

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

Valerio Capecchi^{1,2}, Massimo Perna^{1,2}, and Alfonso Crisci³

¹Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via Madonna del piano 10, Sesto Fiorentino, 50019 Firenze

²Consorzio LaMMA, Via Madonna del piano 10, Sesto Fiorentino, 50019 Firenze

³Istituto di Biometeorologia, Consiglio Nazionale delle Ricerche, Via Caproni 8, 50145, Firenze

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

Abstract. ~~We present~~ Our study is aimed at estimating the added value provided by Numerical Weather Predictions (NWP) data for the modeling and prediction of rainfall-induced shallow landslides. We implemented a quantitative indirect statistical modeling for predicting rainfall-induced shallow landsliding. We consider as input layers both static thematic predictors, such as of such phenomena by using, as input predictors, both geomorphological, geological, climatological information, and numerical weather model's forecast. Two different and numerical data obtained by running a limited area weather model. Two standard statistical techniques are used to combine together the above mentioned predictors the predictor variables: a Generalized Linear Model and Breiman's Random Forests. We tested these two techniques models for two rainfall events that occurred in 2011 and 2013 in Tuscany region (central Italy). Model's evaluation is measured by means Modeling results are compared with field data and the forecasting skill is evaluated by mean of sensitivity-specificity ROC Receiver Operating Characteristic (ROC) analysis. In the 2011 rainfall event, the Random Forests technique performs slightly better, whereas in the 2013 rainfall event the than 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, with Area Under the ROC Curve (AUC) values around 0.91 vs 0.84. In the 2013 rainfall event, both models provide AUC values around 0.7.

Using the variable importance parameter output provided by the Random Forests algorithm, we assess the added value carried by numerical weather forecast, in particular in

. The main results are: (i) for the rainfall event occurred in 2011 most of the NWP data, and in particular hourly rainfall intensities, are classified as "important" and (ii) for the rainfall event characterized by deep atmospheric convection and heavy precipitations, occurred in 2013 only NWP soil moisture data in the first centimeters below ground is found to be relevant for landslide assessment. In the discussions we argue how these results are connected to the type of precipitation observed in the two events.

1 Introduction

In the last years, in the north-western part of Tuscany region and nearby areas, (central Italy), were hit by 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 ; Perna et al., 2015). In particular the north-western part of Tuscany region is very prone to frequent and severe rainstorms, due to its geographical position and its steep orography close to the sea, making this area one of the wettest in northern and central Italy (Brunetti et al., 2009). Because of the fragility of the territory (Avanzi et al., 2013), large 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. To mention a few of these disastrous events, on 24-25 December 2009 heavy rainfalls and snow melting triggered more than 600 landslides in the Serchio river valley and Apuan Alps (Avanzi et al., 2013). During October

2010 a relatively long wet spell played a significant role in predisposing the slopes instability. Subsequently the rainstorm that affected the Massa-Carrara area on 31 October 2010, triggered hundreds of landslides (Avanzi et al., 2013). On 25 October 2011 a mesoscale convective system (Rebora et al., 2013) caused large precipitation amounts in the Magra river valley; flash floods, landslides and debris flows occurred in the area causing huge economic losses and thirteen fatalities (Rebora et al., 2013; Galve et al., 2014).

~~In this work, we~~ To arrange efficient alarm systems and in order to reduce property damage and hazard for human lives, civil protection plans are adopted on the basis of operational weather forecast (Regione Toscana, 2006). Nevertheless, prediction of rainfall-induced landslides is problematic since it is determined by rainfall infiltration, soil characteristics, antecedent soil moisture content and rainfall history (Guzzetti et al., 2007). Regarding the area under exam (see figure 1), several studies deal with the definition of the critical rainfall thresholds (Guzzetti et al., 2007) for the initiation of shallow landsliding and debris flows. Among these studies, we recall Giannecchini (2005, 2006); Giannecchini et al. (2007, 2012); Rossetti (2010); Segoni et al. (2014). Another approach for the prediction of landslides is a susceptibility assessment (Chung and Fabbri, 1999; Guzzetti et al., 2005a), since future landslides are likely to occur under the same conditions that produced them in the past (Guzzetti et al., 1999). For the north-western Tuscany and nearby areas, several works deal with landslide susceptibility mapping (hereafter LSM): Catani et al. (2005); Ermini et al. (2005); Federici et al. (2007); Avanzi et al. (2009); Fabbri et al. (2009); Catani et al. (2013). Among these studies, Catani et al. (2013) implemented a landslide susceptibility estimation in the Arno river basin, by using the Random Forests technique. In particular they considered some issues related to the resolution of the mapping unit and to the optimal number of landslide conditioning variables. In particular they found that a mapping unit of 50 m (pixel size) gives the best results with respect to 10, 20, 100, 250 and 500 m mapping units. Moreover they took into account several input predictors such as geo-morphology factors (digital elevation model, slope, curvatures, slopes), lithology factors, land cover, distance to roads, rivers and faults and climatology factors. They concluded that the optimal number of predictors for the classification ranges from 9 (mapping unit 10 m) to 24 (mapping unit 20 and 50 m). None of the cited papers make use of numerical weather forecast as input predictors.

Recently (Schmidt et al., 2008; Segoni et al., 2009; Mercogliano et al., 2013b; Segoni et al., 2014) and climatological Numerical Weather Predictions (NWP) data are arising as a promising and reliable tool for the prediction of shallow landslides triggered by precipitation. In New Zealand, Schmidt et al. (2008) implemented a landslide forecasting system based on three components: (i) regional NWP data (forecast up to 48 h ahead), (ii) a hydrological model fed with NWP data aimed at simulating soil moisture and

groundwater levels and (iii) a slope stability model aimed at estimating failure probabilities within a hillslope. The spatial resolution of NWP data is 12 km, whereas the resolution of the hydrological model and of the slope stability model is 30 m. For a specific extreme event occurred in the lower North Island of New Zealand on February 2004, the authors achieved hit rates of about 70-90% and false alarm ratios of about 30%. Segoni et al. (2009) implemented a real-time forecasting chain for rainfall-induced shallow landslide to be used for civil protection purposes. The architecture of the forecasting chain is quite complex and takes advantage of techniques and tools from different fields including meteorology, hydrology, geomorphological and geo-technical modelling, remote sensing and GIS. For what concerns the NWP data, the authors downscaled the rainfall forecast provided by a limited area model using a metagaussian model aimed at reproducing small-scale rainfall fields. The final resolution of the precipitation forecast is 1.75 km in space and 10 min in time. Similarly Mercogliano et al. (2013a,b) implemented a forecasting chain for rainfall induced shallow landslides. They coupled the NWP data (Segoni et al., 2014) produced at 2.8 km of resolution and statistically downscaled to 10 m, with a physically based slope stability simulator (hydrological and geo-technical tool). In Mercogliano et al. (2013b), the authors tested the procedure in a pilot site in the north-western part of Tuscany (Lucca, Pistoia and Prato administrative provinces) for a specific rainfall event. A quantitative validation of the results was not performed, nevertheless the authors concluded that additional work is needed (Schmidt et al., 2008; Segoni et al., 2009; Catani et al., 2013) to better understand the spatial organization of the input parameters and improve the quality of the results.

Our work takes its origin from these latter papers. 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 model techniques: 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 section 2.2 for further details~~) characterizing the areas affected by the heavy precipitations from a climatological point of view. Moreover, since recently (Schmidt et al., 2008; Segoni et al., 2009; Mercogliano et al., 2013b), 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 and dynamical NWP data 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 (~ 2 –20 km of spatial resolution), with the micro- γ scale (≤ 20 m of spatial resolution, according to Orlandi, 1975) which is an average value of the mapping unit of landslide size occurring at the basin scale (Guzzetti et al., 1999, 2005b). This goal is achieved without performing any downscaling of the NWP data (grid point about 3 km of horizontal resolution) to a finer resolution. In this way we preserve the original information content provided by the numerical model without introducing any artificial knowledge on the precipitation patterns.

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 modeling results is performed through the analysis of the Receiver Operating Characteristics Curve (ROC) curve 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), characterized the precipitation type. Using the Beside the reliability of the statistical modeling of rainfall-induced shallow landslides, the innovative feature of our study is the assessment, by using 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) and on soil water content at different levels below ground. diagnostics, of the relative importance of NWP data in LSM, since so far there is a lack of studies focusing on this issue.

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) 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 section 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 section 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 section 2.4. The preliminary results for the selected study cases are shown in section 3 and discussed in section 4.

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 We implemented 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, see figure 1) on 25th October 2011 (Lunigiana) and on 18th 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 25th October 2011

The first rainfall event, hereafter 25OCT2011, occurred on 25th October 2011 and involved the Lunigiana area belonging to the administrative province of Massa-Carrara (see inset figure in the left side of figure 1). The area is located along the Apennine Apennines 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 wettest in Italy (Brunetti et al., 2009). 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 km² (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) and Schiavini et al. (2013) studied the 2009 landslide inventory 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).

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~~ precipitations over the same region, i.e. an 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 h, 317 mm/12 h, 243 mm/6 h, 158 mm/3 h and 67 mm/1 h. ~~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~~ For this rain-gauge, Regione Toscana (2011) reported estimates of return period rainfall events for 1-, 3-, 6-, 12- and 24-hour duration: 51, 438, >500, >500 and 293 years respectively.

~~The landslide~~ Landslide inventory map for this event was created by both field surveys of the expert personnel of the Regional Civil Protection Office. The and using Rapid-Eye images. A semi-automatic detection algorithm based on Rapid-Eye pre/post event images (13 October pre-event and 29 October post-event) was applied. The Rapid-Eye data were multispectral 5-bands with 5-meter of spatial resolution. Shadow areas in satellite images created problems in the lower parts of hillslopes, especially in less populated and hardest to reach areas, so possible commission/omission errors may occur. Since technical authorities surveys were conducted mostly close to populated areas, possible omission errors may occur in particular in forested areas. The final inventory map reported 243 shallow landslides in an area of about 212 km² (see the minimum bounding rectangle in the left side of figure 1) while, whereas the convex hull where landslides were observed has an area of the area has an extent of about 123 km². Average area of landslides is around 2260

m² and average perimeter is around 181 m, whereas the average altitude of landslide initiation points is around 588 m. Estimating how much complete is a landslide inventory is a difficult task (Malamud et al., 2004). Nevertheless in figure 2 we show the landslide probability density as a function of landslides area (Malamud et al., 2004) that follows the same trend of frequency-area distributions of landslides as found in previous works (Guzzetti et al., 2002; Malamud et al., 2004).

2.1.2 Study case 18th March 2013

The second rainfall event, hereafter 18MAR2013, occurred on 18th March 2013 and involved the Garfagnana area belonging to the administrative province of Lucca (see inset figure in the right side of figure 1). This is mainly a mountainous area (the average elevation of the main catchment is 717 m) ~~and, as Lunigiana, also this area, and~~ 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 km² plus several other minor rivers.

Geological features of the area are very similar to the ones described in section 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 18th March 2013 (the one under exam here).~~ 2013. Due to the high degree of saturation of the soils and due to the surface runoff, on 18th March 2013, several regional hydro-meters 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 (Regione Toscana, 2013), 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 UTC). 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

mm/h (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 mm/h. ~~See table 4 for summary statistics on observed and modelled 1-hour precipitation intensities. For this rainfall event, it was not possible to estimate return period due to the lack of a long and consistent time series in the area. Nevertheless Regione Toscana (2013) evaluated this event as one of the largest of the previous 20 years.~~

~~The landslide~~ Landslide inventory map for this event was created by field surveys of the expert personnel of the local Genio Civile Office. The map reported 127 shallow landslides in an area of about 2038 km² (see the minimum bounding rectangle in the right side of figure 1), while the convex hull where landslides were observed has an area of about 1416 km². Due to time constraints, it was not possible to integrate field surveys observations with pre/post event images analyses and thus commission/omission errors are very likely to occur. The field surveys did not collect any information about the size (area and perimeter) of landslides, so we cannot report descriptive statistics on the inventory.

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~~. 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 section 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 1.

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 LSM~~ 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
- Slope (SLP): it is directly derived from the DEM layer. It controls the driving forces (component of weight of material in the direction of failure) acting on hillslopes. For these reason, it has been widely used in geomorphology and landslide mapping studies (Guzzetti et al., 1999; Goetz et al., 2011; Catani et al., 2013)
- 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) Goetz et al. (2011) and Catani et al. (2013) and used this parameter (and its standard deviation) in their landslide susceptibility study based on RF model GAM and RF models respectively~~
- 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~~ Goetz et al. (2011) and Catani et al. (2013)

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

$$TWI = \frac{a}{\tan \beta}$$

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 LSM~~ (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 (Mossa et al., 2005; Mancini et al., 2010). Recently it has been used by several authors as a predictor in ~~landslide susceptibility mapping LSM~~ (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 1:10000) and the regional pedological database (scale is 1:50000). 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, the sample areas extent were about 15-25 km² 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~~ personnel.

- 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 LSM~~ 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 literature~~. 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, 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~~ relationship 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. Since it is a directional variable we considered the sine of the SSS.

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~~ return period (expressed in years). In this, our predictor is slightly different from that one considered by Catani et al. (2013) who evaluated the ~~returning return~~ period of a defined precipitation amount occurring in a defined time interval.

– Rainfall 12, 24, 48, 96 hours 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 return~~ period. In the present study, the ~~durations considered are duration~~ considered is 12, 24, 48 and 96 hours and the ~~returning~~ return period is 100 years. It was derived from statistical analysis of rainfall time series of regional rain-gauges network (30 years minimum), that were interpolated over the area of interest.

~~Besides Beside~~ 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~~ The original Corine dataset (Bossard et al., 2000) produced at scale 1:10000 has 45 classes. Only 8 classes belong to the 25OCT2011 area; 95% of the territory is covered by 6 classes namely broad-leaved forest, agriculture with significant areas of natural vegetation, complex cultivation patterns, mixed forest, non-irrigated arable land and transitional woodland-shrub. Ten classes belong to the 18MAR2013 area, namely: broad-leaved forest, mixed forest, complex cultivation patterns, agriculture with significant areas of natural vegetation, transitional woodland-shrub, non-irrigated arable land, discontinuous urban fabric, olive groves, natural grassland and coniferous forest.

– 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 This~~ 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 ~~reflectance~~ reluctance. 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 2, while the geographical extent of the simulation area is depicted in figure ??1 (outer rectangle). 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 UTC 24th October 2011 for the 25OCT2011 event and 00 UTC 17th March 2013 for the 18MAR2013 event. The length of the simulations is 48 hours for both events.

~~The choice of the geographical static instability factors has been discussed in section 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 ~~accumulated~~ accumulated 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 cm, 10-40 cm, 40-100 cm and 100-200 cm below ground. This is the partition of soil implemented in the Noah land surface model (Chen et al., 1996) ~~and which is incorporated in 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 1 (bottom part of the table)

To validate the rainfall forecast provided by the WRF model in terms of quantitative precipitation forecast (QPF), we used the data gathered by the remotely automated weather station network operated by the National Civil Protection Department. Data from 20 automated rain-gauges recording precipitation every hour were collected for 25OCT2011, while 60 rain-gauges were collected for 18MAR2013.

Locations of rain-gauges are shown in figure 3. For each rain-gauge locations, we extracted the predicted values of the numerical simulation and compared them with the observed rainfall amounts registered in a 24 hour period, namely from 00 UTC 25th October 2011 to 00 UTC 26 October 2011 for 25OCT2011 and from 00 UTC 18th March 2013 to 00 UTC 19 March 2013 for 18MAR2013. The ability of the model to simulate the precipitation amounts was analyzed using the contingency tables (Wilks, 2011) for selected rainfall's thresholds and using quantitative indexes, namely Root Mean Square Error (RMSE) and multiplicative bias.

2.4 Description of the statistical modeling of landslide hazard

For the two rainfall events under exam, we ~~developed~~ implemented 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 was performed to resample, geo-statistically,~~ the NWP outputs (~~produced at~~ 3 km of horizontal resolution) to the resolution of the static instability factors (~~produced at~~ 30 m of horizontal resolution). As stated in section 1, this has been done because we want to preserve and evaluate the information provided by NWP data produced at their native scale (meso- γ scale) in a landslide susceptibility framework. 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 the literature~~ (Carrara et al., 1999; Guzzetti et al., 1999; Catani et al., 2005) 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 figure ??.~~ Summarizing, ~~we developed, implemented and tested 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)~~ LSM (Brenning, 2005). The RF classifier was chosen because it is very flexible, recently used in ~~landslide susceptibility mapping LSM~~ (Stumpf and Kerle, 2011b; Catani et al., 2013) and has interesting and useful diagnostics (~~see below~~). ~~No interactions were implemented between the two models.~~ (Brenning, 2005; Stumpf and Kerle, 2011a,b; Vorpahl et al., 2012; Catar

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 $\{X_i\}, i = 1, \dots, n$. In the present case, L is the presence/absence of landslide, while the $\{X_i\}$ variables are the static parameters detailed in section 2.2 and the NWP outputs detailed in section 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). See the review article by Brenning (2005) and references therein for a list of works using logistic regression. To our purposes GLM is appropriate because (a) it handles both categorical and continuous predictor variables, (b) it does not assume a specific distribution of predictors. In the present study the GLM is taken into account because it is considered a sort of benchmark for LSM. A drawback of GLM is the incapacity to model potential non-linearities in the relationship between response and explanatory variables. In the present work, logistic regression is performed after applying an automatic stepwise backward-forward variable selection based on the Akaike Information Criterion (AIC). The GLM is available in R's 'stats' package.

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 assess the predictive skill of the tree. The most common classification among all the tree is the prediction of the RF.

Schematically some Díaz-Uriarte and De Andres (2006) stress some of the features and advantages of the RF technique are RF in bio-informatics classification problems. Schematically RF: 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 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 the literature 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) (Cutler et al., 2007) 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 susceptibility maps in Ecuador using geomorphometric attributes and information on land use. Recently,

In landslide mapping or geomorphological context it was used in Brenning (2009); Stumpf and Kerle (2011a,b); Vorpahl et al. (2012); Catani et al. (2013) applied the RF algorithm to produce landslide susceptibility maps for the Arno river basin (about 9100 km²) at different mapping unit ranging from 10 m to 500 m. They considered a variety of predisposing factors mainly related to the lithology, the land use, the geomorphology, the structural and anthropogenic constraints. found that the RF is feasible and robust provided that a preliminary procedure is performed aimed at assessing the optimal number of trees of a single realization of RF. Brenning (2009) in his study on an automatic rock glacier detection procedure using multispectral remote sensed data, found that RF tends to overfit to the training data, achieving unsatisfactory results in cross-validation. Marmion et al. (2008) found that RF provides low classification accuracy in a geomorphological application if compared to other statistical methods. In the paper by Brenning (2005), the author discuss how RF prediction is problematic outside the training set and in general RF was found to be prone to overfitting.

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 section 3.2 and discussed in section 4, this method provided interesting results and gave hints for meaningful discussions. Beyond these disadvantages, to our purposes, 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). Variable's importance can be assessed by mean of two different measures: mean decrease in accuracy and mean decrease in node impurity. We chose the mean decrease in accuracy which is meaningful in the present context since AUC is chosen as the performance measure of the predictive skill of the model. As done by Brenning (2009) and Catani et al. (2013) the number of trees used in this study is 200.

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

In both statistical methods, order to remove methodological limitations due to the lack of a performance estimation of the results on an independent test set, we ran both statistical models (GLM and RF, we used, as predictors, the same set of layers detailed in section 2.2 and 2.3 and summarized in table 1.

) for both study cases (25OCT2011 and 18MAR2013) dividing randomly each event inventory dataset in two groups: a training set and a validation set. The first group is 70% of each event inventory dataset, whereas the validation set is the remaining 30%. The relative size of the two groups (70% and 30%) is based on Carrara et al. (2008). For modeling fitting, the number of non-landslide locations equals the number of landslide locations according to Brenning (2005) and Goetz et al. (2011). The non-landslide locations were selected randomly in the two areas of interest.

Due to the small number of elements in the resulting training sets (in particular for 18MAR2013), overfitting may occur in predictive relationship of the models. So we decided to perform multiple runs of the model (100 runs for each study case), with different random choice of the training and test sets, since Stumpf and Kerle (2011b) found that RF variable importance is almost stable when running multiple classification runs.

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 section 3.1 we briefly present the validation show the verification of the WRF predictions for the selected dates using rain-gauge observations as ground truth. In section 3.2 we present the landslide hazard maps for the two events along with their accuracy accuracy of the landslide susceptibility maps in terms of ROC plots and curves and of the corresponding underlying area. In section 3.3 we

show the variable importance outputs provided by the RF algorithm.

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 To have an indication of the precipitation patterns simulated by the WRF model for 25OCT2011 event (from 00 UTC 25th October 2011 to 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 figure 3). For, in figure 4 we show a visual comparison between the rainfall data simulated by the WRF forecast with the observed rainfall data collected in the 20 rain-gauges see figure 4, of the rainfall amounts predicted by the model (panel (a) for WRF data and panel (b) 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-hour accumulated precipitation). In figure 5, we show the False Alarm Rate and Probability of Detection for selected rainfall thresholds. In table 3) and the observed data (panel (b)). In table 3 (top), we show the descriptive statistics (average values and percentiles) of the observed and modelled rainfall data for the selected modeled rainfall data. To evaluate the under-estimation of the numerical data, we calculated the False Alarm Rate (FAR) and Probability of Detection (POD) for selected rainfall thresholds (see figure 5). To estimate errors and bias, we computed the RMSE, the multiplicative bias and the correlation coefficient: the RMSE is about 150 mm and the multiplicative bias is about 0.49, whereas the correlation coefficient is about 0.19. To understand if displacement errors occur in the WRF simulation, we calculated descriptive statistics not only for the numerical data extracted for 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 1st quantile of observed data). Nevertheless if we extract the modelled data in the locations, but also for all the numerical data extracted in a larger area affected by the 25OCT2011 event (see the inset rectangle in the picture in the left side of figure 1), we can see from table 3 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 ($218 \text{ mm } 24\text{h}^{-1}$) 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 $286 \text{ mm } 24\text{h}^{-1}$. It has to be noticed, nevertheless, that they used the ECMWF-IFS analysis at 12 UTC 24th October 2011 (instead of 00 UTC 24th October 2011 as done here) and that their model horizontal resolution is 1.5 km (instead of 3 km as setup here). Descriptive statistics of these data are reported in table 3 (top).

For the 25OCT2011 event Since in the rain-gauges dataset we observed high rainfall intensities (up to 67 mm/h in the Pontremoli rain-gauge station), a special regard was devoted to the verification of the maximum rainfall intensities (mm/h) predicted by the model. In the rain-gauges dataset we can observe very high rainfall intensities (up to 67 mm/h in the Pontremoli rain-gauge station). On average, over the 20 rain-gauges considered, the Root Mean Square Error (RMSE) of modelled RMSE of rainfall intensity data is around 6.19 mm/h (summary and the multiplicative bias is about 0.68. Summary descriptive statistics of both observed and modelled modeled 1-hour rainfall amounts intensities are shown in table 4).

For the 18MAR2013 event it was possible to collect 60 rain-gauges recording precipitation every hour (see the locations of the rain-gauges in figure 3). As shown for the 25OCT2011 case, for, to have an idea of simulated rainfall patterns, we show in figure 4 a visual comparison between the rainfall data simulated by the WRF forecast with the observed rainfall data collected in the 60 rain-gauges see figure 4, model (panel (c) for WRF data and) and the observed rainfall data (panel (d) 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 UTC 18th March 2013 to 00 UTC 19 March 2013), the distribution, in terms of percentiles, of the model data and of the observed data is summarized. The quantitative evaluation of forecasting skill reported a RMSE about 45 mm , whereas the multiplicative bias is about 0.59 and the correlation coefficient is about 0.63. To evaluate the distributions of both numerical data and observations we show in table 3. The RMSE is $40 \text{ mm } 24\text{h}^{-1}$ and the correlation coefficient is 0.63. The False Alarm Rate and Probability of Detection plots are shown in figure 6 (bottom), the descriptive statistics. To evaluate over/under-estimation we computed the FAR and POD skill scores (see figure 6), where FAR and POD skill scores values 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, modeled data.

To evaluate the feasibility of an early warning system for rainfall-induced shallow landslide based on NWP data, we reported the computational time required by the WRF simulation. Considering the extent outlined in figure 1 and the resolution of both static predictors and NWP outputs and considering the statistical models adopted, landslide hazard maps were produced requiring less than 10 minutes of CPU time on a HPC multi-core grid spacing of 3 km, the CPU time is approximately 2 hours on a 8 processors Quad-core HPC Linux server.

3.2 Evaluation of landslide hazard maps

To visualize the modeling effects and to enable a few considerations, outputs of the model prediction have been provided. In figures 7 and 8 we show a single output, for each study case, among the 100 runs performed. The results for the study case 25OCT2011 are shown in figure 7(a) for the GLM model and 7(b) for the RF model. While the The results for the study case 18MAR2013 are shown in figure 8(a) for the GLM model and 8(b) 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 section 2.4, one of the main features of To evaluate the accuracy of landslide hazard maps, we calculated ROC and AUC values of the 100 runs performed on the test sets. Summary statistics of the AUC values achieved by the GLM and RF model runs are reported in table 5.

To evaluate the feasibility of an early warning system, based on the forecasting chain proposed, we reported the CPU time required by GLM and RF. Considering the extent 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 ?? shows the variables importance for the two study cases: two areas of interest (see figure 1), the resolution of both static predictors and NWP outputs and considering the statistical models

adopted, landslide hazard maps were produced requiring less than 10 minutes of CPU time on a 8 processors Quad-core HPC Linux server.

3.3 Evaluation of importance of predictor variables

The optimability of the parameter set is evaluated on the basis of the out-of-bag error as explained in Catani et al. (2013). Out of the 34 parameters considered in the analysis, a combination of 22 parameters represents the best configuration for 25OCT2011 on average over the 100 RF runs performed, whereas a combination 12 parameters is the best set, on average, for the 18MAR2013. Since in both cases the cardinality of the optimal set is quite high with respect to the total number of predictors, as indicators for the stability of the variable importance we decided, subjectively, to report (i) which is the average position of each predictor variable and (ii) how many times (in %) a predictor variable was classified as “important” in the first ten (five) places of the ranking for 25OCT2011 (top) and 18MAR2013 (bottom). In other words, we decided to restrict the optimal set of predictor variables classified as important to the first 10 (5) positions and count the average position of each predictor among the 100 RF runs.

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 25OCT2011 elevation, distance from main tectonic features, vegetation index, maximum soil moisture in the four layers 0-10 cm, 10-40 cm, layer 40-100 cm, total NWP rainfall amount and mean NWP rainfall hourly intensity are the six most frequent variables with average rank 4.5, 3.7, 3.4, 6.2, 2.6 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 2.6 respectively. Altitude above channel network, maximum soil moisture in the layer 10-40 cm and maximum NWP rainfall hourly intensity are the next most frequent variables in the first 10 positions for 25OCT2011. The other variables approximately occur in the first 10 positions of the optimal set, in no more than one third of the 100 years and occurring in 12 h, 48 h and 96 h. Finally, two are hydrology-related predictors, namely: time of concentration and Topography Wetness Index. All these layers are highlighted in table 1 with the ★ symbol. For RF runs performed.

Regarding 18MAR2013, nine predictors are classified as important by the RF algorithm. Two are dynamical NWP predictors, namely mean and the most frequent variable is elevation, which is always classified in the first position,

followed by altitude above channel network, 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 hours and having a returning period of, slope structural setting and land-cover, each one occurring in the optimal set in at least half of the 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 1 with the ★ symbol.

RF runs.

Complete results for 25OCT2011 and 18MAR2013 are reported in table 6.

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 an effort to elucidate the impact of NWP data in LSM, we implemented a simple statistical method by using variable importance provided by the RF algorithm. In the procedure implemented, the two models have no interactions between each other and are tested separately (see figure ?? for a schematic flowchart of the modeling). We chose the GLM model because it has been proven (see references cited in section 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 figure 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, following we discuss how the major findings of the work demonstrate the usefulness of NWP in LSM.

In spite the main aim of the work, a few discussions need also to be addressed regarding (i) 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., 2014). The use of NWP data is justified in reliability of NWP data used and (ii) 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 accuracy of the landslide susceptibility maps produced by the RF and the GLM, this last being a sort of benchmark for LSM.

4.1 Reliability of NWP data

For what concerns 25OCT2011, from the analysis of figure 5 and table 3 (top), it is quite clear how the model largely under-estimates the rainfall amounts for high thresholds (broadly greater than 60 mm). This fact is also confirmed by multiplicative bias, RMSE and correlation coefficient achieved (0.49, 150 mm and 0.19 respectively). The lack of the model in predicting the exact locations of the deep convection processes can also be visually appreciated from figure 4, panels (a) and (b). The general under-estimation of NWP model forecast in predicting precipitation maxima during intense precipitation events is a well known issue and was recently addressed for this specific rainfall event by Buzzi et al. (2014). On the other hand, from table 3 (top), we can state that the model is able to capture the characteristics of heavy rainfall event in the area of interest. Mercogliano et al. (2013b) 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 don't 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 sections 2.1.1 and 2.1.2, the study cases selected show different features: In particular the maximum value achieved by the model in the whole area of interest ($218 \text{ mm } 24\text{h}^{-1}$) 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 $286 \text{ mm } 24\text{h}^{-1}$. It has to be noticed, nevertheless, that they used the ECMWF-IFS analysis at 12 UTC 24th October 2011 (instead of 00 UTC 24th October 2011 as done here) and that their model horizontal resolution is 1.5 km (instead of 3 km as setup here). Similarly, from table 4, we can assess that the model is also able to reproduce adequately the hourly rainfall intensities yielding moderate RMSE (19 mm/h) and a multiplicative bias around 0.68. Overall we can conclude that despite the under-estimation of precipitation maxima on a point-to-point verification, for 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 the numerical weather model provide results as much reliable as possible considering the time of integration and the resolution adopted.

Similar considerations can be assessed for the verification of 18MAR2013 was characterized by a weather storm with precipitation spread over a 24-hour 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, rainfall forecast. Table 3 (bottom) and figure 6 show the over-estimation of WRF rainfall data for low thresholds and under-estimation for high thresholds. This is a feature of the model which was also found in similar studies in Italy (Oberto et al., 2012). However for this event, quantitative forecasting skills (multiplicative bias and RMSE) are better than those found in Oberto et al. (2012). This mainly because this event was not characterized by deep convection processes (Regione Toscana, 2013) and thus fine-scale variability is represented more adequately.

Considering that high-intensity rainfall events are often spatially very variable, fine-scale variability cannot be properly modeled by the WRF simulations produced with grid point about 3 km. Previous similar studies, in particular Mercogliano et al. (2013b), used statistical techniques to resample, from meso- γ to micro- γ scale, rainfall data produced by NWP model. In general downscaling statistical methods, by construction, introduce a source of uncertainty because they dependent upon the choice of predictors (mainly topographical variables) and they suppose the stationarity in the predictor-predictand relationship (Fowler et al., 2007). On the other hand regarding downscaling dynamical methods, it is well known that increasing horizontal resolution produces more skillful forecasts (Buzzi et al., 2014). Hopefully, in the near future operational NWP data will be produced with grid point spacing about 1 km (Schwartz, 2014) or below. Moreover, it is commonly accepted that predictions of intensive precipitation episodes of convective origin should be improved through data assimilation techniques (Schwitalla and Wulfmeyer, 2014). For these considerations, at the present stage of the work, we preferred to preserve the information content of NWP output produced at meso- γ scale and to evaluate their value and possible positive impacts on landslides modeling.

4.2 Accuracy of landslide susceptibility maps

Although AUC values, in general, are not comparable among studies in different study areas, the area under the Receiver Operating Characteristics curve was found to be a good representative index of the accuracy of landslide probabilistic forecast (Fratini et al., 2010). Results here achieved and summarized in table 5 are similar to those found in recent literature, considering in particular that the areas under exam here do not contain large “easy-to-predict” portions (i.e. flat valley floors or less steep forelands). 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 susceptibility and the performance estimation technique is the same. In the present study, for the 25OCT2011 study

case, the AUC values obtained are quite high both for the GLM and, even better, for the RF classifier and are comparable to those found in Catani et al. (2013). On the other hand, for the 18MAR2013 case, average AUC values are lower for both models. This fact might be related to the potential incompleteness of the inventory map produced which is a limit of the study case.

In the susceptibility maps produced (see figure 7(a) and (b) in particular a sample output in figure 7), 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 As a limiting factor of the study, it has to be stated that, in the context of landslide prediction, the relevance of 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 figure ?? 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 ?? 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 section 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 maps produced is limited by the fact that the only variable occurring in the importance ranking (see figure ?? 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 section 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).

As stated in the previous section, models were fitted to an event landslide inventory and to NWP data simulated for the same date. In 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 next future and operationally, the idea is to calibrate the models for some study cases (i.e. at least 10-15 rainfall events) and then apply the models to other independent cases, or, on a daily basis, feed the models with NWP data and get the forecast for the following day. In this preliminary stage, it was not possible to consider additional study cases, because it was not possible to collect enough observed landslides. Moreover, in the present work, we did not trained the models for 25OCT2011 and then applied them for 18MAR2013, because 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 downsealed 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 show different characteristics from a meteorological point of view (convective precipitation vs stratiform precipitation) and because the two events occurred in two different areas even if geographically contiguous. However the work here presented is not strictly a hindcast exercise since the NWP data are not a numerical description of 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, rainfall occurred (that is, we did not use reanalysis data as initial and boundary conditions), on the contrary NWP were obtained by numerical integration using global analysis and forecast.

In spite of the simplicity of such the statistical approach, the another drawback of the method methods proposed is that, being data-driven, a model built up for one region and for one particular event, cannot be applied, without any recalibration, to a neighboring area or for a similar rainfall event. Moreover, it.

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) (Capecchi and Brocca, 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 total time required to produce a single landslide susceptibility map is around 2 hours and 10 minutes. From an operative point of view, since global NWP forecast are available each day at about 7 UTC (when the analysis time is 00 UTC), there's room for short term regional weather forecast and GLM or RF runs, since regional daily assessment of landslide risk for the following day needs to be issued approximately before 12 UTC. Obviously, this CPU time can be further reduced by using additional or more recent processors to meet specific needs (for example larger areas of interest) of the end-users.

4.3 Importance of NWP predictor variables

Our study is mainly 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., 2013b). To our knowledge, so far there's no literature regarding the evaluation of the importance of NWP data on LSM. The present study differs from the cited papers because: i) we consider not only NWP rainfall amounts but also NWP hourly rainfall intensities and soil moisture in four layers below ground, ii) we don't perform any statistical downscaling of the NWP data towards the LSM unit and iii) we combine both static predictors and NWP data into a "black-box" statistical model instead of using a physically based model. In particular this latter feature allows us to assess objectively (with the limitations associated to RF) the impact of NWP data on LSM by using the variable importance parameter. From table 6 we can assess that NWP information is crucial for the 25OCT2011 event, whereas has a more limited impact for the 18MAR2013 event. In particular for 25OCT2011, it is worth noting that not only the total NWP rainfall amount is found important, but also the mean, and to some extent the maximum, hourly NWP rainfall intensities are often ranked in the first positions among the 100 RF runs. In particular the positive impact of hourly rainfall intensities, is a relatively innovative finding and here it is found that, as a predisposing factor, it is as much important as the total rainfall amount. In recent literature, only Segoni et al. (2009) used the radar data to extrapolate hourly or sub-hourly rainfall intensities for LSM purposes. On the contrary, regarding 18MAR2013,

NWP rainfall data are occasionally classified as important in RF runs. In fact, 18MAR2013 was characterized mainly by stratiform precipitation and the role of the prolonged antecedent precipitations played a relevant role in landslide activity. Not surprisingly, for this event the maximum soil moisture in the first 10 cm below ground is classified as important in almost 75% of the runs with an average rank of 3. This result confirms the findings of recent studies investigating the strong linkage between soil moisture and landslide occurrence (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 amounts for the triggering of landslides. In addition, this finding suggests the possible use of NWP soil moisture contents for the initialization of hydrological states in modeling, instead of using NWP rainfall data as the main driver of soil moisture patterns (Schmidt et al., 2008). It is also worth noting how for this rainfall event, NWP root-zone soil moisture contents (namely water content in the layer 100–200 cm below ground) is classified as important in more than 25% of the 100 RF runs. One can argue that this is the effect of the prolonged antecedent precipitation. Similarly, for 25OCT2011, NWP soil moisture in the layers 10–40 cm and 40–100 cm below ground are found to be important in 67% and 90% of the RF cases. This effect needs to be further investigated to understand if there is a causal relationship between sub-surface soil moisture contents and land movements or if RF overfitting occurs.

In general, in spite of RF ease of use, the fact that RF overfitting may occur in geomorphological applications, (Brenning, 2005; Marmion et al., 2008; Brenning, 2009), is a possible drawback of the present study and any result regarding predictor variables importance should be analyzed with care. For this reason we tried to mitigate possible accidental effects in the predictor/predictands relationships by performing multiple runs of the models varying the choice of the training and test sets.

A few discussions need to be addressed for what concerns the importance of the static predictors. This work does not add anything innovative for what concerns the relative importance of the static layers in LSM with respect to other similar and recent works (Catani et al., 2013). The common findings with this paper are: elevation is always (or very often) ranked in the first positions, (ii) the slope curvature variables do not seem to have any influence on landslide susceptibility, (iii) it is confirmed the lack of importance that the prediction model assigns to the hydrology-related variable. On the other hand, some discrepancies found are: (i) the land cover has a relevant influence on landslide assessment for 18MAR2013, while Catani et al. (2013) found a relatively marginal impact, (ii) the climate-related predictor variables in Catani et al. (2013) are often classified as important in the first positions at several mapping unit, whereas in the

present paper these variables are classified as important only occasionally. Nevertheless for these predictors it is interesting to stress how: (i) for 25OCT2011 (large precipitation amount in a relatively short time period) the 12 and 24 hours duration and 100 years return period occur in the RF rank, while (ii) for 18MAR2013 the forecasting chain here developed into a fully operative warning system, further 48 and 96 hours duration and 100 years return period occur, taking into account the attitude of the territory to be affected by long-lasting and prolonged precipitations.

5 Conclusions

Our study presents the opportunity to integrate NWP information content into a statistical method to take account in advance of the relationship between rainfall, soil moisture and landslide movements. The relative importance of the NWP data is assessed by using RF diagnostics and it is found that rainfall amounts and hourly intensities are relevant for the convective precipitation case, whereas for the prolonged antecedent precipitation case, the NWP surface soil moisture is as much important as geomorphology-related predictors. We therefore tried to bridge the gap between the micro- γ scale (≤ 20 m), which is the characteristic scale of geomorphology, hydrology and climate related information, with the meso- γ scale, which is the current typical scale of the NWP forecast. The final output of the statistical modeling has been evaluated by dividing the event inventory map in training sets and test sets and the results are estimated by using the AUC values. To account for possible overfitting in model estimates, or, in other words, to evaluate the stability of variable importance, an “ensemble” of model’s run was performed.

A deeper interpretation of the preliminary results obtained is an ongoing task. Further well-documented severe rainfall events need to be addressed and the results have need to be validated with ground observations, independent ground observations to take account of possible overfitting in the procedures and to calibrate properly the model.

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-supported 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 tec-

- tonics driven by Quaternary uplift in the Northern Apennines: insights from the La Spezia-Reggio Emilia geo-transect, *Quaternary International*, 101, 13–26, 2003.
- Atkinson, P. and Massari, R.: Generalised linear modelling of susceptibility to landsliding in the Central Apennines, Italy, *Computers & Geosciences*, 24, 373–385, 1998.
- Avanzi, G. D., Falaschi, F., Giannecchini, R., and Puccinelli, A.: Soil slip susceptibility assessment using mechanical–hydrological approach and GIS techniques: an application in the Apuan Alps (Italy), *Natural hazards*, 50, 591–603, 2009.
- 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*, pp. 239–246, Springer, 2013.
- 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, *Atmospheric Research*, 138, 268–277, 2014.
- Borghuis, A., Chang, K., and Lee, H.: Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery, *International Journal of Remote Sensing*, 28, 1843–1856, 2007.
- Bossard, M., Feranec, J., and Othel, J.: CORINE land cover technical guide: Addendum 2000, 2000.
- Breiman, L.: Bagging predictors, *Machine learning*, 24, 123–140, 1996.
- Breiman, L.: Random forests, *Machine learning*, 45, 5–32, 2001.
- Breiman, L.: Manual on setting up, using, and understanding random forests v3. 1, 2002.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A.: Classification and regression trees, CRC press, 1984.
- Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, *Natural Hazards and Earth System Science*, 5, 853–862, 2005.
- Brenning, A.: Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection, *Remote Sensing of Environment*, 113, 239–247, 2009.
- Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J.: 1961–1990 high-resolution Northern and Central Italy monthly precipitation climatologies, *Advances in science and research*, 3, 73–78, 2009.
- Bryan, G. H. and Morrison, H.: Sensitivity of a Simulated Squall Line to Horizontal Resolution and Parameterization of Microphysics., *Monthly Weather Review*, 140, 2012.
- 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, *Meteorologische Zeitschrift*, 13, 91–97, 2004.
- Buzzi, A., Davolio, S., Malguzzi, P., Drofa, O., and Mastrangelo, D.: Heavy rainfall episodes over Liguria in autumn 2011: numerical forecasting experiments, *Natural Hazards and Earth System Science*, 14, 1325–1340, 2014.
- Capecchi, V. and Brocca, L.: A simple assimilation method to ingest satellite soil moisture into a limited-area NWP model, *Meteorologische Zeitschrift*, 23, 105–121, 2014.
- Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., and Reichenbach, P.: GIS techniques and statistical models in evaluating landslide hazard, *Earth surface processes and landforms*, 16, 427–445, 1991.
- Carrara, A., Guzzetti, F., Cardinali, M., and Reichenbach, P.: Use of GIS technology in the prediction and monitoring of landslide hazard, *Natural hazards*, 20, 117–135, 1999.
- Carrara, A., Crosta, G., and Frattini, P.: Comparing models of debris-flow susceptibility in the alpine environment, *Geomorphology*, 94, 353–378, 2008.
- 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.
- Catani, F., Lagomarsino, D., Segoni, S., and Tofani, V.: Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues., *Natural Hazards & Earth System Sciences*, 13, 2013.
- 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, *Journal of Geophysical Research*, 101, 7251–7268, 1996.
- Chen, S. and Dudhia, J.: Annual report: WRF physics, Air Force Weather Agency, 2000.
- Chung, C.-J. F. and Fabbri, A. G.: Probabilistic prediction models for landslide hazard mapping, *Photogrammetric engineering and remote sensing*, 65, 1389–1399, 1999.
- 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), *Natural Hazards and Earth System Science*, 12, 327–340, 2012.
- Crozier, M. J.: Prediction of rainfall-triggered landslides: A test of the antecedent water status model, *Earth Surface Processes and Landforms*, 24, 825–833, 1999.
- Cruden, D. and Hu, X.: Landslides in the Rocky Mountains of Canada, in: *Geomorphological Hazards in High Mountain Areas*, pp. 133–148, Springer, 1998.
- 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.
- de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, I.: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF, *Quarterly Journal of the Royal Meteorological Society*, 139, 1199–1213, 2013.
- 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, *Natural hazards*, 65, 1481–1506, 2013.
- 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, *Natural hazards*, 65, 135–165, 2013.
- Dharssi, I., Bovis, K. J., Macpherson, B., and Jones, C. P.: Operational assimilation of ASCAT surface soil wetness at the Met Office, *Hydrology and Earth System Sciences*, 15, 2729–2746, 2011.
- Di Naccio, D., Boncio, P., Brozzetti, F., Pazzaglia, F. J., and Lavecchia, G.: Morphotectonic analysis of the Lunigiana and Garfagna

- nana grabens (northern Apennines, Italy): Implications for active normal faulting, *Geomorphology*, 201, 293–311, 2013.
- Díaz-Uriarte, R. and De Andres, S. A.: Gene selection and classification of microarray data using random forest, *BMC bioinformatics*, 7, 3, 2006.
- D'Odorico, P. and Fagherazzi, S.: A probabilistic model of rainfall-triggered shallow landslides in hollows: A long-term analysis, *Water Resources Research*, 39, 2003.
- 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), *Engineering Geology*, 63, 277–289, 2002.
- Ermini, L., Catani, F., and Casagli, N.: Artificial neural networks applied to landslide susceptibility assessment, *Geomorphology*, 66, 327–343, 2005.
- Falaschi, F., Giacomelli, F., Federici, P., Puccinelli, A., Avanzi, G., Pochini, A., and Ribolini, A.: Logistic regression versus artificial neural networks: landslide susceptibility evaluation in a sample area of the Serchio River valley, Italy, *Natural hazards*, 50, 551–569, 2009.
- Federici, P. R., Puccinelli, A., Cantarelli, E., Casarosa, N., D'Amato Avanzi, G., Falaschi, F., Gianneccchini, R., Pochini, A., Ribolini, A., Bottai, M., et al.: Multidisciplinary investigations in evaluating landslide susceptibility—an example in the Serchio River valley (Italy), *Quaternary international*, 171, 52–63, 2007.
- Felicesimo, 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.
- 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, *Atmospheric Research*, 138, 13–29, 2014.
- Floris, M., Iafelice, M., Squarizoni, C., Zorzi, L., De Agostini, A., and Genevois, R.: Using online databases for landslide susceptibility assessment: an example from the Veneto Region (northeastern Italy), *Natural Hazards and Earth System Science*, 11, 1915–1925, 2011.
- Fowler, H., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, *International Journal of Climatology*, 27, 1547–1578, 2007.
- Fratini, P., Crosta, G., and Carrara, A.: Techniques for evaluating the performance of landslide susceptibility models, *Engineering geology*, 111, 62–72, 2010.
- Galve, J. P., Cevasco, A., Brandolini, P., and Soldati, M.: Assessment of shallow landslide risk mitigation measures based on land use planning through probabilistic modelling, *Landslides*, pp. 1–14, 2014.
- Gao, X., Huete, A. R., Ni, W., and Miura, T.: Optical–biophysical relationships of vegetation spectra without background contamination, *Remote Sensing of Environment*, 74, 609–620, 2000.
- 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 Sensing Letters*, 1, 45–54, 2010.
- Gianneccchini, R.: Rainfall triggering soil slips in the southern Apuan Alps (Tuscany, Italy), *Advances in Geosciences*, 2, 21–24, 2005.
- Gianneccchini, R.: Relationship between rainfall and shallow landslides in the southern Apuan Alps (Italy), *Natural Hazards & Earth System Sciences*, 6, 2006.
- Gianneccchini, R., Naldini, D., D'Amato Avanzi, G., and Puccinelli, A.: Modelling of the initiation of rainfall-induced debris flows in the Cardoso basin (Apuan Alps, Italy), *Quaternary international*, 171, 108–117, 2007.
- Gianneccchini, R., Galanti, Y., and D'Amato Avanzi, G.: Critical rainfall thresholds for triggering shallow landslides in the Serchio River Valley (Tuscany, Italy), *Natural Hazards and Earth System Science*, 12, 829–842, 2012.
- Gislason, P. O., Benediktsson, J. A., and Sveinsson, J. R.: Random forests for land cover classification, *Pattern Recognition Letters*, 27, 294–300, 2006.
- Glade, T., Crozier, M., and Smith, P.: Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical “Antecedent Daily Rainfall Model”, *Pure and Applied Geophysics*, 157, 1059–1079, 2000.
- Goetz, J. N., Guthrie, R. H., and Brenning, A.: Integrating physical and empirical landslide susceptibility models using generalized additive models, *Geomorphology*, 129, 376–386, 2011.
- 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.
- Guzzetti, F., Aleotti, P., Malamud, B., and Turcotte, D.: Comparison of three landslide event inventories in central and northern Italy, in: *Proceedings of the 4th EGS Plinius Conference*, Spain: Universitat de les Illes Balears, 2002.
- Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., and Ardizzone, F.: Landslide hazard assessment in the Staffora basin, northern Italian Apennines, *Geomorphology*, 72, 272–299, 2005a.
- Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., and Ardizzone, F.: Probabilistic landslide hazard assessment at the basin scale, *Geomorphology*, 72, 272–299, 2005b.
- Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, *Meteorology and atmospheric physics*, 98, 239–267, 2007.
- 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-Science Reviews*, 112, 42–66, 2012.
- Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes., *Monthly Weather Review*, 134, 2006.
- 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 sensing of environment*, 83, 195–213, 2002.
- 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., et al.: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP, *Weather & Forecasting*, 23, 2008.
- Klemp, J., Skamarock, W., and Dudhia, J.: Conservative split-explicit time integration methods for the compressible nonhydrostatic equations., *Monthly Weather Review*, 135, 2007.

- Liaw, A. and Wiener, M.: Classification and Regression by random-Forest, *R news*, 2, 18–22, 2002.
- Malamud, B. D., Turcotte, D. L., Guzzetti, F., and Reichenbach, P.: Landslide inventories and their statistical properties, *Earth Surface Processes and Landforms*, 29, 687–711, 2004.
- Mancini, F., Ceppi, C., and Ritrovato, G.: GIS and statistical analysis for landslide susceptibility mapping in the Daunia area, Italy, *Natural Hazards and Earth System Science*, 10, 1851–1864, 2010.
- Mărgărint, M., Grozavu, A., and Patriche, C.: Assessing the spatial variability of weights of landslide causal factors in different regions from Romania using logistic regression, *Natural Hazards and Earth System Sciences Discussions*, 1, 1749–1774, 2013.
- Marjanovic, M., Kovacevic, M., Bajat, B., Mihalic, S., and Abolmasov, B.: Landslide assessment of the Starca basin (Croatia) using machine learning algorithms, *Acta Geotechnica Slovenica*, 8, 45–55, 2011.
- Marmion, M., Hjort, J., Thuiller, W., and Luoto, M.: A comparison of predictive methods in modelling the distribution of periglacial landforms in Finnish Lapland, *Earth Surface Processes and Landforms*, 33, 2241–2254, 2008.
- McCullagh, P.: Generalized linear models, *European Journal of Operational Research*, 16, 285–292, 1984.
- McCullagh, P. and Nelder, J.: General linear models, Chapman and Halls, London, 1989.
- Mercogliano, P., Casagli, N., Catani, F., Damiano, E., Olivares, L., Picarelli, L., Rossi, G., Schiano, P., Segoni, S., Sikorski, B., et al.: Short Term Weather Forecasting for Shallow Landslide Prediction, in: *Landslide Science and Practice*, pp. 121–129, Springer, 2013a.
- Mercogliano, P., Segoni, S., Rossi, G., Sikorski, B., Tofani, V., Schiano, P., Catani, F., Casagli, N., and Tiranti, D.: A prototype forecasting chain for rainfall induced shallow landslides, *Natural Hazards & Earth System Sciences*, 13, 2013b.
- 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 sensing of environment*, 115, 1743–1757, 2011.
- Moore, I. D. and Wilson, J. P.: Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation, *Journal of soil and water conservation*, 47, 423–428, 1992.
- Moriondo, M., Stefanini, F., and Bindi, M.: Reproduction of olive tree habitat suitability for global change impact assessment, *Ecological Modelling*, 218, 95–109, 2008.
- Mossa, S., Capolongo, D., Pennetta, L., and Wasowski, J.: A GIS-based assessment of landsliding in the Daunia Apennines, southern Italy, *Polish Geological Institute Special Papers*, 20, 86–91, 2005.
- Nefeslioglu, H. A., Duman, T. Y., and Durmaz, S.: Landslide susceptibility mapping for a part of tectonic Kelkit Valley (East-Black Sea region of Turkey), *Geomorphology*, 94, 401–418, 2008.
- 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.
- Nefeslioglu, H. A., San, B., Gokceoglu, C., and Duman, T.: An assessment on the use of Terra ASTER L3A data in landslide susceptibility mapping, *International Journal of Applied Earth Observation and Geoinformation*, 14, 40–60, 2012.
- Oberto, E., Milelli, M., Pasi, F., and Gozzini, B.: Intercomparison of two meteorological limited area models for quantitative precipitation forecast verification, *Nat. Hazards Earth Syst. Sci.*, 12, 591–606, 2012.
- Orlanski, I.: A rational subdivision of scales for atmospheric processes, *Bulletin of the American Meteorological Society*, 56, 527–530, 1975.
- Pal, M.: Random forest classifier for remote sensing classification, *International Journal of Remote Sensing*, 26, 217–222, 2005.
- Parodi, A., Boni, G., Ferraris, L., Siccaldi, F., Pagliara, P., Trovatore, E., Foufoula-Georgiou, E., and Kranzlmüller, D.: The “perfect storm”: From across the Atlantic to the hills of Genoa, *Eos, Transactions American Geophysical Union*, 93, 225–226, 2012.
- Perna, M., Crisci, A., Capecchi, V., Bartolini, G., Betti, G., Piani, F., Gozzini, B., Barsanti, B., Bigio, T., Bonciani, F., et al.: Sensitivity Analysis for Shallow Landsliding Susceptibility Assessment in Northern Tuscany, in: *Engineering Geology for Society and Territory-Volume 2*, pp. 197–200, Springer, 2015.
- Peters, J., Baets, B. D., Verhoest, N. E., Samson, R., Degroove, S., Becker, P. D., and Huybrechts, W.: Random forests as a tool for ecohydrological distribution modelling, *Ecological Modelling*, 207, 304–318, 2007.
- 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.
- 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, *Environmental monitoring and assessment*, 184, 715–727, 2012.
- Ray, R. L. and Jacobs, J. M.: Relationships among remotely sensed soil moisture, precipitation and landslide events, *Natural Hazards*, 43, 211–222, 2007.
- 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 sensing of environment*, 114, 2624–2636, 2010.
- Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, E., Pignone, F., Siccaldi, F., Silvestro, F., Tanelli, S., and Parodi, A.: Extreme Rainfall in the Mediterranean: What Can We Learn from Observations?, *Journal of Hydrometeorology*, 14, 2013.
- Regione Toscana, G. R.: Indirizzi operativi per la gestione organizzativa e funzionale del sistema di allertamento nazionale, statale e regionale per il rischio idrogeologico e idraulico ai fini di protezione civile (in italian), Tech. rep., Regione Toscana, 2006.
- Regione Toscana, S. I. R.: Report evento meteo-idrologico del giorno 25 Ottobre 2011 (in italian), Tech. rep., Centro Funzionale della Regione Toscana, 2011.
- Regione Toscana, S. I. R.: Report eventi meteo-idrologici occorsi nel periodo 5-19 Marzo 2013 sul territorio regionale (in italian), Tech. rep., Centro Funzionale della Regione Toscana, 2013.
- Rosi, A., Segoni, S., Catani, F., and Casagli, N.: Statistical and environmental analyses for the definition of a regional rainfall threshold system for landslide triggering in Tuscany (Italy), *Journal of Geographical Sciences*, 22, 617–629, 2012.
- Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., and Petrucci, S.: Optimal landslide susceptibility zonation based on

- multiple forecasts, *Geomorphology*, 114, 129–142, 2010.
- Sacchi, J. A.: Analisi meteorologica degli eventi alluvionali eccezionali tra l'Alta Toscana e l'estremo Levante ligure dal 2009 al 2011 (in italian), *Atti della Società Toscana di Scienze Naturali, Memorie serie A*, 117–119, 75–88, 2012.
- 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.
- Schmidt, J., Turek, G., Clark, M., Uddstrom, M., and Dymond, J.: Probabilistic forecasting of shallow, rainfall-triggered landslides using real-time numerical weather predictions, *Natural Hazards and Earth System Science*, 8, 349–357, 2008.
- 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, *Monthly Weather Review*, 142, 1525–1541, 2014.
- Schwartz, C. S.: Reproducing the September 2013 Record-Breaking Rainfall over the Colorado Front Range with High-Resolution WRF Forecasts, *Weather and Forecasting*, 29, 393–402, 2014.
- Schwitalla, T. and Wulfmeyer, V.: Radar data assimilation experiments using the IPM WRF Rapid Update Cycle, *Meteorologische Zeitschrift*, 2014.
- Segoni, S., Leoni, L., Benedetti, A., Catani, F., Righini, G., Falorni, G., Gabellani, S., Rudari, R., Silvestro, F., and Rebori, N.: Towards a definition of a real-time forecasting network for rainfall induced shallow landslides., *Natural Hazards & Earth System Sciences*, 9, 2009.
- Segoni, S., Battistini, A., Rossi, G., Rosi, A., Lagomarsino, D., Catani, F., Moretti, S., and Casagli, N.: Technical Note: An operational landslide early warning system at regional scale based on space-time variable rainfall thresholds, *Natural Hazards and Earth System Sciences Discussions*, 2, 6599–6622, 2014a.
- Segoni, S., Rosi, A., Rossi, G., Catani, F., and Casagli, N.: Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional scale warning systems, *Natural Hazards and Earth System Sciences Discussions*, 2, 2185–2213, 2014b.
- Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, *Journal of Computational Physics*, 227, 3465–3485, 2008.
- 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, 2005.
- Stumpf, A. and Kerle, N.: Combining Random Forests and object-oriented analysis for landslide mapping from very high resolution imagery, *Procedia Environmental Sciences*, 3, 123–129, 2011a.
- Stumpf, A. and Kerle, N.: Object-oriented mapping of landslides using Random Forests, *Remote Sensing of Environment*, 115, 2564–2577, 2011b.
- 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., *Monthly Weather Review*, 136, 2008.
- Varnes, D. J., the International Association Engineering Geology Commission on Landslides, and on Slopes, O. M. M.: *Landslide hazard zonation: a review of principles and practice*, 1984.
- Vorpahl, P., Elsenbeer, H., Märker, M., and Schröder, B.: How can statistical models help to determine driving factors of landslides?, *Ecological Modelling*, 239, 27–39, 2012.
- Wilks, D. S.: *Statistical methods in the atmospheric sciences*, vol. 100, Academic press, 2011.
- Yilmaz, I.: Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine, *Environmental Earth Sciences*, 61, 821–836, 2010.

~~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.~~

~~RF's variable importance for classification of landslide occurrence in the 25OCT2011 study case (top) and in the 18MAR2013 study case (bottom). The meaning of the variables on the y-axis is reported in table 1, while extended descriptions are reported in sections 2.2 and 2.3.~~

~~Basic options of the numerical WRF simulations~~
Variable
Value projection Lambertrows×columns 440×400vertical levels 35horizontal 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)

~~Statistics of the observed and modelled 24-hour 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 18.2 43.8 59.3 70.4 90.7 218.9 (whole area)
 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

~~Statistics of the observed and modelled 1-hour 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

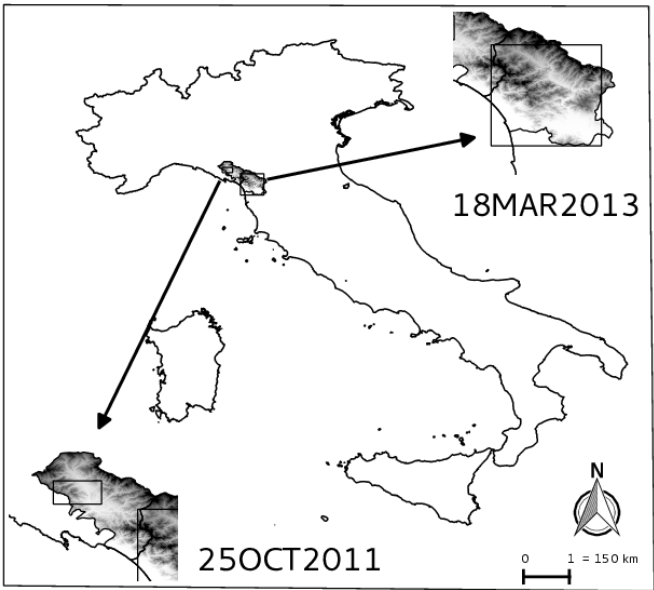


Fig. 1. Italy domain with the two areas of interest: in the inset figure on the left side, it is shown the Lunigiana area belonging to the administrative province of Massa-Carrara. This area was interested by the event 25OCT2011. In the inset figure on the right, it is shown the Garfagnana area belonging 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 black rectangles) where the statistical models were implemented and tested, while, Outer black rectangle represents the cross-signs represent domain of the observed landslides WRF simulations.

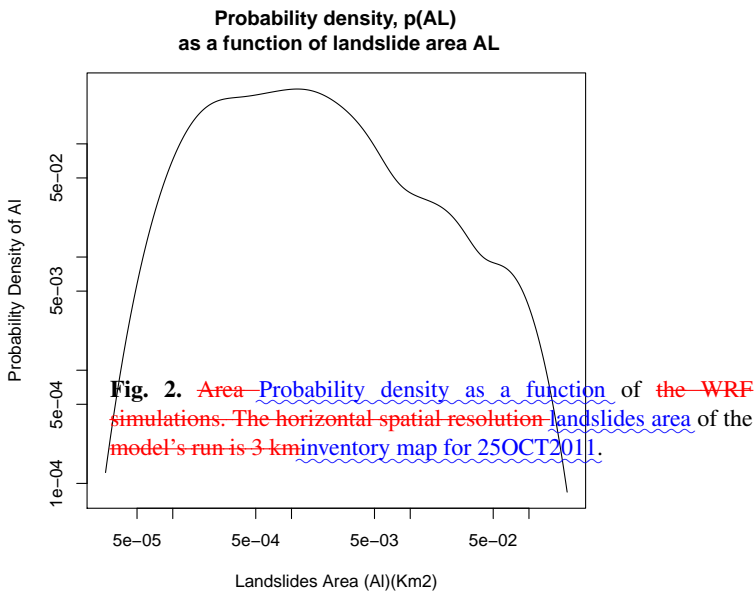


Table 1. List of the predictors taken into account in both GLM model and RF classifier for landslide hazard mapping. Using the RF's variable importance parameter, predictors relevant for the 25OCT2011 rainfall event have a ★ symbol, whereas those relevant for the 18MAR2013 event have a ● symbol.

Geomorphology-related predictors		
Variable/Code	Short description	Unit of measure
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	convexity Concavity/convexity of the land perpendicular to aspect	h/m
prof_curv_geo/PRFC ★	convexity Concavity/convexity of the land parallel to aspect	h/m
Hydrology-related predictors		
Variable/Code	Short description	Unit of measure
conv_index_geo/COVI	convergence 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		
Variable/Code	Short description	Unit of measure
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		
Variable/Code	Short description	Unit of measure
p12h_100_clima/R12 ★	Rainfall 12 hours duration and 100 years return period	mm
p24h_100_clima/R24 ●	Rainfall 24 hours duration and 100 years return period	mm
p48h_100_clima/R48 ★	Rainfall 48 hours duration and 100 years return period	mm
p96h_100_clima/R96 ★	Rainfall 96 hours duration and 100 years return period	mm
Other predictors		
Variable/Code	Short description	Unit of measure
corine_c_landscape/COR ●	Corine Land Use	categorical/8 classes for 25OCT2011 categorical/10 classes for 18MAR2013
evi_media_land/EVI ●	Vegetation Index	—
NWP predictors		
Variable/Code	Short description	Unit of measure
farw3km SOILW0_10cm ★●	Soil moisture 0-10 cm below ground layer	m ³ /m ³
farw3km SOILW10_40cm ★	Soil moisture 10-40 cm below ground layer	m ³ /m ³
farw3km SOILW40_100cm ★	Soil moisture 40-100 cm below ground layer	m ³ /m ³
farw3km SOILW100_200cm ★	Soil moisture 100-200 cm below ground layer	m ³ /m ³
farw3km APCP	Precipitation Total precipitation amounts	mm/24h
farw3km APCP.RI ★	Precipitation Hourly precipitation intensity	mm/h

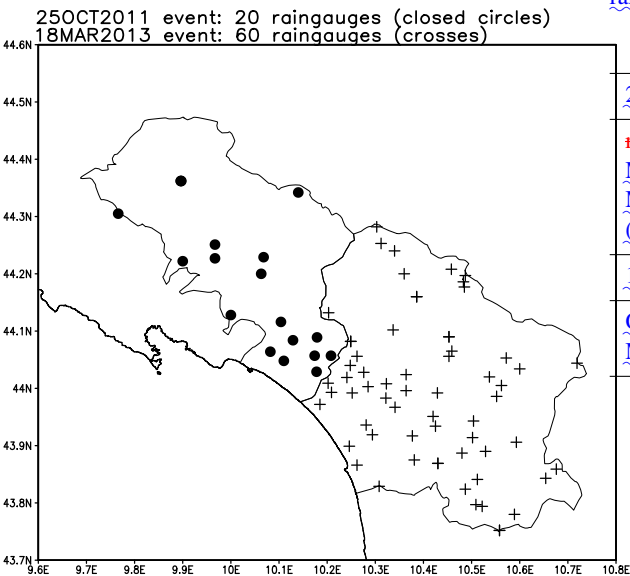


Fig. 3. Locations of the rain-gauges (closed circles) for the two study cases: 25OCT2011 in gray, (closed circles) and 18MAR2013 in black. Crosses indicate the location of the landslides for the two study cases: 25OCT2011 in gray, 18MAR2013 in black. (crosses)

Table 3. Descriptive statistics of the observed and modeled 24-hour rainfall amounts for both study cases (unit is mm)

25OCT2011	Min	1st Qu	Max
Observed data	mean hourly rainfall intensity-21	68	100
Modeled data	40	50	61
Modeled data (whole area)	18	43	59
18MAR2013	Min	1st Qu	Max
Observed data	0	49	64
Modeled data	13	21	38

Table 4. Descriptive statistics of the observed and modeled 1-hour rainfall values for 25OCT2011 (unit is mm)

25OCT2011	Min	1st Qu	Median	Mean	3rd Qu	Max
Observed data	6	26	28	31	33	67
Modeled data	10	16	18	21	26	41

Table 2. List Basic options of the ten NWP input predictors classified as “important” by the RF classifier for the 25OCT2011 event, numerical WRF simulations

Coded-name	Variable	Description	Value
projection		Importance (ranking)	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)	

Table 5. AUC values achieved by the GLM and RF models for both study cases. Summary statistics are computed considering the 100 runs performed (δ stands for standard deviation)

25OCT2011	Min	Mean	Max	δ
GLM	0.76	0.85	0.91	0.031
RF	0.84	0.91	0.96	0.022
18MAR2013	Min	Mean	Max	δ
GLM	0.53	0.69	0.82	0.054
RF	0.59	0.71	0.83	0.047

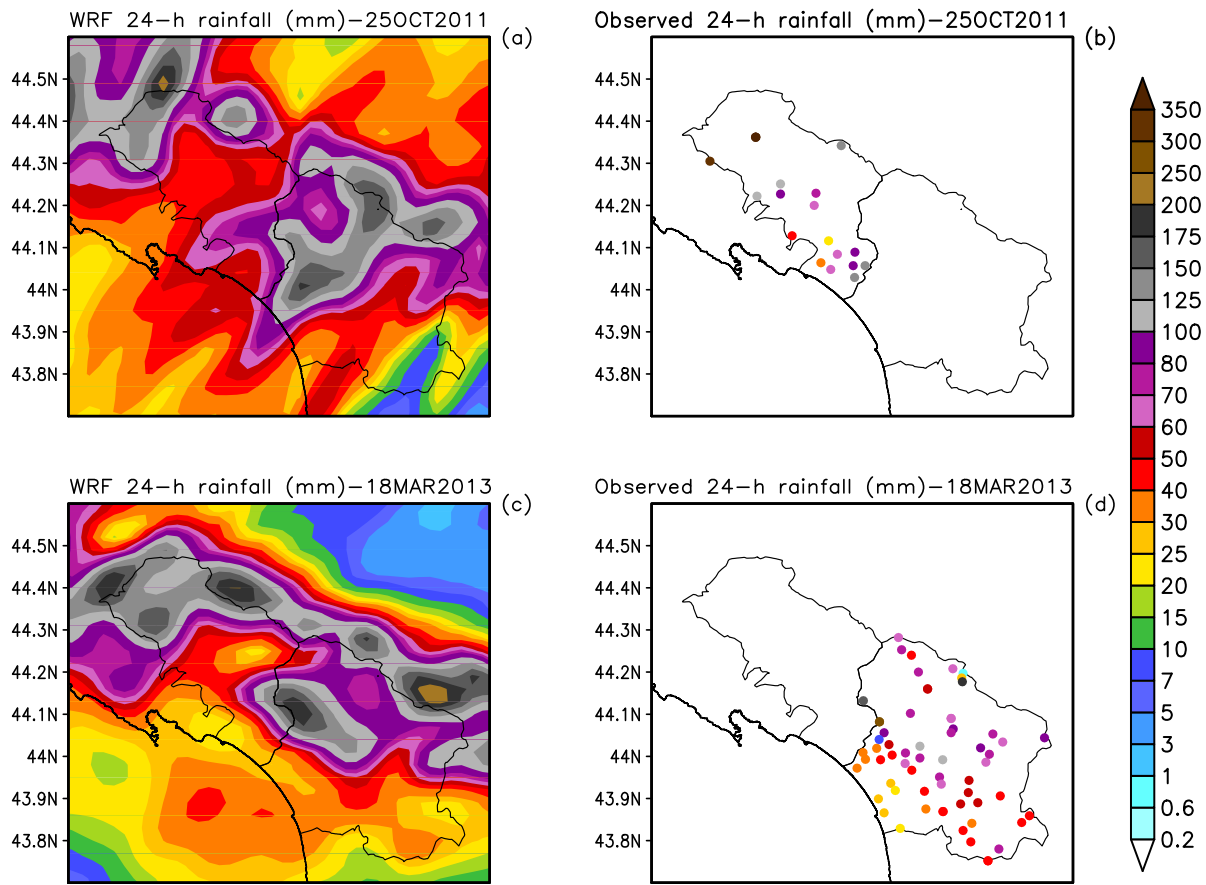


Fig. 4. Panels (a) and (b): ~~modelled~~modeled and observed precipitation in mm accumulated in the 24h period starting at 00:00 UTC of 25th October 2011. Panels (c) and (d): ~~modelled~~modeled and observed precipitation in mm accumulated in the 24h period starting at 00:00 UTC of 18th March 2013.

Table 6. List of predictors classified as “important” by the RF algorithm for 25OCT2011 and 18MAR2013. The average position and the percentage of occurrences in the first ten (five) positions in 25OCT2011 (18MAR2013) ranking are computed over the 100 runs of RF.

Variable/Code	25OCT2011 Average rank	Percentage in the first ten
Geomorphology-related predictors		
dem_30m_topo/DEM	4.5	
h_channel_geo/AaCN	7.1	
aspect_topo/ASP	7.6	
slope_topo/SLP	8.6	
ls_factor_geo/LSF	7.7	
plan_curv_geo/PLAC	8.0	
prof_curv_geo/PRFC		
Hydrology-related predictors		
conv_index_geo/COVI		maximum soil moisture (100–200 cm below ground)
tc_geo/ToC		
TWI_geo/TWI	9.6	
r_max_d_aste_fluvi_geo/DfCN		maximum soil moisture (40–100 cm below ground)
Geology-related predictors		
d_lineamenti_tettonici_geo/DfTF	3.7	
litho_geo_int/BLT	8	
r_RI_max_a_distacco_geo/LMS		max hourly rainfall intensity
r_pedopaesaggi_int/SKST		
quaternario_frane_geo/LaSD		
assetto_geo/SSS		
Climate-related predictors		
p12h_100_clima/R12	7.8	
p24h_100_clima/R24	7.3	
p48h_100_clima/R48		
p96h_100_clima/R96		
Other predictors		
corine_c_landscape/COR		
evi_media_land/EVI	3.4	
NWP predictors		
arw3km_SOILW0_10cm_max(max)		maximum soil moisture (0–10 cm below ground) 8.3
arw3km_SOILW0_10cm_mean	159.5	
r_arw3km_SOILW10_40cm_max(max)		maximum soil moisture (10–40 cm below ground) 7.8
arw3km_SOILW10_40cm_mean	169.2	
r_arw3km_SOILW40_100cm_mean-mean soil moisture (40–100 cm below ground)(max)	176.2	
r_meanarw3km_SOILW40_100cm_mean		mean soil moisture (10–40 cm below ground) 9.4
arw3km_SOILW100_200cm_max	19	
r_meanarw3km_SOILW100_200cm_mean		mean soil moisture (0–10 cm below ground)
arw3km_APCP_total	202.6	
r_mean_arw3km_APCP_RI(max)		mean soil moisture (100–200 cm below ground) 7.2
arw3km_APCP_RI_mean	222.6	

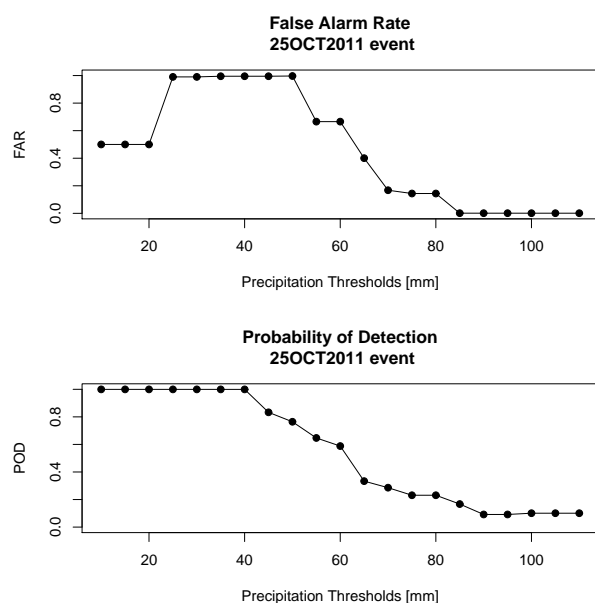


Fig. 5. Study case 25OCT2011: False Alarm Rate (top) and Probability of Detection (bottom) skills obtained when validating the accumulated 24-h rainfall predicted by the WRF model with the rainfall data observed at rain-gauge locations displayed in figure 3.

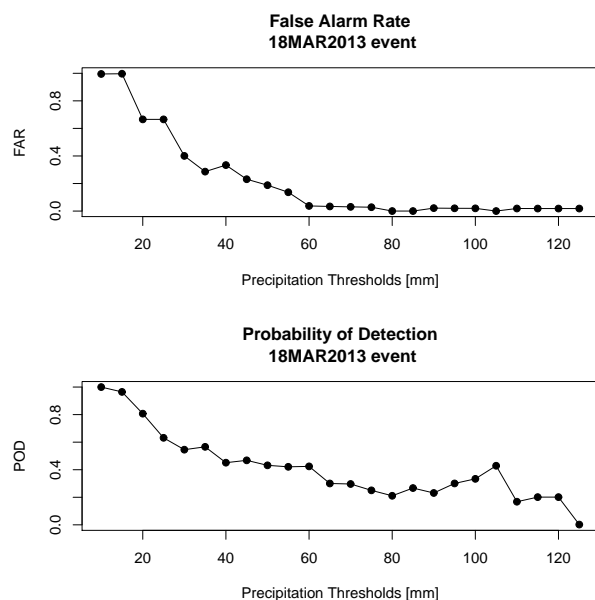


Fig. 6. Study case 18MAR2013: False Alarm Rate (top) and Probability of Detection (bottom) skills obtained when validating the accumulated 24-h rainfall predicted by the WRF model with the rainfall data observed at rain-gauge locations displayed in figure 3.

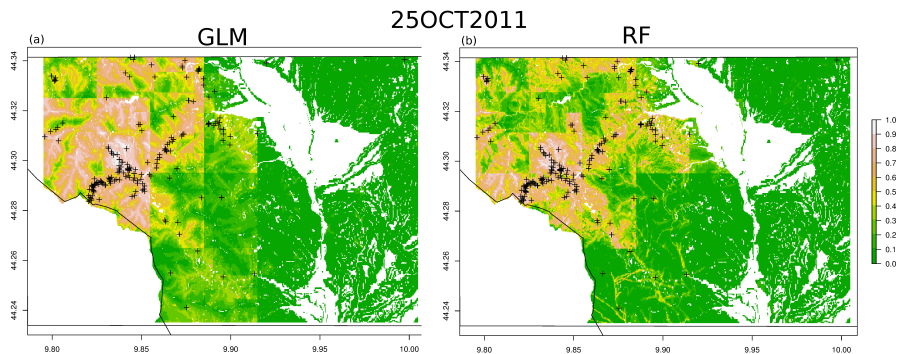


Fig. 7. 25OCT2011 landslide hazard maps and corresponding Receiver Operating Characteristics Curves (ROC) with the values of the underlying area (AUC). Results from the GLM model (a) and from the RF classifier (b). Crosses points are the event inventory maps produced by field surveys a few days after the rainfall event.

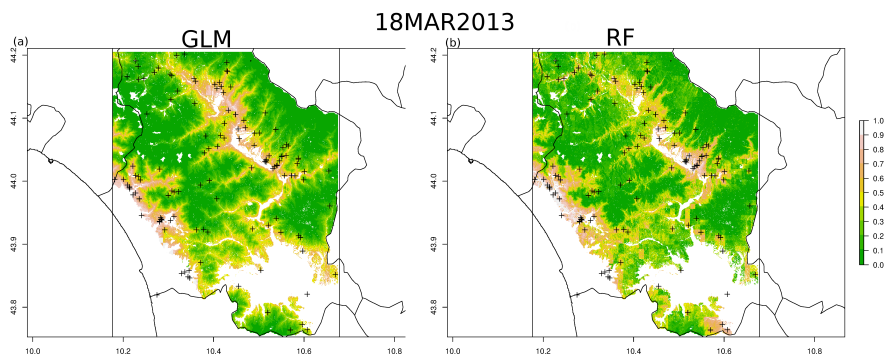


Fig. 8. 18MAR2013 landslide hazard maps and corresponding Receiver Operating Characteristics Curves (ROC) with the values of the underlying area (AUC). Results from the GLM model (a) and from the RF classifier (b). Crosses points are the event inventory maps produced by field surveys a few days after the rainfall event.