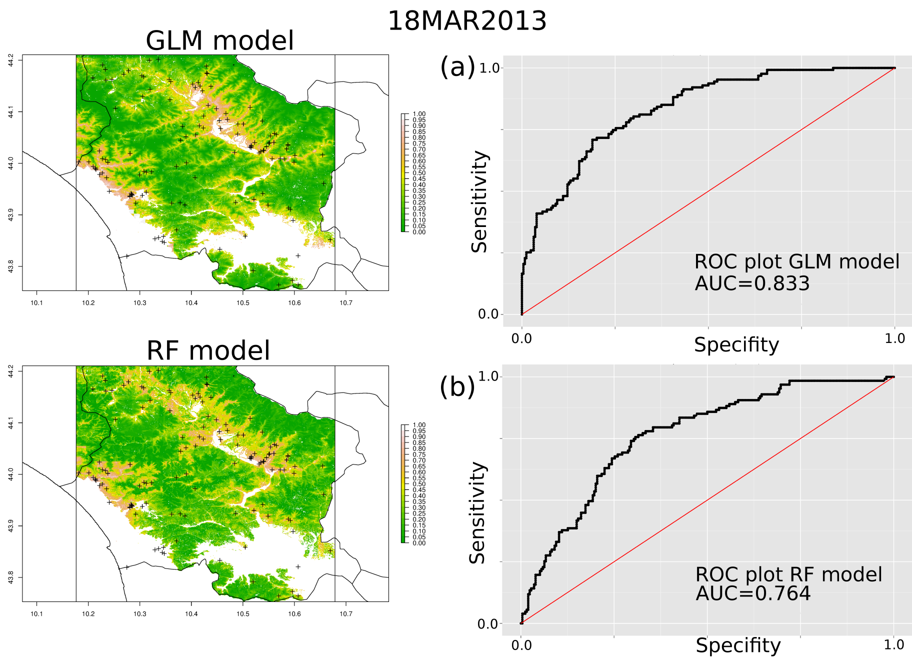

Figure 9. 18MAR2013 landslide hazard maps and corresponding Receiver Operating Characteristics Curves (ROC) with the values of the underlying area (AUC). Results from the GLM model **(a)** and from the RF classifier **(b)**. Crosses points are the event inventory maps produced by field surveys a few days after the rainfall event.

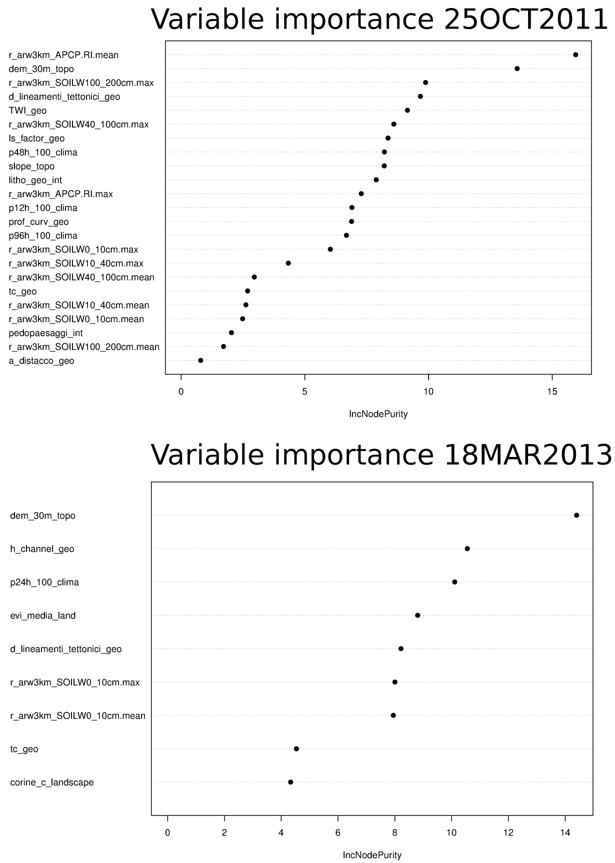

Figure 10. RF's variable importance for classification of landslide occurrence in the 25OCT2011 study case (top panel) and in the 18MAR2013 study case (bottom panel). The meaning of the variables on the *y* axis is reported in Table 4, while extended descriptions are reported in Sects. 2.2 and 2.3.