Chrono-validation of near-real-time landslide susceptibility models via plug-in statistical simulations

Luigi Lombardo\(^a,\)*, Hakan Tanyas\(^b,\)\(^c\)

\(^a\) University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), PO Box 217, Enschede, AE 7500, Netherlands
\(^b\) Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, United States
\(^c\) USRA, Universities Space Research Association, Columbia, MD, United States

ARTICLE INFO

Keywords:
Temporal validation
Landslide susceptibility
Statistical simulations
INLA
Slope unit

ABSTRACT

The idea behind any validation scheme in landslide susceptibility studies is to test whether a model calibrated on a certain data can predict an unknown dataset of the same nature (landslide presences/absences and covariates). Almost the entirety of landslide susceptibility studies are validated by subsetting a single dataset into a training and test sets. This dataset usually corresponds either to event-specific or to historical inventories. Very rarely, a multi-temporal inventory is available and, in the few cases where this condition is met, the validation practices involve training a model on a specific landslide inventory, deriving a single predictive equation and validating it on a subsequent landslide inventory. This commonly leads landslide predictive studies, even those with a strong statistical rigor, to neglect the uncertainty estimation in their modeling scheme.

In statistics, validation can also be performed via statistical simulations. This means that after fitting a given model, one can generate any number of predictive functions and test their predictive skills on any type and number of unknown datasets. In this work, we take a similar direction and we apply it to model and validate three separate co-seismic inventories, including an uncertainty estimation phase. We mapped these inventories within the same area in Indonesia, for three earthquakes occurred in 2012, 2017 and 2018. Specifically, we build three event-specific Bayesian Generalize Additive Models of the binomial family. From each model we then simulate 1000 predictive realizations over the remaining two inventories, by using a plug-in scheme where all the morphometric covariates are kept fixed and only the ground motion is replaced according to the prediction target. By doing so, we introduce a new analytical tool for near-real-time landslide predictive purposes, which is able to produce a probabilistic model which stands in between the definitions of susceptibility and hazard. In fact, our model is able to accurately estimate “where” and “when” – although not “how frequently” – landslide have occurred by featuring the multitemporal information of the trigger.

In our findings, the simulations are quite similar to the fitted models; and the nine combinations we analyse produce excellent performance. This result confirms the assumption that “the past is the key to the future”, as we show that the relative contribution of each variable and their interactions in each probabilistic model remains practically the same across temporal replicates. This information is not trivial because it supports the routines implemented in global near-real-time applications.

1. Introduction

Landslide susceptibility is defined as the probability of a population of landslides to occur over space on the basis of a set of predisposing factors (Brabb, 1984). Therefore, susceptibility usually neglects the contribution to the landslide initiation brought by the main trigger. And, it does not feature the temporal component which controls the landslide onset. This properties are instead part of the hazard definition, together with the magnitude of the landslide event (Guzzetti et al., 1999, 2005).

Validation is the main requirement of any predictive model (e.g., Beguera, 2006; Chung and Fabbri, 2003). In landslide susceptibility research, model validation is conducted in different ways for models based on historical inventories (e.g., Frattini et al., 2010; Castro Camilo et al., 2017) and event-specific inventories (e.g., Lee et al., 2008; Nowicki et al., 2014).

The former corresponds to landslide inventories spanning over a long or ill-defined temporal window. In this case, susceptibility models
are built irrespective of the specific trigger. And, the predisposing factors’ effects are assumed to be stationary over time. As a result, a susceptibility model developed on the basis of historical landslide inventories is, by definition, a time-invariant product representing geomorphologically landslide-prone hillslopes (Varnes, 1984).

As for the latter, related to event-specific inventories, these are mapped right after a major event (e.g., earthquake, rainfall or snowmelt), which triggered widespread slope failures (Guzzetti et al., 2012). As a result, these inventories are associated with a triggering event of known date. And, the component expressed by the trigger is assumed to vary over time. A common approach to make use of such inventories is to build a predictive model over a prior event and then by forecasting a new one by replacing (or plugging-in) in the predictive equation the proxy used to express the trigger (e.g., Nowicki et al., 2014; Nowicki Jessee et al., 2018). Therefore, the applicability of such models only exists if there is another event that has already occurred or assumed to occur in the future with a known extent and intensity. However, apart from the replacement of the data expressing the trigger, in analogy to the historical inventory case, the effects of the covariates (both predisposing and triggering factors) and their interaction are also assumed to be time-invariant (e.g., Tanyas et al., 2019).

In the last two decades, the plug-in structure illustrated above for event-specific cases have been used to build several near-real-time models, for both earthquake-triggered (e.g., Godt et al., 2008; Nowicki Jessee et al., 2018; Tanyas et al., 2019) and rainfall-triggered (e.g., Kirschbaum and Stanley, 2018) landslides. For instance, the model developed by Nowicki Jessee et al. (2018) is part of a recent near-real-time products (i.e., the Ground Failure model) of the U.S. Geological Survey. The Ground Failure model generates probabilistic maps in near-real-time with earthquake occurrences (Allstadt et al., 2018). Specifically, the estimated ground motion of newly occurred earthquake provided by the USGS ShakeMap system (Garcia et al., 2012) is plugged-in in a reference regression equation.

We stress to the reader that these models contain more information than a traditional susceptibility because they feature the trigger signal and because they essentially predict landslide occurrences for a specific time, which are both part of the landslide hazard definition (Guzzetti et al., 1999). However, they do not contain information on the landslide event magnitude and temporal recurrence, which are also fundamental components of the hazard. As a result, these near-real-time predictive models do not have a specific definition for they actually stand in between the susceptibility and the hazard concepts, though for conservativeness they are often still referred to as susceptibility (e.g., Parker, 2013). In the rest of the manuscript, we will also use the susceptibility term to describe the probability of landslide occurrence in the context of near-real-time applications rather than long-term territorial planning. Whenever this will be the case, we will add a prefix (acronym for near-real-time) to the susceptibility, to avoid any confusion in terminology.

Near-real-time models, such as the one presented in Nowicki Jessee et al. (2018) and others mentioned above, are validated on the assumption that the multivariate interactions and their contributions to the estimated NRT-susceptibility are stationary over space and time. In other words, these models are built based on the assumption that the past is the key to the future (Carrara et al., 1995). However, the validation is generally performed by using landslide-event inventories collected in different locations (spatially) rather than using repeated landslide-event inventories collected in the same area (temporally). Thus, validation is hardly performed by testing near-real-time models over time-series of event-specific inventories.

Notably, this is also valid outside near-real-time applications and generally in landslide science, time-dependent validation (or chronovalidation) can seldom be found (Reichenbach et al., 2018). The reason behind this is due to: i) the complexity of accessing multi-temporal scenes of a given study area; ii) the complexity of a standardized mapping procedure (Guzzetti et al., 2012). Nevertheless, valid examples can be found in the landslide literature (e.g., Remondo et al., 2003; Guzzetti et al., 2005; Meusburger and Alewell, 2009; Van Den Eeckhaut et al., 2010; Cama et al., 2015) where the authors test temporal validation routines, albeit the temporal dependence between unstable slopes has been neglected. In fact, more recently, the work proposed by Samia et al. (2017a, 2017b) has postulated that there is an effect between consecutive failure mechanism at the same slope, which should be accounted for in multi-temporal landslide susceptibility models. And, this concept has been further developed by Lombardo et al., 2020a, where the authors account for space-time dependencies in landslide intensity models (Lombardo et al., 2018).

Nevertheless, these models are generally built through historical landslide inventories, therefore, the direct signal or effect of the trigger is not included in their analytical routines. To our knowledge, there is only one study where chrono-validation has been applied to landslide-event inventories (Chang et al., 2014) featuring the signal of the trigger. Specifically, Chang et al. (2014) use nine typhoon-triggered landslide inventories; however, they note that, for each inventory they use, they map landslides using pre-event images acquired approximately seven months before the event and post-event images collected one month after the event. Therefore, the inventories may include some landslides triggered by agents other the identified rainfall proxies. In this work, we take inspiration from the chrono-validation procedure mentioned above (in analogy to near-real-time applications) and we test it over event-specific co-seismic landslide inventories. We use earthquake-induced landslide inventories for which the intensity and spatial distribution of the trigger has been estimated and made accessible to the public. In doing so, we also combine the plug-in method implemented in near-real-time models and extend it even further. In fact, the general practice is to use a single predictive function, and use it to predict landslides triggered by a newly occurred earthquakes. Conversely, what we propose here is to initially fit a Bayesian version of a binomial Generalized Additive Model (GAM; e.g., Goetz et al., 2011; Lombardo et al., 2020a) which provides a full distribution for each model component. Then, from each distribution we randomly sample and solve 1000 simulated predictive functions to include an uncertainty estimation step in our model, which is traditionally neglected in most landslide predictive studies.

We test this framework in Indonesia where the area close to the Palu region has been recently affected by three earthquakes in 2012, 2017 and 2018, triggering widespread landslides in each case. Specifically, from three separate binomial GAM models built for each earthquake-induced landslide inventory, we generate 1000 NRT-susceptibility realizations to validate the two remaining inventories. This procedure features a plug-in step where we substitute the component related to the ground motion with the ground motion of the simulation target. As a result, our modeling protocol sheds light on the uncertainty affecting the estimates. Notably, these uncertainties can come from various sources, from the resolution of the predisposing and triggering factors, to errors in the mapping procedure and the stochasticity in the model itself. The uncertainty we estimate is an aggregate of these three sources and cannot be deconvolved in their respective origins. This topic will be discussed in Section 5.2.

The article is organized as follows: Section 2 introduces the three co-seismic inventories we mapped within the study area. Section 3 describes the mapping unit we chose and the algorithmic architecture behind our model and associated simulations. Section 4 presents our results which we discuss in Section 5. Ultimately, Section 6 provides an overview of our work and its implications for near-real-time seismically-induced landslide prediction.

2. Study area and co-seismic landslide inventories

We mapped landslides triggered by three different earthquakes occurred in a portion of the Sulawesi Island of Indonesia (Fig. 1). We selected the mapping window as the largest area contextually covered
by cloud-free PlanetScope (3–5 m) and Rapid Eye (5 m) images acquired from Planet Labs (PlanetTeam, 2017) across the three earthquakes.

The first earthquake we studied is the Sulawesi earthquake (Mw = 6.3, 18th August 2012), which occurred on a strike-slip fault. For this earthquake, we examined five different satellite images (acquired on 17th and 21st August 2012, 4th September 2012, 4th February 2013 and 20th August 2013). From these scenes, we mapped 520 landslides with a total surface of 1.3 km² (Fig. 2). However, the first cloud-free image covering the whole study area and showing post-seismic conditions was the last one in August 2013 (roughly one year after the earthquake occurrence). The other images provided us partial information regarding the triggered landslides. Based on our investigation on the available satellite images, we interpreted the majority of landslides to be co-seismic. However, we cannot fully reject the hypothesis that some post-seismic landslides could have been rainfall-triggered, due to the time span between the earthquake occurrence and a clear cloud-free scene.

The second event is the Kasiguncu (Mw = 6.6, 29th May 2017) earthquake which occurred on a normal fault. Notably, this is a rare observation for a landslide event since most of the available earthquake-induced landslide inventories are associated with either strike-slip or thrust faults (Tanyas et al., 2017). We examined two pre-seismic (acquired on 11th and 25th May 2017) and two post-seismic (acquired on 7th and 26th June 2017) images close to the earthquake date of occurrence, without cloud disturbances. As a result, we clearly mapped 384 co-seismic landslides with a total surface of 0.5 km² (Fig. 2).

The third case we investigated is the Palu earthquake (Mw = 7.5, 28th September 2018), which is the biggest landslide events observed in this region, and which occurred on a strike-slip fault (Watkinson and Hall, 2019). The Palu earthquake caused not only hillslope failures but also lateral spreading of unconsolidated sediments due to seismically-induced liquefaction in nearly flat areas along rupturing zone (Bradley et al., 2019). However, in our study area, we only observed hillslope failures as a consequence of hilly topography. For the Palu case, we examined five cloud-free satellite images close to the earthquake date of occurrence (acquired on 21st and 26th September 2018, 2nd, 5th and 22nd October 2018) and precisely mapped 722 co-seismic landslides with a total surface of 2.5 km² (Fig. 2).

We interpreted the vast majority of landslides involved in the three co-seismic failures to be shallow rock and debris slides. Our interpretation originates from a favorable mapping situation where the extremely dense vegetation in the area clearly defined the landslide boundaries, making the interpretation relatively easy on high resolution imagery (3-5 m).

3. Modeling strategy

3.1. Mapping units and presence/absence assignment

Any landslide susceptibility model is strictly dependent on the mapping unit one chooses (Rossi and Reichenbach, 2016). A mapping unit essentially subdivides any study area into geographical objects upon which the model ultimately assigns the susceptibility estimates. Among the mapping units proposed, grid-cells and Slope Units (SUs) are the most common in the literature (Reichenbach et al., 2018). Here, we opt for a Slope Unit partition for this unit reflects near-homogeneous slope responses to landslide occurrences (e.g., Alvioli et al., 2016) rather than the point-specific information conveyed by the grid-cell.

We compute the SUs by using the r.slopeunits software made available by (Alvioli et al., 2016). In particular, we used the r.slopeunits version which is able to mask out flat topographies before the actual generation of the final polygonal structure. Our r.slopeunits parameterization features a flow accumulation threshold of 800,000 m², a minimum slope unit area of 50,000 m² and a circular variance of 0.35. As a result, we obtained a SU partition where the mean SU extent is 0.40 km² and the standard deviation is 0.35 km², which implies that the local topography is relatively rough.

Ultimately, we assign a presence status to each SU intersecting a landslide polygon and an absence condition to the complementary case for each co-seismic inventory, respectively. As a result, from the initial number of 520, 384 and 722 landslides, the aggregated information of slope instability per slope unit has become 185, 138 and 183 for Sulawesi, Kasiguncu and Palu, respectively.

Fig. 1. Overview of the study area. Stars correspond to the epicenters of the three earthquakes. Lines are the Peak Ground Acceleration contours for each of the three cases.
3.2. Covariates

The stability of a single slope is governed by factors that contribute to driving and resisting forces along the sliding plane. At the slope scale, the dynamics that led to a landslide can be reconstructed with relative ease (e.g., Juanget al., 1998; Luet al., 2017) via geotechnical surveys. However, when the objective is to model regional landslide occurrences (Crozier, 2005), the estimation of factors controlling the slope stability condition and their interactions becomes more complex. This complexity is mostly due our inability to systematically survey and reliably estimate geotechnical properties over large regions. For this reason, the geoscientific community adopts proxies of the instability factors. As a result, the selection and interpretation of the best set of variables need to be performed following an initial geomorphological criterion. This is valid both for rainfall and co-seismic landslides. For the latter, among various causative factors, seismic shaking heavily controls the landslide patterns in space (Nowicki et al., 2014). For instance, both landslide area and count density decrease with increasing distance from the rupturing zone (e.g., Gorum et al., 2011; Massey et al., 2018). Therefore, seismic shaking is commonly introduced in co-seismic landslide probabilistic models (especially in NRT-susceptibility) and it is expected to positively contributes to the landslide occurrences.

Following the same logic, we can identify other factors possibly controlling the slope stability conditions. Hillslope steepness is by far the most important predisposing factors (e.g., Reichenbach et al., 2018). Overall, steeper hillslopes can be associated with more susceptible condition for steepness contributes to driving forces (i.e., gravitational forces) by increasing the force component acting towards the downslope direction (Wu and Sidle, 1995; Parise and Jibson, 2000).

Local relief, which may be partially correlated with slope steepness, is another common proxy in landslide susceptibility models (e.g., Reichenbach et al., 2018). It gives the maximum difference in elevation
within a defined neighborhood and can be related to slope instability caused by differences in potential energy across the landscape and/or to tectonic uplift (e.g., Tanyas et al., 2019).

Similar to the local relief, topographic curvature (i.e., profile and plan curvatures) and roughness indicate localized hillslope changes (e.g., Dou et al., 2015; Ohlmacher, 2007; Tanyas et al., 2019) and therefore are commonly used as predisposing factors (e.g., Budimir et al., 2015). For the curvatures, it has been demonstrated that they can control overland flows, therefore contributing to variations in the pore pressure regimes (Ohlmacher, 2007).

Another common predisposing factor is distance to river channels. It is a proxy used to account the external disturbance to the slope stability caused by fluvial undercutting. In fact, in areas close to river channel, common erosional processes may remove material from the two hydrological sides. Therefore, this may lead to the removal of lateral support and to increased shear stresses (e.g., Korup, 2004). Notably, at times earthquakes preferentially trigger landslides at ridge tops because of topographic amplification (Meunier et al., 2008) which would primarily affect areas far from the river network, leading to an apparent inverse relation with this hydrological feature.

To test the statistical significance of the properties mentioned above, we examine a covariate set made of six morphometric and one ground shaking properties: i) Distance to Stream (Dist2St); ii) Slope steepness (Zevenbergen and Thorne, 1987); iii) Profile and iv) Planar Curvatures (PRC and PLC, Heerdegen and Beran, 1982); v) Vector Ruggedness Measure (Roughness, Sappington et al., 2007); vi) Local Relief (LR, Jasiewicz and Stepinski, 2013); vii) Peak Ground Acceleration (PGA, Garca et al., 2012).

We compute Dist2St as the Euclidean distance from a 30 m lattice to the nearest channel. Being the covariates calculated from a 30 m DEM (SRTM, Jarvis et al., 2008), and the SU being on a coarser scale, we further summarized each covariate distribution per SU via its mean and standard deviation (e.g., Amato et al., 2019; Rossi et al., 2010). We do this for every morphometric covariate. The actual representation of the covariate distribution into its mean and standard deviation is suitable if the distribution per SU is Gaussian or near-Gaussian. For this reason, we report in Fig. 3 a summary plot of the covariate distributions per SU. This plots supports the compression of the covariate distribution uniquely via two of its main statistical moments.

Moreover, we do not compress the distribution of the PGA per SU for which the native coarse resolution at a 1-km pixel implies very small to no variations per SU. Notably, the 1-km resolution is the global default resolution at which the ShakeMap system estimates ground motion and makes it available to the community. In reality, the ground motion can varies over very small distances (even tens of meters) but because such variations cannot be captured at the global scale, the 1-km resolution of the ShakeMap products is the most common global representation of the shaking levels experienced in response to an earthquake. Notably, despite all the limitations such a global product may have, the numerical description of the PGA is quite detailed. We show in Fig. 4 that even if the global nature of the data could smooth the PGA over space, the patterns in such a relatively small area are still well captured and different from each other. As a result, the information accessed via the ShapeMap system of the USGS is suitable for our analyses although we recall the reader that a better characterization could be achieved with data coming from local seismic stations.

We stress here that before any modeling steps, we have performed a mean zero unit variance rescaling to each covariate. This procedure ensures that the associated estimates can be compared one another.
3.3. Model fit and simulations

To fit our three reference NRT-susceptibility models for each co-seismic inventory, we opt for a binomial Bayesian Generalized Additive Model (GAM, Goetz et al., 2015; Lombardo et al., 2020b) by using the R package INLA (Bakka et al., 2018). This class of models is an extension of the most common Generalized Linear Model (Lombardo and Mai, 2018) upon which the vast majority of landslide susceptibility studies is traditionally built (e.g., Nefeslioglu et al., 2008; Reichenbach et al., 2018). A GAM allows for the inclusion of linear properties (or fixed effects, Steger et al., 2020) together with any type of non-linear functions (or random effects, Lombardo et al., 2019), these being expressed as categorical, ordinal, or any residual effect at the latent level (Bivand et al., 2015). Here we implement a GAM with all covariates modeled as fixed effects. Only the slope steepness is modeled as an ordinal random effect with an adjacent-class dependency governed by a random walk of the first order (RW1, Lindgren and Rue, 2008). We recall here that most of the literature, whenever the slope or another covariate is reclassified into a finite number of bins, features the use of models that assume each class to be independent (iid) from the others (e.g., Nandi and Shakoor, 2010; Kritikos et al., 2015). This is certainly the way to operate for categorical properties such as Geology or Land Uses. However, in case of slope steepness (in our case), a iid model would neglect the ordinal structure that exists among classes. For instance, classes 0°-5°, 5°-10° and 10°-15° are not sequentially independent one another and informing the model about it via a RW1, leads to better estimates overall. Similar use of adjacent-class models can be found in data mining- (e.g., Conoscenti et al., 2016), as well as in frequentist (e.g., Knevels et al., 2020, see Fig. 6) and Bayesian (e.g., Lombardo et al., 2018, see Fig. 4) GAM-based landslide predictive models.

On the basis of this model structure, we separately fitted three co-seismic inventories for which the morphometric covariates have been kept the same while the ground motion changed for each specific earthquake. Being the three reference fits implemented in a Bayesian framework, each model component is estimated with an associated estimated distribution (see, Bakka et al., 2019). This makes it easy to simulate any number of realizations by sampling at random from each distribution. As a result, we implemented a subsequent step where we simulated from the posterior distribution of each model component. And, for each fit we validate with respect to the other two co-seismic inventories. More specifically, we used the plug-in simulation framework tested for earthquake-induced landslides by Lombardo and Tanyas (2020) and graphically explained in Fig. 6 of Luo et al. (2020). The plug-in phase comes for each simulation before transforming the linear combination of each model component into probability via the logit link. We substitute the PGA of the reference fit with the PGA or the co-seismic inventory we wanted to validate, for a total of six combinations and 1000 simulations for each case.

This procedure ensures that the uncertainty in the estimation of the three reference NRT-susceptibility models (Sulawesi 2012, Kasinguncu 2017 and Palu 2018) propagates to the complementary predicted (chrono-validated) co-seismic NRT-susceptibility estimates.

Fig. 4. PGA maps summary. The 2D-kernel density scatterplots show the variation of a PGA maps against the others, showing that a suitable description of the ground shaking parameter is presented. In fact, the respective PGA ranges are substantially different and able to spatially describe the ground motion of each earthquake.
4. Results

In this section we initially provide an overview of the significant covariates’ effects to interpret the geomorphological reasonability of each reference model (§4.1). Subsequently, we discuss the performance for the three reference GAM fits and for the six chrono-validated simulation batches (§4.2). We then conclude it showing the differences among the nine NRT-susceptibility estimates in map form (§4.3).

4.1. Covariates’ effects

Fig. 5 shows the marginal distribution of the significant fixed effects namely, each global intercept (the $\beta_0$ term appearing in the NRT-susceptibility equation), together with the standard deviations of $\text{Dist2St}$, $\text{Slope}$ and $\text{Roughness}$ within each SU, and the mean $\text{PGA}$.

We show the three global intercepts because they provide an interesting situation to be discussed. The estimation of the intercept is intrinsically linked to the proportion of landslide presences/absences in a given study area (see Supplementary material of Lombardo and Mai, 2018). The relatively small variation in the respective means indicates that almost irrespective of the earthquake magnitude and rupturing type, the landscape has responded with a similar pattern of unstable SUs. Similarly, the overall distribution of the other fixed effects for each reference NRT-susceptibility model appears to be consistent in terms of signs. And, they show limited variations with the exception of $\text{Slope}$ and $\text{PGA}$. The latter cases differ among co-seismic cases only in the amplitude of the regression coefficients, although as mentioned above the sign of the respective means remain the same across models.

To provide an interpretation from a geomorphological standpoint, in Fig. 6 we show the raw data and the relative distribution of each covariate (before rescaling) with respect to stable/unstable SUs.

In Fig. 6, the difference in the medians between stable and unstable SUs is evident for each of the four covariates and for each earthquake. In fact, the median $\text{Dist2St}$, associated to stable slope units is 205, 205 and 208 m for SUlawesi, Kasiguncu and Palu, respectively. The median values for the unstable SUs are instead 275, 317 and 230 m. This is an extremely surprising and interesting result because it implies that the stable and unstable SUs have a clear common behavior across co-seismic events. The same situation occurs for $\text{Slope}$, where the medians are 7.5, 7.6 and 7.5 degrees for stable SUs. And, they are 9.0, 9.3 and 9.0 degrees for the unstable cases. Again, the same situation is repeated for $\text{Roughness}$, with medians equal to 0.007, 0.007 and 0.007 for stable conditions and 0.012, 0.011 and 0.012 for unstable SUs. As for the PGA, the pattern shows variations which are due to the different shaking levels. But, the difference between the stable and unstable medians is respected and clearly shown in each of the three cases.
The three random effects shown in Fig. 7 for the Slope also show a broad agreement. We remind here we modeled Slope as a non-linear effect with adjacent-class dependency. Specifically, the three functions, expressed as the posterior mean and associated 95% credible interval, show similar patterns which clearly overlap between 12 and 25 degrees of steepness and slightly diverge towards the left and right tails of the slope distribution where the uncertainty is larger.

This is a remarkable situation which demonstrates that the effects, or contribution of each model component, for earthquake-induced landslides can be stationary over time, both when covariates are modeled linearly and non-linearly. This is something that was currently and partially demonstrated over space for the vast majority of contributions in the literature so far (e.g., Tanyas et al., 2019) without significant evidences in the temporal case.

4.2. Performance overview

The performance assessment in this work needs to be separated into two types. The first corresponds to a calibration stage where we separately fitted the three reference models to the three co-seismic inventories. And, the second consists of the simulation we have run for each fit, with respect to the complementary two co-seismic scenarios. The overall performance is assessed via Receiver Operating Characteristic curves and their integrated Area Under the Curve (ROC and AUC, respectively; Hosmer and Lemeshow, 2000). For clarity, we remind here the reader that the ROC curves are constructed in a plane defined between the Sensitivity and 1-Specificity. The former is also commonly referred to as True Positive Rate (TPR) and it can be calculated from any confusion matrix – for specific probability intervals – as the ratio between True Positives (TP) over the sum of TP and False Negatives (FN) (see, Rahmati et al., 2019). The latter is also commonly referred to as False Positive Rate (FPR) and it can be calculated from any confusion matrix – for specific probability intervals – as the ratio between False Positives (FP) over the sum of FP and True Negatives (TN) (see, Rahmati et al., 2019).

Fig. 8 summarizes both information, where along the diagonal the
the most is that, with the exception of very localized differences, the susceptibility of the 1000 simulations, for each case. What stands out is that, with the exception of very localized differences, the susceptibility of the 1000 simulations, for each case. What stands out, whereas the remaining panels correspond to the mean NRT-susceptibility mapping statistical model and simulate from it rather than fitting a new model. This supports the idea in near-real-time studies that one can create a robust predictions to the NRT-susceptibility do not substantially change. This result has been already examined in the literature. Scholars have primarily validated it purely over space for co-seismic landslides (e.g., Nowicki Jessee et al., 2018), or over space (e.g., Lombardo et al., 2014) and time (e.g., Chang et al., 2014) for rainfall-triggered landslides. Therefore, the same concept has not been temporally validated for co-seismic landslides.

In this study, we created three co-seismic landslide inventories. We used high-resolution satellite images and scanned the entire study area manually. As a part of the mapping procedure, we eliminated pre- and post-seismic landslides. As a result, the three co-seismic landslide inventories, created for the same study area, offer the chance to temporally address the time-invariant assumption of causative and triggering effects to the NRT-susceptibility. We remind here that by time-invariance we do not refer to the NRT-susceptibility per se, but to the effect or role that the covariates and their interactions play in the NRT-susceptibility model.

This is shown in Fig. 5 where the marginal distribution of each covariate does not change in sign for the three earthquakes. And, they only slightly shift in amplitude along the abscissa. This is even more stressed at the data level. In fact, we show in Fig. 6 that, to our surprise, the median of morphometric covariates practically does not change from one co-seismic dataset to another. The only change for the PGA and this is intrinsically due to the three different shaking levels.

Our finding showed that, for each case, there is only a minor variation between modeling performance of fitted versus chrono-simulated NRT-susceptibilities (Fig. 9). This situation is particularly evident for the 2017 and 2018 inventories. The AUC value of the fitted 2017 model is 0.89 while its chrono-simulations from the 2012 and 2017 are 0.86 and 0.87, respectively. Similarly, the AUC of the fitted 2018 model is 0.92 while its chrono-simulation from 2012 and 2017 are 0.90 and 0.91, respectively. This difference is relatively larger for chrono-simulated models from the Sulawesi 2012 inventory, which is relatively poor compared to others. In fact, the first cloud-free satellite image covering the entire study area is dated almost one year after the earthquake (see Section 2). Therefore, although we recognized the associated landslides to be primarily linked to the Sulawesi earthquake, the noise due to subsequent rainfall-triggered landslides may be present in the data.

The residuals between fitted and chrono-simulated models show the same situation (Figs. 11). Therefore, we can argue that the contribution of each factor is practically stationary through time and the minor variations can be linked with several sources of uncertainties in the input data as well as in the modeling procedure. Ultimately, we stress the relevance of our plug-in simulation...
procedure. In fact, any other statistically-based susceptibility study with a transferability component either through space (e.g., Lombardo et al., 2015) or through time (e.g., Guzzetti et al., 2006) have so far used a single predictive equation, thus neglecting the uncertainty in the modeling procedure. Here we demonstrate that a number of predictive equations can be simulated from a given fit and used to assess not only the mean behavior but also the uncertainty around it, especially for chrono-validation studies. In other words, we have taken advantage of the full estimated distribution, for each model component, shown in Figs. 2 and 5, which otherwise has been done so far by using the corresponding theoretical mean.

5.2. Opposing arguments

The data sets we used have some limitations which are worth to be discussed more in detail.

First of all, the hypothesis we initially made are only valid within the area we examined, which is a coinciding subset of the three co-
seismic landslide affected areas.

Also, the one-year period between the Sulawesi event and the next fully clear imagery may be problematic for a few reasons. The humid climate of New Guinea is likely to produce rainfall events sufficient to trigger landsliding on a very frequent basis. As a result the seismic signature on the landslide distribution may be partially masked because of the noise introduced by rainfall-induced landslides. We actually believe this to be one of the possible reasons for the lower performance of the Sulawesi model, compared to the Kasiguncu and Palu. However, the Sulawesi model still produced excellent performance. This inevitably implies that the presence of rainfall-induced landslides, if present at all, was negligible compared to the co-seismic landslides. In fact, whether the number and distribution of rainfall-induced landslides would have been comparable to the co-seismic sample, the performance of the model, which does not include any rainfall-proxy covariate, would have likely dropped.

Furthermore, due to the relatively small time span, especially between Kasiguncu 2017 and Palu 2018, some temporal dependence can
exist in terms of slope stability per slope unit. This temporal dependence could be due to the path-dependency proposed by Samia et al. (2017b), which assumes that slope failures can affect the same mapping units over time because of the disturbance of previous landslides to the overall slope equilibrium. Or, this temporal dependence could be due to the legacy effect of an earthquake, as proposed by (Parker et al., 2015), which assumes that the weakening effect to the earth surface caused by a specific seismic event can persist in time. Therefore it may contribute to increase the number of landslides in a subsequent (but close in time) earthquake.

The models we propose in this work do not consider temporal dependence. This could have been implemented by introducing the distribution of previous unstable slope units as a linear covariates, as also proposed by Samia et al. (2017b). However, this would have hindered
our simulation scheme from one event to the other. In fact, it would have not be reasonable to use the distribution of unstable slope units from Palu 2018 in the model built for the Sulawesi 2012. Hence, our chrono-validation scheme could have only been possible in a forward direction.

Another alternative could have been implemented by rigorously modeling the temporal dependence using techniques that are common practices in time-series analyses. For instance, we could have introduced a random effect modeled via an autoregressive structure (Blangiardo and Cameletti, 2015; Lombardo et al., 2020a). However, we envisioned our model for engineering purposes in near-real-time applications and in such cases, even for the rare ones where earthquakes may repeatedly hit the same area, the pre-existing landslide inventories are usually not available. For this reason, we opted to build a modeling framework that could be directly operational for the vast majority of the situations and near-real-time products. Moreover, the Sulawesi 2012 and Palu 2018 ruptured with a strike-slip mechanism whereas the Kasiguncu 2017 occurred in a normal fault. The literature refers to different landslide characteristic (number and areal distribution) in response to different rupture mechanisms and the associated ground motion (Fan et al., 2019; Gorum and Carranza, 2015). This signal would be ideally carried by the PGA. However, in the study area, the raw data used to create the ShakeMap was recorded by few seismic stations, not enough to calibrate an accurate model. Therefore, the PGA is simply interpolated using ground-motion prediction equations (Garcia et al., 2012). Ideally, one should use covariates expressed at the same resolution or scale to avoid co-registration issues. However, the current state of technology and science does not offer the chance to have perfectly co-registered predisposing and triggering factors. For instance, apart from the difference in resolution between terrain properties and PGA in this study, the same problem exist whenever practitioners use a Geological and/or Land Use maps (e.g., Nowicki

Fig. 11. Residual maps expressed as the difference between the posterior mean of the three reference GAM fits and the mean of the 1000 simulations, for each combination. The diagonal here is masked out because we did not simulate from the fit to explain the same co-seismic landslide distribution.
Jesse et al., 2018), vegetation indices (e.g., Lee et al., 2008), rainfall measurements (e.g., Kirschbaum and Stanley, 2018) and any other property collected at sites and then interpolated at a higher resolution via downsampling methods. Notably, co-registration issues affect pixel-based susceptibility and hazard models to a much larger extent than they do in case of a SU partition where the different resolution of the data is upscaled to the coarser SU scale.

Despite these limitations, our models produced at least excellent performance. Because the PGA map and the landslide distribution are data collected independently from each other, such high performance suggests that the quality of the ground motion data played a minor role in the overall modeling procedure. Or, that the quality of the ShakeMap was enough to explain the distribution of landslides per SU.

We conclude this section by noting that the modeling framework we proposed for co-seismic landslides cannot be used for long-term planning unless a wide range of scenario based PGA would be available. In fact, at the present stage, we only used three events, but a potential future earthquake may exhibit entirely different patterns, making our probabilistic maps useless if not specifically during post-disaster emergency phases – for which near-real-time models– have been proposed.

6. Conclusions

In this work, we tested a multi-temporal chrono-validation scheme for earthquake-induced landslides via plug-in statistical simulations. Overall, our work is a rare case where multi-temporal co-seismic inventories have been mapped within the same study site, although the extent of this site is only a part of the whole landslide affected area. Together with our chrono-validation scheme, where a Bayesian GAM is used to calibrate and simulate from for specific scenarios, our work offers a different experimental design compared to other earthquake-induced landslide studies. In fact, the ground motion and the other causative factors appeared to be stationary over space and time in their contribution to the NRT-susceptibility. Or, at least their differences are negligible with respect to the overall performance. And, we propose a rigorous validation pipeline where a plug-in simulation framework is temporarily tested for the first time.

This work supports the assumption held in near-real-time predictive studies where the regression coefficients are typically fixed only to leave the trigger pattern to change over space. Notably, the chronova lidity showed signs of reduced predictive performance, although still excellent, when the co-seismic inventory may have been affected by mapping uncertainties. This may indicate that, even more for traditional chrono-validation routines where the uncertainty is totally missed, a complete co-seismic inventory map is of crucial importance. Here at least, we account for potential sources of variability in the data by simulating thousands of NRT-susceptibility realizations.

We stress that our model does not exactly estimate the susceptibility because it features the signal of the trigger, making the results dependent on a specific event rather than expressing the long-term proneness of the landscape to fail. However, our model does not estimate the landslide hazard either because it does not satisfy the requirement on the temporal recurrence nor the landslide magnitude of the event (Guzzetti et al., 2006). As a result, we consider it a hybrid whose use can be chiefly implemented in near-real-time application.

Given these findings, we believe our plug-in simulation approach to be a more informative modeling procedure than current NRT-susceptibility models and we envision future applications in this direction, not only for co-seismic landslide occurrences but also for their rainfall-triggered counterparts.

Declaration of Competing Interest

None.

References

Allstadt, K.E., Jibson, R.W., Thompson, E.M., Massey, C.I., Wald, D.J., Godt, J.W., Rengers, F.K., 2018. Improving near-real-time coseismic landslide models: lessons learned from the 2016 Kaikoura, New Zealand, Earthquake. Bull. Seismol. Soc. Am. 108 (3B), 1649–1664.

Alvioli, M., Marchesini, L., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., Guzzetti, F., 2016. Automatic delineation of geomorphological slope units with r.sloopeunits v1.0 and their optimization for landslide susceptibility modeling. Geosci. Model Dev. 9 (11), 3975–3991.

Amato, G., Eisank, C., Castro-Camilo, D., Lombardo, L., 2019. Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment. Eng. Geol. 260 (In print).

Bakka, H., Rue, H., Fuglstad, G.-A., Ruben, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., Lindgren, F., 2018. Spatial modeling with R-INLA: a review. Wiley Interdiscip. Rev. 10 (6), e1443.

Bakka, H., Vanhatalo, J., Illian, J.B., Simpson, D., Rue, H., 2019. Non-stationary Gaussian models with physical barriers. Spatial Stat. 29, 266–288.

Beguería, S., 2006. Validation and evaluation of predictive models in hazard assessment and risk management. Nat. Hazards 37 (3), 315–329.

Bivand, R., Gómez-Rubio, V., Rue, H., 2015. Spatial Data Analysis With R-INLA With Some Extensions.

Blangiardo, M., Cameletti, M., 2015. Spatial and Spatio-Temporal Bayesian Models With R-INLA. John Wiley & Sons.

Brabb, E.E., 1984. Innovative approaches to landslide hazard and risk mapping. In: International Landslide Symposium Proceedings, Toronto, Canada. 1 pp. 17–22.

Bradley, K., Mallick, R., Andikagumi, H., Hubbard, J., Meiliandsa, S., Switzer, A., Du, N., Brocard, G., Alifan, D., Benazir, B., et al., 2019. Earthquake-triggered 2018 Pulu Valley landslides enabled by wet rice cultivation. Nat. Geosci. 12 (11), 935–939.

Budimir, M.E.A., Atkinson, P.M., Lewis, H.G., 2015. A systematic review of landslide probability mapping using logistic regression. Landslides 1–18.

Cama, M., Lombardo, L., Conoscenti, C., Agnesi, V., Rotigliano, E., 2015. Predicting storm-triggered debris flow events: application to the 2009 lotionurian landslide disaster in Sibila, Italy. Nat. Hazards Earth Syst. Sci. 15 (8), 1785–1796.

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

Castro-Camilo, D., Lombardo, L., Mai, P., Dou, J., Huser, R., 2017. Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized Generalized Linear Model. Environ. Model. Softw. 97, 145–156.

Chang, C.-H., Chiang, S.-h., Chen, Y.-c., Mondini, A.C., 2014. Mapping the spatial occurrence of shallow landslides triggered by typhoons. Geomorphology 208, 137–148.

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

Conoscenti, C., Rotigliano, E., Cama, M., Carabalbo-Arias, N.A., Lombardo, L., Agnesi, V., 2016. Exploring the effect of absence selection on landslide susceptibility models: a case study in Sicily, Italy. Geomorphology 261, 222–235.

Crozier, M., 2005. Multiple-occurrence regional landslide events in New Zealand: hazard management issues. Landslides 2 (4), 247–256.

Dou, J., Yamagishi, H., Pourghaemi, H.R., Yunes, A.P., Song, X., Xu, Y., Zhu, Z., 2015. An integrated artificial neural network model for the landslide susceptibility assessment of Osaka Island, Japan. Nat. Hazards 78 (3), 1749–1776.

Fan, X., Scaringi, G., Korup, O., West, A.J., van Westen, C.J., Tanyas, H., Hovius, N., Hales, T.C., Jibson, R.W., Allstadt, K.E., et al., 2019. Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. Rev. Geophys. 57 (2), 421–502.

Fratini, P., Crosta, G., Carrara, A., 2010. Techniques for evaluating the performance of landslide susceptibility models. Eng. Geol. 111 (1), 62–72.

Garca, D., Mah, R., Johnson, K., Hearne, M., Marano, K., Lin, K., Wald, D., Worden, C., So, E., 2012. ShakeMap Atlas 2.0: an improved suite of recent historical earthquake ShakeMaps for global hazard analyses and loss model calibration. In: World Conference on Earthquake Engineering.

Garca, D., Wald, D.J., Hearne, M., 2012. A global earthquake discrimination scheme to optimize ground-motion prediction equation selection. Bull. Seismol. Soc. Am. 102 (1), 185–203.

Godt, J., Sener, B., Verdin, K., Wald, D., Earle, P., Harel, E., Jibson, R., 2008. Rapid assessment of earthquake-induced landsliding. In: Proceedings of the First World Landslide Forum. 4 United Nations University, Tokyo 3116–11.

Goeteb, J.N., Guthrie, R.H., Brenning, A., 2011. Integrating physical and empirical landslide susceptibility models using generalized additive models. Geomorphology 129 (3–4), 376–386.

Goeteb, J., Brenning, A., Petschko, H., Leopold, P., 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput. Geosci. 81, 1–11.

Gorum, T., Carranza, E., 2015. Control of style-of-faulting on spatial pattern of earth- quake-triggered landslides. Int. J. Environ. Sci. Technol. 12 (10), 3189–3212.

Gorum, T., Fan, X., van Westen, C.J., Huang, R.Q., Xu, Q., Tang, C., Wang, G., 2011. Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake. Geomorphology 133 (3–4), 152–167.

Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study. Central Italy. Geomorphology 31 (1), 181–216.

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., Ardizzone, F., 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72 (1-4), 272–299.

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., Cardinali, M., 2006. Landslide
L. Lombardo and H. Tanyas
Engineering Geology 278 (2020) 105818

hazard assessment in the Collazzone area, Umbria, Central Italy. Nat. Hazards Earth Syst. Sci. 6 (1), 115–131.

Guzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: New tools for an old problem. Earth Sci. Rev. 112 (1–2), 42–66.

Heerden, R.G., Beran, M.A., 1982. Quantifying source areas through land surface curvature and shape. J. Hydrol. 57 (3–4), 359–373.

Homer, D.W., Leneshow, S., 2000. Applied Logistic Regression, Second edition. Wiley, New York.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-Filled SRTM for the Globe Version 4.

Jasiewicz, J., Stepinski, T.F., 2013. Geomorphicosa—a pattern recognition approach to classification and mapping of landforms. Geomorphology 182, 147–156.

Jiang, C., Jhi, Y.-Y., Lee, D.-H., 1998. Stability analysis of existing slopes considering uncertainty. Eng. Geol. 49 (2), 111–122.

Kirschbaum, D., Stanley, T., 2018. Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness. Earth’s Future 6 (3), 505–523.

Knevels, R., Petschko, H., Proske, H., Leopold, P., Maraun, D., Brenning, A., 2020. Event-based landslide modeling in the Styrian Basin, Austria: accounting for time-varying rainfall and land cover. Geosciences 10 (6), 217.

Korup, O., 2004. Geomorphic implications of fault zone weakening: slope instability along the Alpine Fault, South Westland to Fiordland. N. Z. J. Geophys. 47 (2), 257–267.

Kritikos, T., Robinson, T.R., Davies, T.R., 2015. Regional coseismic landslide hazard assessment without historical landslide inventories: a new approach. J. Geophys. Res. Earth Surf. 120 (4), 711–729.

Lee, C., Huang, C., Lee, J., Pan, K., Lin, M., Dong, J., 2008. Statistical approach to storm event-induced landslides susceptibility. Nat. Hazards Earth Syst. Sci. 8 (4), 941–960.

Lindgren, F., Rue, H., 2008. On the second-order random walk model for irregular locations. Scand. J. Stat. 35 (4), 691–700.

Lombardo, L., Mai, P.M., 2018. Presenting logistic regression-based landslide susceptibility results. Eng. Geol. 244, 14–24.

Lombardo, L., Tanyas, H., 2020. From scenario-based seismic hazard to scenario-based landslide hazard: rewinding to the past via statistical simulations. arXiv preprint arXiv:2004.00539.

Lombardo, L., Cama, M., Märker, M., Rotigliano, E., 2014. A test of transferability for landslide susceptibility zonation based on multiple forecasts. Geomorphology 114 (3), 129–142.

Samia, J., Temme, A.J., Bregt, A., Wallinga, J., Guzzetti, Fausto, Ardizzone, F., Rossi, M., 2017a. Characterization and quantification of path dependency in landslide susceptibility. Geomorphology 292, 16–24.

Samia, J., Temme, A.J., Bregt, A., Wallinga, J., Guzzetti, Fausto, Ardizzone, F., Rossi, M., 2017b. Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. Landslides 14, 547–558.

Sappington, J.M., Longshore, K.M., Glade, T., 2020. The (f)utility to account for pre-failure topography and environmental factors in probabilistic landslide susceptibility zonation, version 1.0. Geosci. Model Dev. 9 (10), 3553.

Sterr, S., Schmutz, E., Glade, T., 2020. The (f)utility to account for pre-failure topography and environmental factors in probabilistic landslide susceptibility zonation, version 1.0. Geosci. Model Dev. 9 (10), 3553.

Steger, S., Schmaltz, E., Glade, T., 2020. The (f)utility to account for pre-failure topography and environmental factors in probabilistic landslide susceptibility zonation, version 1.0. Geosci. Model Dev. 9 (10), 3553.