Saponaro, A., Pilz, M., Wieland, M., Bindi, D., Moldobekov, B., Parolai, S. (2015): Landslide susceptibility analysis in data-scarce regions: the case of Kyrgyzstan. - *Bulletin of Engineering Geology and the Environment*, *74*, 4, p. 1117-1136.

DOI: http://doi.org/10.1007/s10064-014-0709-2

Landslide susceptibility analysis in data-scarce regions: the case of Kyrgyzstan

Annamaria Saponaro[1], Marco Pilz[1], Marc Wieland[1], Dino Bindi[1], Bolot Moldobekov[2] and Stefano Parolai[1]

[1]{ Centre for Early Warning, Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam, Germany}

[2]{Central Asian Institute for Applied Geosciences, Bishkek, Kyrgyzstan}

Correspondence to: Annamaria Saponaro

Email: asapon@gfz-potsdam.de

Postal address: Centre for Early Warning, Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Helmholtzstr. 7, 14467 Potsdam,

Telephone: +49 (0)331 288-28661

Fax: +49 (0)331 288-1204

Abstract

Kyrgyzstan is one of the most exposed countries in the world to landslide hazard. The large variability of local geological materials, together with the difficulties in forecasting heavy precipitation locally and in quantifying the level of ground shaking, call for harmonized procedures to better quantify the hazard and the negative impact of slope failures. By exploiting new advances in Geographic Information System (GIS) technology, together with

concepts from Bayesian statistics, and promoting the use of open-source tools, we aim to identify areas in Kyrgyzstan where the potential for landslide activation exists.

A range of conditioning factors and their potential impact on landslide occurrence are quantitatively assessed on the basis of the spatial distribution of landslides by applying Weights-of-Evidence modelling based on 1) a landslide inventory of past events, 2) terrain-derived variables of slope, aspect and curvature, 3) a geological map, 4) a distance from faults map, and 5) a seismic intensity map. A spatial validation of the proposed method has been performed, indicating sufficient measures of significance to predicted results.

Initial results are promising and demonstrate the applicability of the method to the entire Kyrgyzstan, allowing the identification of areas that are more susceptible to landslides with a level of accuracy greater than 70%. The presented method is, therefore, capable of supporting land planning activities at the regional scale in places where only scarce data are available.

Keywords: Landslides; Susceptibility; GIS; Weights-of-Evidence; Kyrgyzstan.

1 Introduction

Increasing populations and the expansion of urban settlements towards landslide-prone slopes contribute to the worldwide destructive impact of landslides (Rosenfeld, 1994). In particular, recent studies (Petley, 2012) highlighted that most of losses of life are concentrated in less developed countries, where there is a relatively little investment in understanding the hazards and risks associated with landslides, due largely to a lack of appropriate resources. In line with this, it is well recognised that Kyrgyzstan represents a hotspot in terms of natural hazard with a specific link between earthquakes and landslides (Nadim et al., 2006). According to recent surveys referring to the time period between 1988 and 2007 (Risk Assessment for Central Asia and Caucasus, 2009), 18% and 27% of yearly reported disasters in Kyrgyzstan are due to earthquakes and landslides, respectively. In particular, more than 300 large-sized landslides occurred between 1993 and 2010, resulting in an average of 256 deaths per year (Torgoev et al., 2012) with substantial associated economic losses (an average of 2.5 million USD per year).

The roots of the strong hazardous component in Kyrgyzstan are based on a combination of several factors. The presence of substantial geodynamic activity linked to the collision of Eurasian and Indian tectonic plates is responsible for the development of the Tien Shan Mountains (Molnar, 1975). With elevations rising up to 7,439 m a.s.l., the Tien Shan is characterised by high topographic relief and mainly trends in a W-E direction. Correspondingly, the presence of many active faults is the principal source for the high seismic activity in this region (Trifonov, 2002). A relatively large variability characterizes the geological units, which generally consist of Cenozoic-Mesozoic deposits, mainly composed of sandstone, siltstone with gypsum inter-beds, and conglomerates. Although continental climate conditions are prevalent, the occurrence of heavy precipitation, which mainly occurs

in spring, acts as an added component to hazards in cases where rocks are unconsolidated, leading to the potential to trigger landslides and to initiate floods.

Actions with the aim of understanding and controlling slope instability phenomena are, therefore, necessary to implement appropriate landslide risk mitigation measures. Authorities and decision makers who are responsible for regional land use planning are in the constant need for maps that show the areas that may be endangered by landslides. To this purpose, the first step is typically represented by preparing a landslide susceptibility map. Landslide susceptibility is the probability of the spatial occurrence of known slope failures, given a set of geoenvironmental conditions (Guzzetti, 2005). It is the potential of a terrain to be affected by slope movements providing an estimate of "where" landslides are likely to occur in the future. Additionally, landslide hazard is defined as the probability of a landslide occurrence within a specified time and within a given area of potentially damaging phenomenon. In situations where the temporal information on past landslide occurrences is missing, determination of the probability of landslide occurrence within a defined time interval is prevented. Therefore, the analysis of landslide susceptibility is highly supportive under these limited circumstances.

The problems of landslide susceptibility and hazard have been addressed in several ways – according to the scale of analysis and the aim of investigation - during the last decades. The foundations for landslide hazard analysis were laid by Varnes (1984). In his work, he clarified how it is possible to identify areas where a potential for landsliding exists by exploiting the uniformitarian principle, which states that "the past and the present are the keys for the future": slope failures in the future are more likely to happen under the same conditions that led to past and current instability. General overviews of research in the topic of landslide susceptibility can be found in the works of Soeters and van Westen (1996), Aleotti and

Chowdhury (1999), Carrara et al. (1999), Guzzetti et al. (1999), Dai et al. (2002), van Westen et al. (2006), and Fell et al. (2008). Among statistical approaches, bivariate methods – e.g. the Weights of Evidence (WOE) - offer the advantage to determine the influence of a single parameter in landsliding, an aspect that is particularly useful in regional scale analysis where data are poorly defined.

The reliability of landslide susceptibility and hazard maps depends on the amount and the quality of input data. Geographic Information System (GIS) are currently adopted for improving landslide inventory mapping and spatial data analysis. Numerous studies already showed that the spatial distribution of landslides can be better understood through GIS-based susceptibility assessments and successful examples of regional scale analyses can be found in the work of Chung and Fabbri (2003), van Westen et al. (2003), Neuhäuser and Terhost (2006), Oh and Lee (2010), Schicker and Moon (2012), Holec et al. (2013).

Landslide processes and their impacts have already been investigated in Kyrgyzstan, at the local scale. Examples can be found for the Suusamyr (Havenith et al., 2006) and the Mailuu-Suu (Torgoev and Havenith, 2013) regions, as well as for South Kyrgyzstan (Roessner et al., 2005). Nevertheless, a robust statistical analysis of country-wide landslide susceptibility as well as of the existing relationships among the most influential factors across the country, at a regional scale, is not achieved yet.

With these premises, the principal aim of this manuscript is to investigate landslide susceptibility and the related contributing factors in Kyrgyzstan, using WOE and GIS environment. Topographic and tectonic causative factors have been selected and corresponding maps have been prepared in GIS. Weights for different categories of these factors have been statistically determined and then integrated into the GIS to prepare landslide susceptibility maps for the country.

Furthermore, an inventory of past landslides, including information about the date of occurrence and a detailed description of failure mechanism is currently missing. Evaluating landslide susceptibility when limited background information and data are available constitutes a constant challenge for engineers and geologists in their efforts to cooperate with planners and government bodies. This work will try to overcome this difficulty. The presented approach will promote the implementation of open source software (QGIS, GRASS, R), and take advantage of their ease of distribution, an aspect that is particularly desirable in developing countries such as those in Central Asia.

2 Study area

A range of different slope failures have affected the study area, mainly landslides, rockslides and rock avalanches. Based on the type of dominant process, slope failures in Kyrgyzstan are subdivided into: 1) mass movements along the slope surface occurring in soft and semi-hard rocks, and 2) rockfalls occurring in steep cliffs and semi-hard and hard-rocks. In particular, the majority of slope failures are confined to mountainous regions, at an altitude ranging from 800 to 1,200 m.

To date, 5000 potential active landslide sites have been identified in Kyrgyzstan (Kalmetieva et al, 2009) with 80% of landslides located along the borders of large depressions in Cenozoic-Mesozoic deposits where the presence of fine dispersed lithologies – e.g. clays, algillities – in combination with the level of natural moisture enhances the development of slope instability. In most of the cases, these materials are covered by Quaternary loess deposits that – being relatively thin –generally develop as earth flows and superficial wash-out. On the other hand, rockslides and rockfalls tend to be close to tectonic faults, and their size is not considerable.

A significant number of large landslides (size in the order of millions of m^3) have been recorded in the Kyrgyz territory, some of which are shown in Fig. 1. Among others, on 14 April 1994, a major landslide in the Osh-Jalal-Abad region killed 111 people, affected (injured, damaged houses) another 58,500 (Moldobekov et al., 1997). In April 2003, a landslide in the Uzgen district killed 38 people and affected another 211. In April 2004, two different landslide events in the Alay district and the Kara-Sogot region killed a total of 38 people and affected 96 (CAC DRMI, 2009). Additionally, historical evidence reports about the probable connection between induced-slope failures and the 1992 Suusamyr and 1946 Chatkal earthquakes (Fig. 1) – among the strongest known earthquakes (I = IX-X) to have occurred in the region (Havenith et al., 2006; Kalmetieva et al., 2009). Furthermore, Strom and Korup (2006) provided evidence of two rockslides, the Chukurchak and Sarychelek slope failures, which date back to prehistoric time.

The Jalad-Abad province has been selected as study area, given the wide variability of landslide factors in the area. Additionally, as it can be seen from histograms of frequencies (Fig. 2), the distribution of landslide factors in Jalal-Abad area is comparable to the one concerning the entire Kyrgyz country. For this reason, this area is considered representative of the existing relationships among landslide factors in Kyrgyzstan.

Among different types of slope failures in the area, for the purpose of this study only landslides occurring in soft materials are addressed, as they are the most represented class.

3 Data preparation

The primary steps in landslide susceptibility assessment are data collection and the construction of a spatial database. Usually, the identification of factors correlated with slope instability is based on the choice of physically-based indicators (Guzzetti et al., 1999). For the

present study, slope gradient, slope aspect, profile curvature, geology, distance from faults and seismic intensity have been selected. An inventory of landslides in the region has been compiled and used as a reference dataset. Slope gradient, slope aspect, and profile curvature have been derived from the NASA released Shuttle Radar Topographic Mission digital elevation model, with a spatial resolution of 81.7 m (SRTM, 2004). Geology, distance from faults and seismic intensity maps have been converted into raster format and then resampled to the same spatial resolution of the topography.

3.1 Slope gradient

The stability of a slope is known to be highly dependent upon the slope angle and its material properties (Terzaghi and Peck, 1967). The slope is presented in degrees ranging from 0° to 88° (Fig. 3a); for the purpose of this study, slope values have been divided into four bins with approximately the same number of features (quantile classification), 0°-6.6°, 6.6°-16.6°, 16.6°-27.5°, >27.5°, to facilitate the comparison of results.

3.2 Slope aspect

Aspect is defined as the direction of maximum slope of the terrain surface. In landslide susceptibility analysis at the regional scale, the influence of slope aspect is typically taken into consideration, although in some cases its importance has been questioned (Guzzetti et al., 1999). For the selected areas (Fig. 3b), a classification based on azimuth being divided into eight bins, North, North-East, East, South-Est, South, South-West, West, North-West, has been carried out.

3.3 Profile Curvature

Curvature represents one of the topographic attributes which are also commonly included in a landslide susceptibility analysis (Alalew et al., 2004). Two possible outputs of this variable are possible: profile curvature, which is defined as the second derivative of the slope with respect to the maximum steepness direction, and plan curvature, which is perpendicular to the direction of the maximum slope. The profile curvature is known in that it affects the acceleration and deceleration of flow and, therefore, influences erosion and deposition. For the selected area (Fig. 3c), profile curvature values have been classified (quantile classification) into four bins, -0.02507− -0.00101, -0.00101− -0.00005, -0.00005− 0.00095, 0.00095 -0.01891.

3.4 Geology

Lithology plays an important role in landslide susceptibility studies because different geological units have different slope failure behaviours. For example, old rocks are clearly less prone to landslide occurrence than more recent lithologies due to their higher compactness.

For our study, geological information has been obtained from “The Geological Map of Central Asia and Adjacent Areas” (Tingdong et al., 2008), scaled 1:2,500,000; overall, the study area is covered by a range of different sedimentary formations, mostly dated to the Quaternary, Neogene, Paleogene, Cretaceous, Jurassic, and Triassic. Igneous rocks related to the Palaeozoic epoch are also present.

Based on this information, stratigraphic units have been digitised and classified into Paleozoic, Mesozoic and Cenozoic units (Fig. 3d).

199 *3.5 Distance from faults*

200 The presence of major lineaments is among the important factors governing the stability of
201 slopes (Varnes, 1984). Tectonic structures form zones of weakness in rocks and might
202 accelerate the process of slope failures. In landslide susceptibility studies, distance from
203 lineament features (i.e., faults) is typically used to investigate any cause-effect relationships
204 between lineaments and landslide occurrence (Gemitzi et al., 2011; Pradhan et al., 2010). The
205 entire Kyrgyz territory is covered by a large number of active faults. The largest fault is the
206 NW-SE trending Talas-Fergana fault accounting for a total horizontal displacement of 200 km
207 (Kalmetieva et al., 2009). For our analysis, fault lines were derived from the 1: 2,500,000
208 scale geology map and four-buffer zone maps (< 1km, 1 - 5km, 5 - 10km, > 10km) were
209 prepared in GIS (Fig. 3e).

210 *3.6 Seismic intensity*

211 As the study area is strongly affected by earthquakes, it is necessary to take seismic ground
212 shaking into account as a triggering factor for landsliding. In this study, ground shaking is
213 expressed through the observed macro-seismic intensity (MSK 64), as part of a
214 comprehensive database recently produced for Central Asian countries (Bindi et al., 2012).
215 Information about historical earthquakes occurring in various locations in Kyrgyzstan reveals
216 that the resulting observed effects are associated with maximum intensity values ranging from
217 I = VII to I = IX. Starting from these observations, seismic intensity values with a 10%
218 probability of being exceeded in 50 years have been estimated by Bindi et al. (2012) for all
219 Central Asian countries including the study area. Overall, this study returned intensities of
220 VIII and IX as expected in the future for all Kyrgyz territory. After importing values to GIS, a
221 raster map has been created and the values have been categorized into 3 classes: I = VII, I =
222 VIII, and I = IX (Fig. 3f). Although intensity values have not been provided at each single

223 landslide location, such a map provides a good suggestion about future intensities, and hence
224 the potential for landslide triggering.

225 *3.7 Landslide locations*

226 A landslide inventory represents an essential ingredient in order to carry out landslide hazard
227 analysis at the regional scale (Guzzetti et al., 1999). It helps in identifying locations of
228 previous landslides in order to be able to predict future failures. There is no agreement on the
229 technique for the preparation of landslide inventory maps, researchers usually adopt different
230 inventory maps where landslides are shown as points, scarp, seed cells (Yilmaz, 2010). Small-
231 scale maps may only show landslide locations (points strategy) as, due to the scale of the map,
232 it is not possible to draw the landslide extension. On the other hand, large-scale maps may
233 distinguish between source and deposit areas (Yilmaz, 2010).

234 For the purposes of this study, we applied point strategy and considered only one type of mass
235 movement. Therefore, a selection of landslide locations has been defined on the basis of
236 published information (Kalmetieva et al., 2009), and their distribution has been mapped (Fig.
237 3g).

238 All the input parameters have been arranged in raster format with a cell-size equivalent to the
239 SRTM digital elevation model.

240

241 **4 Weights-of-Evidence Method**

242 In this study, the Weights-of-Evidence method is used for generating a landslide susceptibility
243 map. The Weights-of-Evidence method is a data–driven quantitative method used to combine
244 evidences in support of a hypothesis (Bonham-Carter, 1994). The method was originally

developed for medical studies and subsequently has been applied to other disciplines like e.g. identifying mineral deposit potential (Bonham-Carter et al., 1989).

With respect to other methods, the Weights-of-Evidence method has been successfully used in previous landslide susceptibility studies for examining the distribution and spatial relationships of particular features. The method offers a flexible way of testing the importance of various input factors to the potential of slope failure, providing a simple statistical and straightforward tool that allows the calculated weights to be interpreted. Although already carried out in large scale contexts, the method has not been previously tested in data-scarce regions.

Within the context of landslide susceptibility analysis, the influence of landslide potential factors (evidence) on the occurrence of landslides themselves (hypothesis) is assessed. Weights for each landslide causative factor are calculated based on the presence or absence of landslides within the study area.

4.1 Theory of Weights-of-Evidence

Considering a given number of cells affected by landslide phenomena ($N\{L\}$), then the prior probability of landslide occurrence $P\{L\}$ within the studied area T is expressed by (Bonham-Carter, 1994):

$$P\{L\} = \frac{N\{L\}}{N\{T\}} \qquad 1)$$

This initial estimate can be increased or decreased based on the relationships between landslide potential factors and landslides. In particular, the probability of finding a landslide potential factor within the studied area is given by: $P\{F\} = N\{F\} / N\{T\}$.

Suppose that a landslide potential factor occurs in the studied area, and that a number of known landslides occur preferentially within the factor, it is possible to indicate the probability of finding a landslide given the presence (F) or the absence ($\overline{F}$) of a factor, through the definition of conditional probabilities:

$$P\{L \mid F\} = \frac{P\{L \cap F\}}{P\{F\}} = P\{L\}\frac{P\{F \mid L\}}{P\{F\}} \quad 2)$$

$$P\{L \mid \overline{F}\} = \frac{P\{L \cap \overline{F}\}}{P\{\overline{F}\}} = P\{L\}\frac{P\{\overline{F} \mid L\}}{P\{\overline{F}\}}, \quad 3)$$

where $P\{F \mid L\}$ and $P\{\overline{F} \mid L\}$ are the conditional probabilities of being and not being within the factor, given the presence of a landslide.

Equations 2) and 3) can be expressed in an odds-type formulation, where the odds, O, are defined as: $O = P/(1-P)$.

The weights for a landslide potential factor are, hence, defined as:

$$W^+ = \ln\frac{P\{F \mid L\}}{P\{F \mid \overline{L}\}} \quad 4)$$

$$W^- = \ln\frac{P\{\overline{F} \mid L\}}{P\{\overline{F} \mid \overline{L}\}}. \quad 5)$$

In the Weights-of-Evidence method, the natural logarithms of both sides of equations are taken, resulting in:

$$\ln O\{L \mid F\} = W^+ + \ln O\{L\} \quad 6)$$

$$\ln O\{L \mid \overline{F}\} = W^- + \ln O\{L\}. \quad 7)$$

In case the influence of several factors on the distribution of landslides is taken into analysis, then the summation of the weights of each factor is used, provided that these factors are conditionally independent. The general expression for combining $i = 1,2,3...n$ landslide factors is:

$$\ln O\{L \mid F_1 \cap F_2 \cap F_3 \cap \ldots \cap F_n\} = \sum_{i=1}^{n} W^+ + \ln O\{L\}. \quad 8)$$

If the $i-th$ factor is absent, then W^+ becomes W^-. The difference between positive and negative weights is known as the weight contrast ($C = W^+ - W^-$) and provides a useful measure of the overall spatial correlation between a certain class of factor and the occurrence of landslides (Bonham-Carter, 1994). A positive C indicates that the causative factor is present at the landslide location, and its magnitude is a measure of the positive correlation between the presence of the causative factor and landslides. On the other hand, a negative C is used to assess the importance of the absence of the factor in landslide occurrence. Factors with contrast values around 0 have no significant connection with the occurrence of landslides. The statistical significance of the weights can be verified by calculating their variances (S^2) together with the studentized contrasts ($C / S(C)$) by means of the following equations:

$$S^2(W^+) = [1/N\{F \cap L\} + 1/N\{F \cap \overline{L}\}] \quad 9)$$

$$S^2(W^-) = [1/N\{\overline{F} \cap L\} + 1/N\{\overline{F} \cap \overline{L}\}] \quad 10)$$

$$S^2(C) = S^2(W^+) + S^2(W^-) \quad 11)$$

A script code in R has been prepared to perform the necessary calculations.

292 *4.2 Test for Conditional Independence*

293 In Weights-of-Evidence modelling, it is typically assumed that factors, which have been
294 outlined in section 3, are conditionally independent with respect to landslide occurrences.

295 It can be shown that equation 8) is equivalent to:

$$N\{F_1 \cap F_2 \cap L\} = \frac{N\{F_1 \cap L\} N\{F_2 \cap L\}}{N\{L\}} \qquad 12)$$

296 The left side of equation 12) indicates the observed number of occurrences in the overlap zone
297 where both F_1 and F_2 are present, while the right side represents the expected number of
298 occurrences in this overlap zone. A contingency table can be prepared based on this
299 relationship for testing the conditional independency of two factors (Table 1). The chi-square
300 values are calculated at the 99% significance level and 1 degree of freedom, and compared
301 with table values. Calculated chi-square values greater than 6.64 suggest that the pairs are not
302 significantly different. In our study, conditional dependency exists between distance from
303 faults and seismic intensity ($\chi^2 = 9.579$), while geology and distance from faults are
304 conditional independent of each other ($\chi^2 = 0.320$). This means, for example, that the pair
305 distance from faults-seismic intensity should not be used together to map landslide
306 susceptibility. On the contrary, the pair geology-distance from faults could be combined.
307 Therefore, four models are derived starting from independent factors (Table 2).

308

309 **5 Susceptibility results**

310 The landslide susceptibility analysis has been conducted for the Jalal-Abad region over a
311 landslide sample consisting of 1,347 landslide locations. Specifically, 50% of the total
312 number of locations have been randomly selected in the study areas, and then used as the

313 “training dataset”. The remaining 50% of locations have been used as the “test dataset” for
314 validating results. The “training dataset” of landslides has been overlaid with each landslide
315 potential factor to calculate weights and the statistical parameters representative of existent
316 spatial relationships (Table 3) by applying equations 4, 5, 8, 9, 10, and 11. As can be seen, the
317 most noteworthy classes of parameters with a positive impact on slope instability are: slope
318 gradient 6° - 16.6°, north-facing slope aspect, profile curvature -0.00101(1/m) − -0.00005
319 (1/m), mesozoic-aged lithologies, distance from faults greater than 10km, and seismic
320 intensity values (MSK 64) equal to IX. Furthermore, the highest contrasts ($C / S(C)$) values
321 are found for the geology factor, while lowest ones are for profile curvature.

322 Based on the weights calculations, landslide factors maps have been re-classified according to
323 their positive or negative correlation with landslide locations. A landslide susceptibility
324 zonation map has hence been obtained by combining previously calculated contrast values
325 with re-classified factors maps based on:

$$LSI = \sum_{j=1}^{n} C_{ij}, \qquad 13)$$

326 where LSI indicates a Landslide Susceptibility Index and C_{ij} represents the contrast for the
327 i-th bin of the j-th factor. Precisely, five different landslide susceptibility maps have been
328 prepared by combining those landslide factors which have been found conditional
329 independent of each other (Table 4), and by considering all factors together. The result of the
330 summation is a continuous interval of values ranging from -6.720 to 4.531 (Model A), from -
331 3.771 to 3.503 (Model B), from -3.438 to 3.124 (Model C), from -2.848 to 2.011 (Model D),
332 from -9.062 to 6.021 (Model E), indicative of various degrees of landslide susceptibility.
333 Normalised values have been calculated by dividing the difference between a value and the
334 minimum result by the maximum minus the minimum, allowing them to take on values

between 0 and 1: from 0.033 to 0.955 (Model A), from 0.060 to 0.945 (Model B), from 0.087 to 0.932 (Model C), from 0.103 to 0.903 (Model D), from 0.016 to 0.971 (Model E). Afterwards, susceptibility values have been classified into 10 equal-sized sectors corresponding to different levels of susceptibility (Fig. 4).

In particular, it can be observed (Fig. 4) that high LSI levels are present in the southern area, precisely along the eastern border of Fergana valley, where slope values mostly range from 0.0º to 16.6º, and the majority of past landslides are also distributed. In addition, high landslide susceptibility are recognized across the Jalal-Abad province region, though the lack of landslide observations. This latter result might indicate the potential for landslide activation and, therefore, serve as input when estimating landslide risk.

Based on these results, the landslide susceptibility model calibrated for this region can be considered reliable as the input landslide factors are good indicators of existing variability conditions.

6 Validation

To check the predictive capabilities of any model, an essential requirement is to carry out a validation of the results. Without some kind of validation, such results are useless since they lack knowledge of the degree of confidence in the model, a crucial element for transferring results to end users and stakeholders (Chung and Fabbri, 2003).

Cross-validation is commonly used for assessing the capability of results from a statistical analysis to be generalized to an independent data set. The procedure consists of partitioning a sample of data into complementary subsets. The analysis is, then, performed on one subset – named the “training set”, and the validation carried out on the other subset – named the “test set”.

In the field of landslide hazard assessment, the cross-validation of the results is commonly carried out by partitioning the data in time or in space (Chung and Fabbri, 2003). When the temporal approach is chosen, landslide occurrences are subdivided into two subsets referring to different time periods, typically named “past” and “future” landslides. This approach is meant to construct the prediction model based on “past” occurrences and then to validate the results with respect to “future” ones. When temporal information is missing, space robustness validation is commonly applied.

The validity and accuracy of landslide susceptibility maps are typically ascertained with the help of success- and prediction-rate curves in combination with the area under the curves. The curves provide information about the relationship between the percentage of area identified as being landslide susceptible and the actual landslide occurrences. In particular, success-rate curves show how good the susceptibility model is in fitting the already occurred landslides; prediction-rate curves provide quantitative information about landslides that might occur in the future. The area under the curve provides a measure of the total accuracy based on the rate curves, where a total area equal to one indicates perfect accuracy. The common procedure consists of sorting in descending order the calculated index values that refer to the total number of cells in the study area. Landslide susceptibility results are hence cross-tabulated with landslide locations and presented as a cumulative frequency diagram.

When comparing the landslide training dataset with landslide susceptibility maps, 77.211%, 60.420%, 57.721%, 79.160%, and 68.966% of landslides are found in the 20% of highest susceptibility classes of model A, B, C, D, and E, respectively. Related accuracy values are equal to 0.794, 0.734, 0.723, 0.686, 0.805 (Fig. 5a). On the other hand, comparing the landslide test dataset with the five landslide susceptibility models, 75.537%, 60.299%, 58.060%, 79.104%, and 66.119% of landslides are found in the 20% of highest susceptibility

classes of model A, B, C, D, and E, respectively. Correspondent accuracy values are equal to 0.788, 0.733, 0.715, 0.691, 0.801 (Fig. 5b).

Model E shows the highest accuracy (AUC = 0.800), and has hence been used as a basis for extension to all the Kyrgyz territory. The resulting landslide susceptibility map (Fig. 6) emphasizes the relatively high potential for landslides over the entire country, specifically along the eastern boarder of Fergana valley, in the South of Talas province, and in Issyk-kul district. It has to be noted that high levels of susceptibility can be found in the Naryn province where the conditions for slope failures exist though the scarce occurrence of past landslides.

7 Discussion

The analysis of results reveal that the most influential factors to slope instability, sorted according to their values of contrast, are: the class of Mesozoic materials, the class IX of seismic intensity, the class 6.6° – 16.6° of slope gradient, distance from faults greater than 10km, the class 292.5° – 22.5° of slope aspect, and the class -0.00101 − -0.00005 of profile curvature. In addition, the classes linked to the highest slope stability probabilities are: the class of Paleozoic materials, the class of slope gradient greater than 27.5º, the class VII of seismic intensity, distance from faults less than 1km, the class 157.5º – 202.5º of slope aspect, and the class -0.00005 − 0.00095 of profile curvature.

Precisely, relatively high slope values do not imply landslide occurrence, in agreement with the existing evidence (Havenith et al., 2006) that most landslides in Kyrgyzstan occur on relatively low slope angles (< 20°). Clearly, a relatively low slope angle is associated with the presence of soft materials due to their mechanical properties. Moreover, given that most settlements are located in relatively low-slope regions, an increase in the overall landslide risk assessment is to be expected.

The influence of geology is also clear, since there is an increase in the contrast values from the older geological units to the more recent ones. This agrees with the fact that, especially in the northern Kyrgyzstan (Kalmetieva et al., 2009), most landslides are found on Quaternary materials. A relatively high degree of rock fracturing might be responsible for the occurrence of slope failures in Mesozoic rocks, as revealed by the high contrast values. On the other hand, it is clear that Palaeozoic materials have no connection with landslide initiation.

At a relatively large distance from the faults (> 10km), an increase of landslide occurrences can be observed. The presence of relatively deep hypocentres might be the reason for a considerable surficial distance between fault lines and landslides, even beyond 10km. This evidence, in line with what has been previously shown (Gemitzi et al., 2011), confirms the influence of neotectonic lineaments and fault density on landsliding phenomena.

With regard to slope aspect, small contrast values clearly indicate the partial contribution of this factor in the susceptibility analysis. Nevertheless, the presence of southwest monsoon winds, might make south-facing slopes relatively wet and undisturbed, while leaving north-facing slopes drier, less vegetated, and, consequently more exposed to landslide phenomena. In agreement with this statement, similar evidences have already been provided for Central Asia (Strom, 2013).

Most of past landslides are linked to negative values of profile curvature, which correspond to convex-shaped slopes. Although having small contrast values, the control of convex morphologies to slope instability can be explained with local seismic amplification phenomena occurring in topographic ridges.

As far as the seismic input is concerned, earthquakes of magnitude around 7 are expected in order to approximate the identified seismic intensity values of VIII and IX. According to observed seismicity, it is possible to find evidence of past earthquakes with magnitude of 6.8 -

7 having an impact on the investigated areas. It has to be underlined that, despite alternative seismic input parameters are possible, in this case, the intensity assigned at the site accounts for all the possible combination of magnitude and distance determining ground shaking at the site. Therefore, for the purpose of the actual landslide susceptibility analysis, seismic intensity is for sure more reliable than other energy-related parameters like PGA.

Even though conventional landslide susceptibility analyses do not incorporate triggering information (Fell et al., 2008), the present study considers also the inclusion of the seismic input, in line with susceptibilities analyses previously presented by Schicker and Moon (2012), who included rainfall, and Holec et al. (2013), who included both seismicity and rainfall. Besides, given our focus on investigating the potential for landslide activation over a large territory, only seismicity has been considered as the triggering mechanism. While having an impact on slope instability, the effect due to precipitations might only have a local influence and, therefore, is out of the scope of this study.

An important issue in landslide susceptibility studies is represented by the influence on the final susceptibility values of transforming continuous variables into discrete variables. In this respect, Remondo et al. (2003) demonstrated how the predictive capability of validation curves obtained from input data, which were classified into only a few intervals, and the outcome from almost continuous variables, is quite similar. Based on this consideration, landslide potential factors have been classified in such a way as not to have many classes in the analysis. As a general observation, it can be stated that increasing the number of classes in landslide factors leads to unstable results. Overall, the operation of classifying susceptibility values is an ongoing topic of debate within the scientific community, given that there are no reference rules on categorizing data (Alalew et al., 2004). In this study, susceptibility values have been normalised and hence classified into 10 equal-sized sectors corresponding to

different levels of susceptibility. In this way, a quantitative comparison among landslide susceptibility maps is achievable, together with a statistically reproducible framework.

Based on the analysis of accuracy values for susceptibility models (Fig. 5), it can be seen that most of susceptibility models show AUC values greater than 0.70 and can, therefore, be accepted as significant. Model E has finally been chosen as applicable to the entire country, being the most accurate one. Given that the condition of total independency among factors in never completely verified in nature (Bonham-Carter, 1994), we have not excluded the combination of those factors known as the most relevant to landslide initiation in Kyrgyzstan, similarly to what has been carried out by Dahal et al. (2008) and Pradhan et al. (2010) in Nepal and Malaysia, respectively.

Our results demonstrate the applicability of the Weights-of-Evidence method for landslide susceptibility mapping in a data-scarce region like Kyrgyzstan. The selected study area offers a significant variability in parameters together with statistical consistency among factors, the reason why an attempt of extension to the all Kyrgyzstan has been carried out. Our method can be effectively used for detecting landslide-prone areas at the regional scale, and for identifying the most important factors inducing slope failures.

Moreover, for regional scale analyses of landslide susceptibility, overestimating the number of landslides is a common problem. In order to tackle this problem, the adoption of one single landslide point per unit area has been considered in the present study (Neuhäuser and Terhorst, 2006).

8 Conclusions

Landslides pose a serious threat to human life and human facilities. One of the preliminary steps in the direction of minimizing landslide risk is represented by susceptibility map

preparation. A landslide susceptibility map divides an area into categories ranging from more stable (or less susceptible) to less stable (or more susceptible). Moreover, areas currently free of landslides with a potential for future slope instability are identifiable. Consequently, the map can be used as a tool to support disaster management and planning activities at the regional level.

We propose an approach with the aim of preparing the landslide susceptibility map using the relationships between known landslide locations and the most significant factors to landslide activation, in high seismically-prone regions like Central Asia. As an example, we presented an application of this method - together with a sensitivity analysis - to the Jalal-Abad province in Kyrgyzstan, and we provided a landslide susceptibility map for the whole Kyrgyzstan. Results have been cross-validated by the comparison between training-control landslide locations and susceptibility values. The method, successfully applied to the data-scarce country of Kyrgyzstan, helped in identifying landslide-prone areas with greater than 70% accuracy in fitting the input data and in predicting future landslides. As a consequence, the map offers sufficient reliability measures for planning purposes and represents a basis for extension to all Central Asia.

Based on the outcomes of this study, we conclude that geology plays a critical role in guarantying slope stability. Mesozoic materials are found to be the most responsible for landslide initiation. The huge variability of these materials (soft and semi-hard rocks, principally deposits of clays, argillites, sandstones, limestones - often covered by Quaternary loess), prevents us to discriminate the influence of specific rock types, which would add more value to the analysis. The contribution of rock structure to instability is also clear given the fact that most of landslides occurred in places with relatively moderate slope gradient values (6-16°). Additionally, due to the presence of neotectonic lineaments all over the country, the

strength of rock materials is reduced and slopes are made unstable. It has been observed that, at a certain distance from faults, slope failure phenomena are quite sever. Our results show that the highest level of susceptibility to landslide activation exists along the eastern border of the Fergana valley, as well as in the south of Talas province and in the Naryn province.

In remote areas like Central Asia a complete landslide inventory is not easily achievable. Therefore, actions to minimize uncertainty in known landslide locations should be considered in order to improve the spatial correlations between landslides and causative factors. This action will in turn help in better constraining the influence of classified susceptibility values.

Furthermore, for landslide hazard evaluation purposes, the inclusion of the temporal aspect to results would be necessary. The susceptibility model might be constructed by adopting a landslides map referring to a “previous” period; the validation would subsequently be performed by using current or past landslides, provided that such temporal information about the landslides’ occurrence is available. In order to achieve this task, a more complete landslide inventory is under construction.

Considering the inclusion of seismic information, the present study may be considered as a step forward, since conventional susceptibility analyses typically do not incorporate triggering information. In Central Asia regions, the influence of precipitation on landslide initiation cannot be excluded, though it might only have local influence. Accordingly, we are currently collecting precipitation data covering the Kyrgyz country, which will later be incorporated into the statistical analysis. Moreover, in order to account for the influence of human activity on slope instability, the inclusion of environmental factors like distance from roads and land use will be part of future work.

Although landslide studies have been already started in Kyrgyzstan, a sound statistical methodology for landslide susceptibility and risk mapping applicable to the entire country

was not tested before. Thus, this paper, far from being a local investigation, provides national authorities with a regional landslide susceptibility map which will serve as a prelude for cross-border landslide risk mitigation activities.

Acknowledgements

The research presented in this paper has been carried out under the project *TIPTIMON – Tien Shan-Pamir Monitoring Program* (Project 03G0809), funded by the German Ministry of Education and Research (BMBF), and with the support from FP7 project SENSUM (grant agreement 312972). We thank Z. Kalmetyeva and A. Meleshko from the Central Asian Institute for Applied Geosciences (CAIAG) for providing the landslide data. K. Fleming kindly revised our English. Finally, the authors would like to thank the editor and two anonymous reviewers for their useful comments, which improved the manuscript.

9 References

Alalew L, Yamagishi H, Ugawa N (2004) Landslide susceptibility mapping using GIS-based weighted linear combination the case of Tsugawa area of Agano River. Niigata Prefecture Japan. Landslides 1: 73-81

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

548 Bindi D, Abdrakhmaov K, Parolai S, Mucciarelli M, Grüntal G, Ischuk A, Mikhailova N,
549 Zschau J (2012) Seismic hazard assessment in Central Asia: outcomes from a site approach.
550 Soil Dyn Earthq Eng 37: 84-91
551 Bonham-Carter GF, Agterberg FP, Wright DF (1989) Weights of evidence modelling: A new
552 approach to mapping mineral potential. Statistical applications in the Earth Sciences,
553 geological survey of Canada, Paper 89-9, 171-183
554 Bonham-Carter GF (1994) Geographic Information Systems for Geoscientists: modelling
555 with GIS Computer methods in the geosciences. Pergamon Press Oxford, vol.13, 398p
556 Carrara A, Guzzetti F, Cardinali A, Reichenbach P (1999) Use of GIS Technology in the
557 Prediction and Monitoring of Landslide Hazard. Nat Hazards 20: 117-135
558 CAC DRMI - Central Asia and Caucasus Disaster Risk Management Initiative, Desk Study
559 Review, 2009
560 Chung C-JF, Fabbri AG (2003) Validation of Spatial Prediction Models for Landslide Hazard
561 Mapping. Nat Hazards 30: 451-472
562 Dahal RK, Hasegawa S., Nonoumra A, Yamanaka M, Dhakal S, Paudyal P (2008) Predictive
563 modelling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on
564 weights-of-evidence. Geomorphology 102: 496-510
565 Dai FC, Lee CF, Ngai YY (2002) Landslide risk assessment and management: an overview.
566 Eng Geol 64: 65-87
567 Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savage WZ (2008) Guidelines for
568 landslide susceptibility, hazard and risk zoning for land use planning. Eng Geol 102: 85-98

569 Gemitzi A, Falalakis G, Eskioglou P, Petalas C (2011) Evaluating landslide susceptibility
570 using environmental factors, fuzzy membership functions and GIS. Global NEST Journal 13:
571 28-40

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

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

577 Havenith H-B, Strom A, Caceres F, Pirard E (2006) Analysis of landslide susceptibility in the
578 Suusamyr region Tien Shan: statistical and geotechnical approach. Landslides 3: 39-50

579 Holec J, Bednarik M, Šabo M, Minár J, Yilmaz I, Marschalko M (2013) A small-scale
580 landslide susceptibility assessment for the territory of Western Carpathians. Nat Hazards 69:
581 1081-1107

582 Kalmetieva ZA, Mikolaichuk AV, Moldobekov BD, Meleshko AV, Janaev MM, Zubovich
583 AV (2009) Atlas of Earthquakes in Kyrgyzstan edited by Central-Asian Institute for Applied
584 Geosciences and United Nations International Strategy for Disaster Reduction Secretariat
585 Office in Central Asia. Bishkek, p 75

586 Moldobekov B, Sarangoev A, Usupaev S, Meleshko A (1997) Prognosis of natural hazards on
587 the territory of the Kyrgyz Republic. All-Press Bishkek Kyrgyzstan, 172p (in Russian)
588 Molnar P, Tapponier P (1975) Cenozoic tectonics of Asia Effects of a continental collision.
589 Science 189: 419-426
590 Nadim F, Kjekstad O, Peduzzi P, Herold C, Jaedicke C (2006) Global landslide and avalanche
591 hotspots. Landslides 3: 159-173

592 Neuhäuser B, Terhorst B (2006) Landslide susceptibility assessment using “weights-of-
593 evidence” applied to a study area at the Jurassic escarpment (SW-Germany). Geomorphology
594 86: 12-24

595 Oh H-L, Lee S (2010) Landslide susceptibility mapping on Panaon Island, Philippines using a
596 geographic information system. Environ Earth Sci 62: 935–951

597 Petley D (2012) Global Patterns of loss of life from landslides. Geology 40 (10): 927-930

598 Pradhan B, Oh H-J, Buchroithner M (2010) Weights-of-evidence model applied to landslide
599 susceptibility mapping in a tropical hilly area. Geomatics Nat Hazards Risk 1(3): 199-223

600 Remondo J, Gonzales A, De Teran JR D, Cendrero A, Fabbri A, Chung C-JF (2003)
601 Validation of Landslide Susceptibility Maps: Examples and Applications from a Case Study
602 in Northern Spain. Nat Hazards 30: 437-449

603 Roessner S, Wetzel H-U, Kaufmann H, Samagoev A (2005) Potential of satellite remote
604 sensing and GIS for landslide hazard assessment in Southern Kyrgyzstan (Central Asia). Nat
605 Hazards 35: 395-416

606 Rosenfeld C (1994) The geomorphological dimensions of natural disasters. Geomorphology
607 10: 27-36

608 Schicker R, Moon V (2012) Comparison of bivariate and multivariate statistical approaches in
609 landslide susceptibility mapping at a regional scale. Geomorphology 161-162: 40-57

610 Soeters R, Van Westen CJ (1996) Slope stability: recognition, analysis and zonation. In:
611 “Landslides: investigation and mitigation”, Turner AK, Shuster RL (eds). Transportation
612 Research Board, Special Report 247, 129–177

SRTM Shuttle Radar Topography Mission (2004) SRTM digital topographic data US Geological Survey's EROS Data Center. [URL]: ftp://e0mss21uecsnasagov/srtm/ (2004-11-12)

Strom A (2013) Geological Prerequisites for Landslide Dams' Disaster Assessment and Mitigation in Central Asia. In: Progress of Geo-Disaster Mitigation Technology in Asia, F. Wang, M. Miyajima, T. Li, W. Shan, T. F. Fathani (eds.), 17-53

Strom AL, Korup O (2006) Extremely large rockslides and rock avalanches in the Tien Shan Mountains Kyrgyzstan. Landslides 3: 125-136

Terzaghi K, Peck RB (1967) Soil mechanics in engineering practice. Wiley and Sons, 752p

Tingdong L, Ujkenov BS, Kim BC, Tomurtogoo O, Petrov OV, Strelnikov SI (2008) Geological map of Central Asia and Adjacent Areas edt. by Geological Publishing House. Beijing China

Torgoev A, Havenith H-B (2013) Landslide Susceptibility, Hazard and Risk Mapping in Mailuu-Suu, Kyrgyzstan. In: Landslide Science and Practice, Vol. 1, C. Margottini, P. Canuti, K. Sassa (eds.), 505-510

Torgoev I, Alioshin Yu G, Torgoev A (2012) Monitoring landslides in Kyrgyzstan. In: Freiberg Online Geology, B. Merkel (Ed.), 130-139

Trifonov VG, Soboleva OV, Trifonov RV, Vostrikov GA (2002) Recent geodynamics of the Alpine-Himalayan collision belt. Transactions of the Geological Institute RAS 541, 224p (in Russian)

Van Westen CJ, Rengers N, Soeters R (2003) Use of geomorphological information in indirect landslide susceptibility assessment. Nat Hazards 30: 399-419

635 Van Westen CJ, van Asch TWJ, Soeters R (2006) Landslide hazard and risk zonation – why
636 is it still so difficult? Bull Eng Geol Environ 68: 297-306

637 Varnes JD (1984) IAEG Commission on Landslides and Other Mass Movements, Landslide
638 hazard zonation: a review of principles and practice. The UNESCO Press, Paris, 63p

639 Yilmaz I (2010) The effect of the sampling strategies on the landslide susceptibility mapping
640 by conditional probability and artificial neural networks. Environ Earth Sci 60: 505-519
641

642 **Figure captions**

643 **Fig. 1** Location map of the study area and some among the strongest past earthquakes and
644 landslides. In particular, the 1992 Suusamyr (1) and the 1946 Chatkal (2) earthquakes are
645 known to have triggered a large number of slope-failures. Among the largest landslides, the
646 1994 Osh-Jalal-Abad (3), the 2004 Aloy (4), the 2004 Kara-Sagot (5), the Chukurchak (6),
647 and the Sarychelek (7) are shown

648 **Fig. 2** Histograms of frequencies relative to classified landslide potential factors in Jalal-Abad
649 province (left) and over all Kyrgyzstan (right)

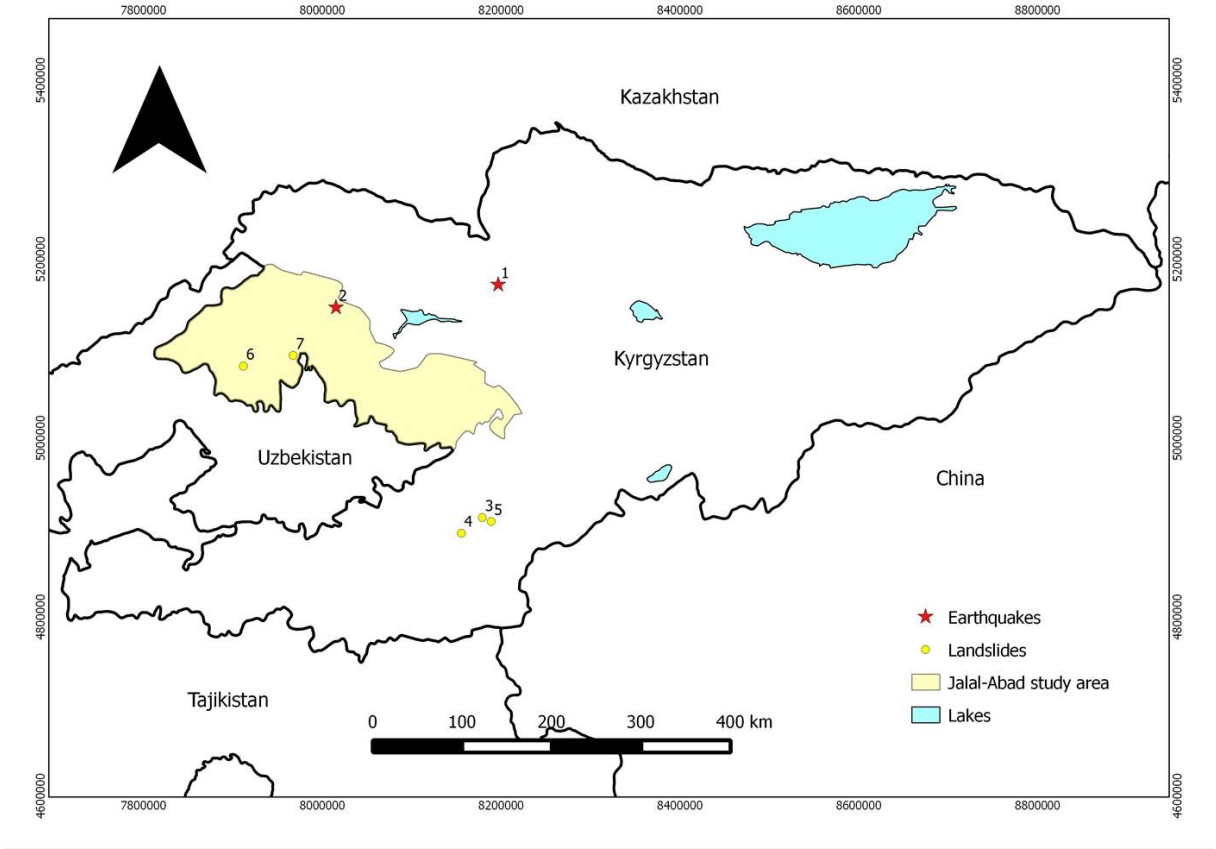
650 **Fig. 3** Landslide potential indicators for Jalal-Abad province, western Kyrgyzstan (see Fig.
651 1). (a) Slope gradient, (b) slope aspect, (c) profile curvature, (d) geology, (e) distance from
652 faults, (f) seismic intensity, and (g) landslide locations maps (all have a spatial resolution of
653 81.7m, as provided by SRTM DEM). Sources of these parameters are outlined in the text

654 **Fig. 4** Landslide Susceptibility Index (LSI) maps for the Jalal-Abad study area, Kyrgyzstan,
655 based on combinations of conditional independent factors (Model A, B, C, D), and a
656 combination of all factors (Model E). Normalised susceptibility values are shown. The yellow
657 circles indicate previous landslide locations (training dataset) (Fig. 3f)

658 **Fig. 5** Accuracy assessment of landslide susceptibility models for training (a) and test (b)
659 databases, respectively

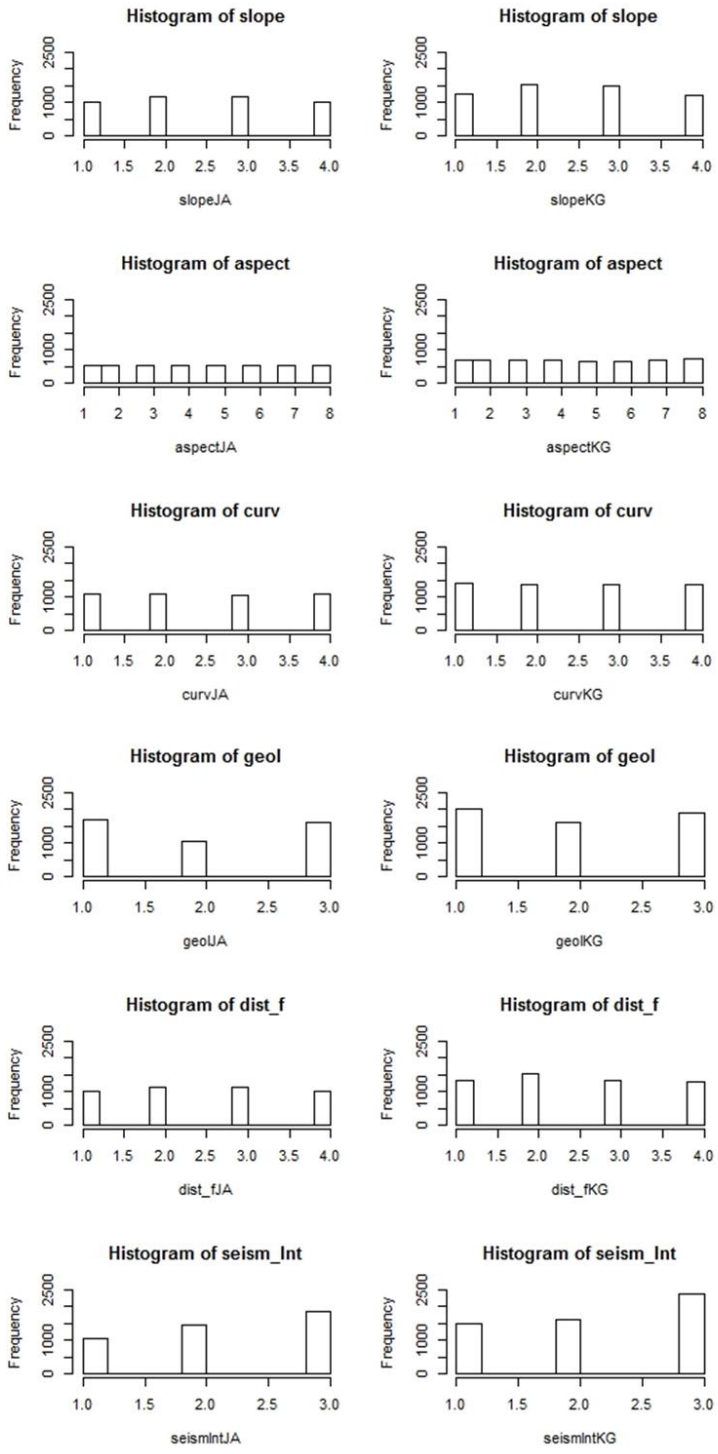
660 **Fig. 6** Landslide Susceptibility Index (LSI) maps for Kyrgyzstan calculated with respect to
661 slope gradient, slope aspect, profile curvature, geology, distance from faults, and seismic
662 intensity factors. As in Fig. 4, normalised values are shown
663

664 Fig1


7800000
8000000
8200000
8400000
8600000
8800000
5400000
5200000
5000000
4800000
4600000
Kazakhstan
Kyrgyzstan
Uzbekistan
China
Tajikistan
1
2
3
4
5
6
7
0
100
200
300
400 km
Earthquakes
Landslides
Jalal-Abad study area
Lakes

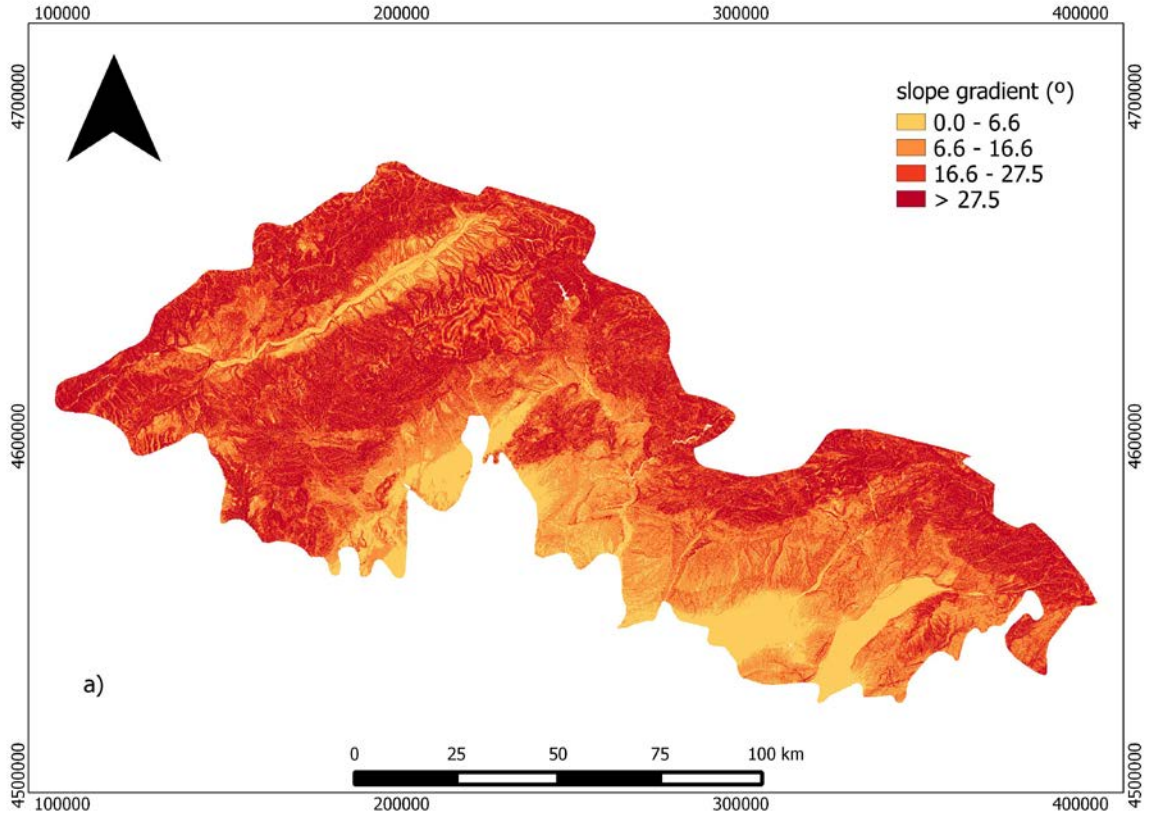
665

666 Fig2

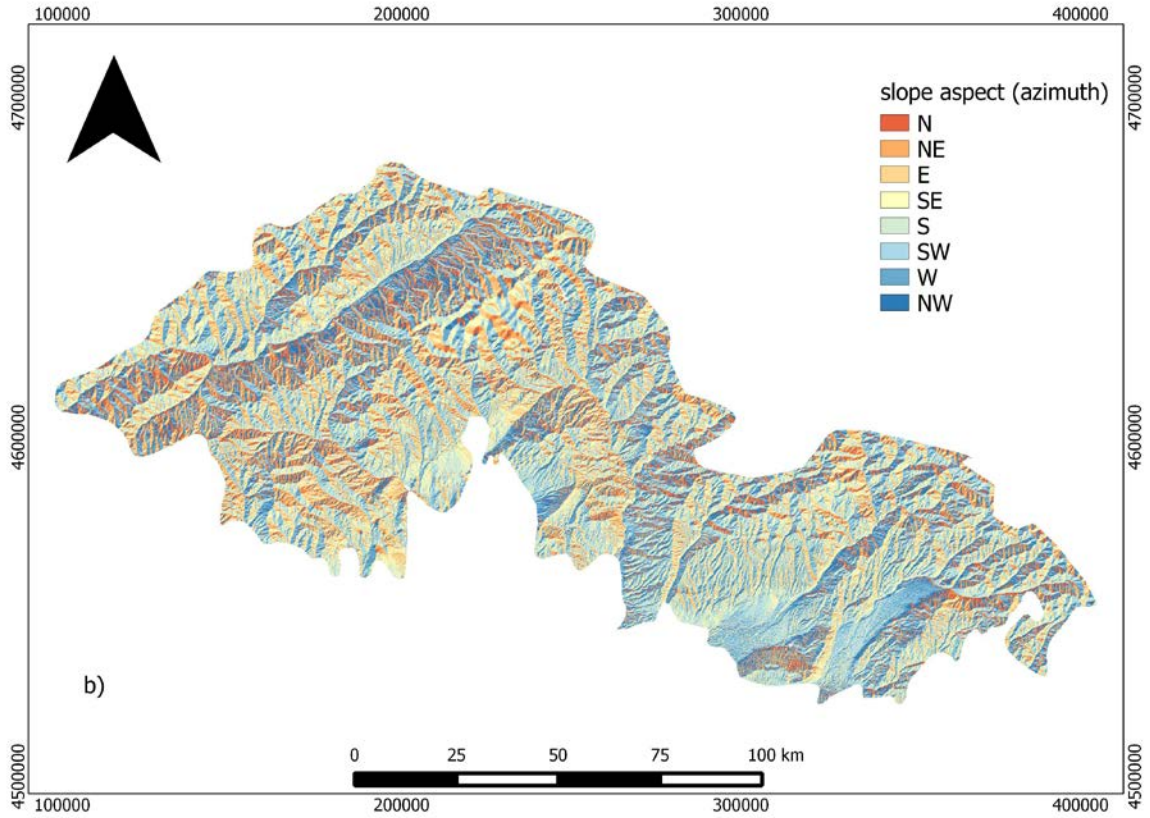

Histogram of slope
Frequency
slopeJA
Histogram of slope
Frequency
slopeKG
Histogram of aspect
Frequency
aspectJA
Histogram of aspect
Frequency
aspectKG
Histogram of curv
Frequency
curvJA
Histogram of curv
Frequency
curvKG
Histogram of geol
Frequency
geolJA
Histogram of geol
Frequency
geolKG
Histogram of dist_f
Frequency
dist_fJA
Histogram of dist_f
Frequency
dist_fKG
Histogram of seism_Int
Frequency
seismIntJA
Histogram of seism_Int
Frequency
seismIntKG

667

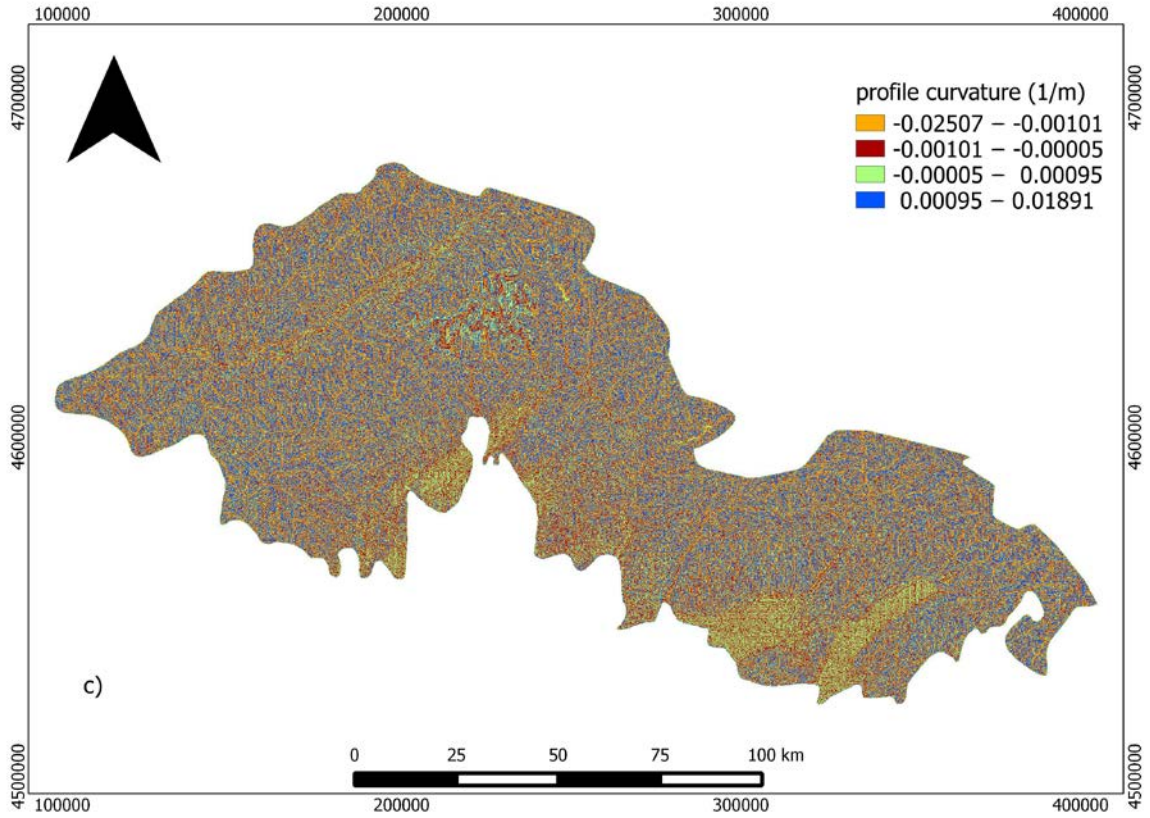
668 Fig3



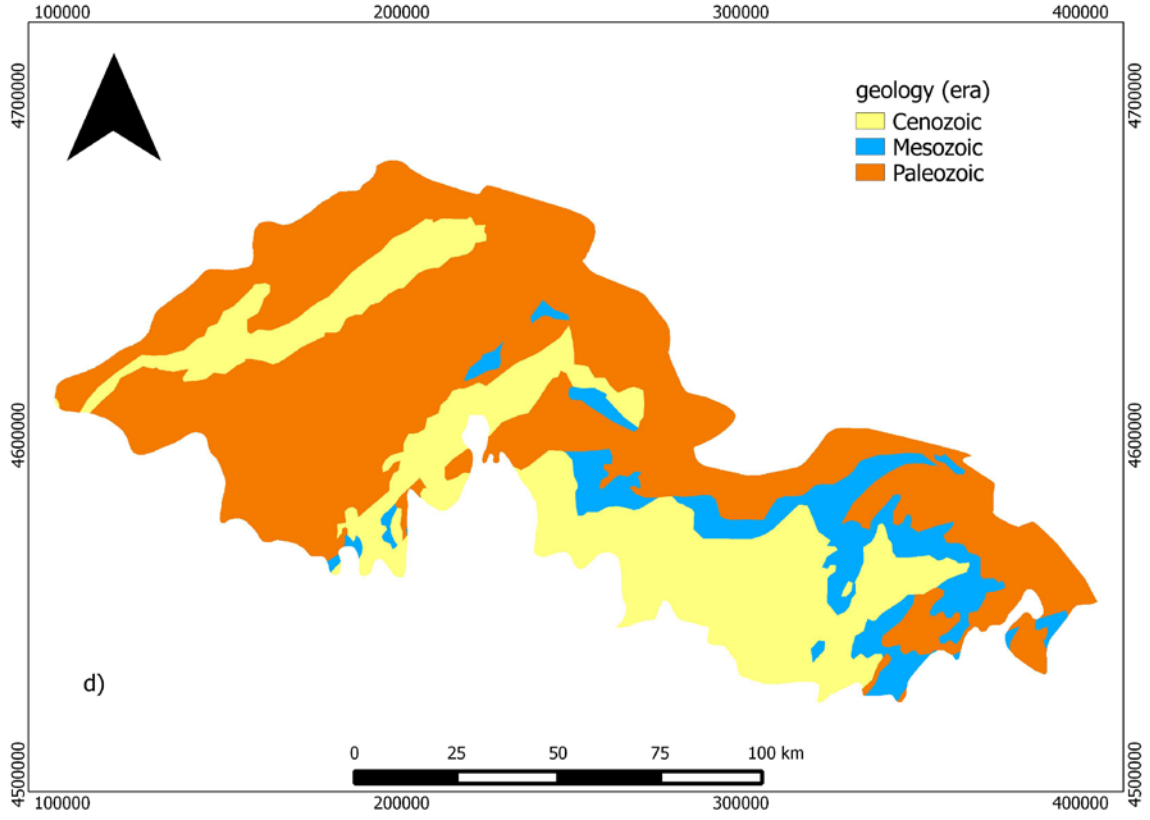
669



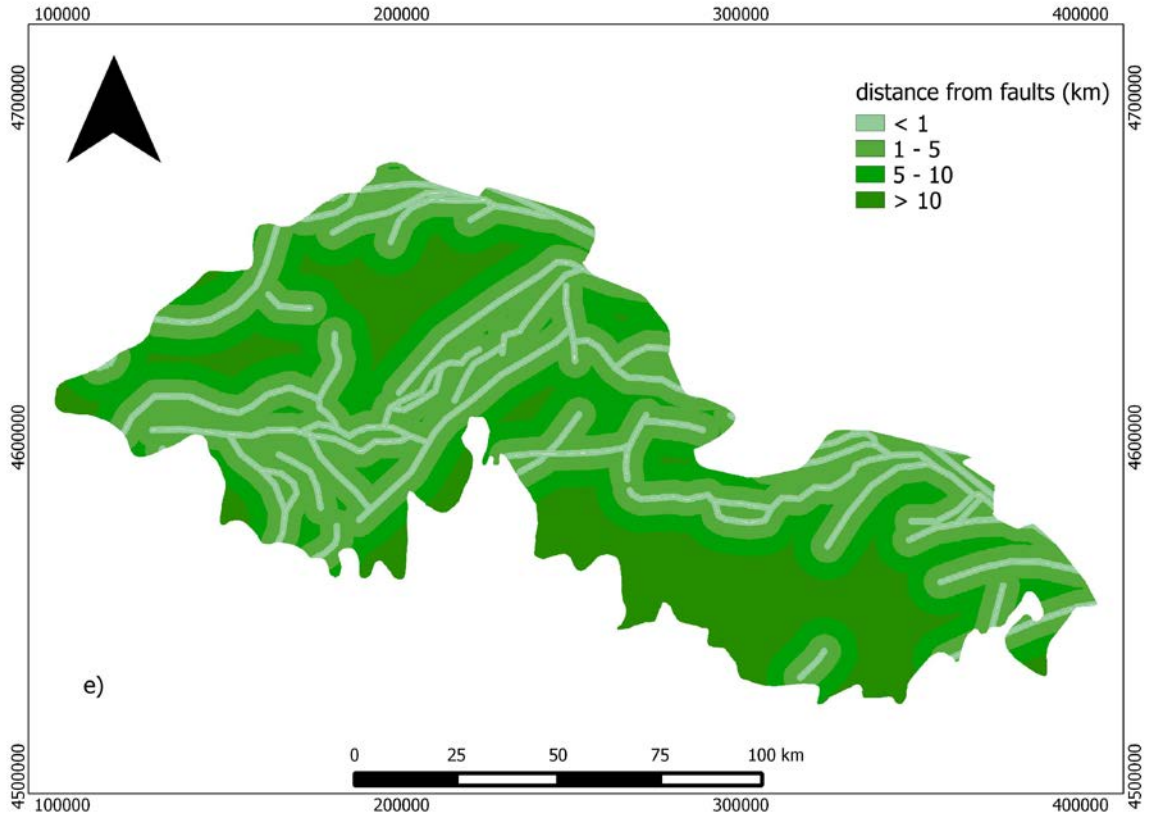
670


100000
200000
300000
400000
4700000
4600000
4500000
profile curvature (1/m)
-0.02507 – -0.00101
-0.00101 – -0.00005
-0.00005 – 0.00095
0.00095 – 0.01891
c)
0
25
50
75
100 km

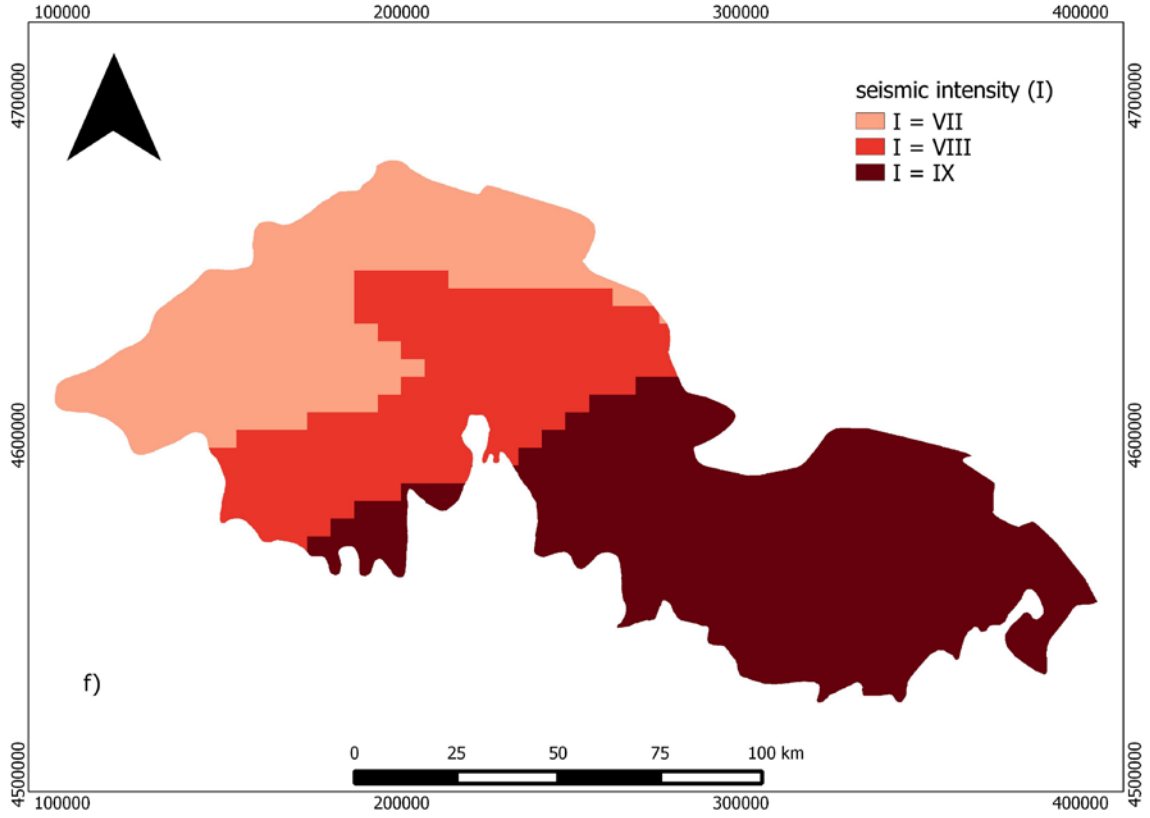
671


100000
200000
300000
400000
4700000
4600000
4500000
geology (era)
Cenozoic
Mesozoic
Paleozoic
d)
0
25
50
75
100 km

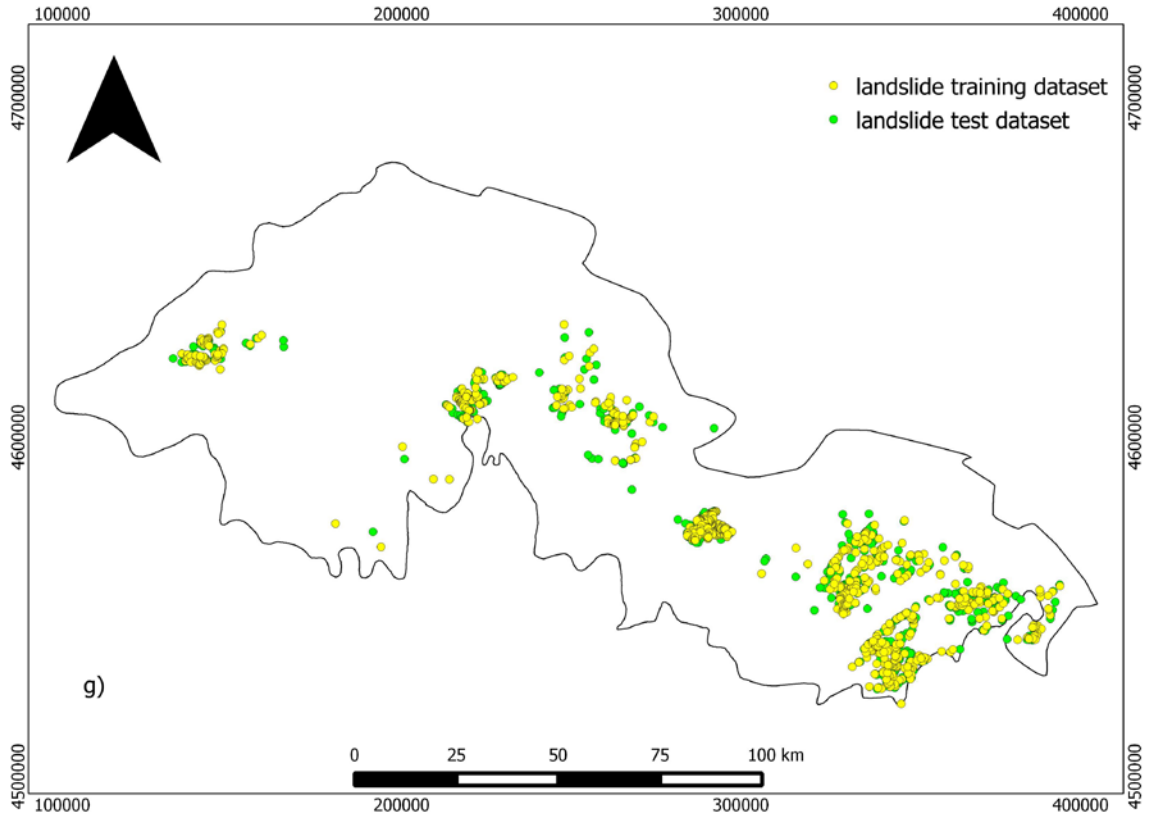
672


100000
200000
300000
400000
4700000
4600000
4500000
distance from faults (km)
< 1
1 - 5
5 - 10
> 10
e)
0
25
50
75
100 km

673


100000
200000
300000
400000
4700000
4600000
4500000
seismic intensity (I)
I = VII
I = VIII
I = IX
f)
0
25
50
75
100 km

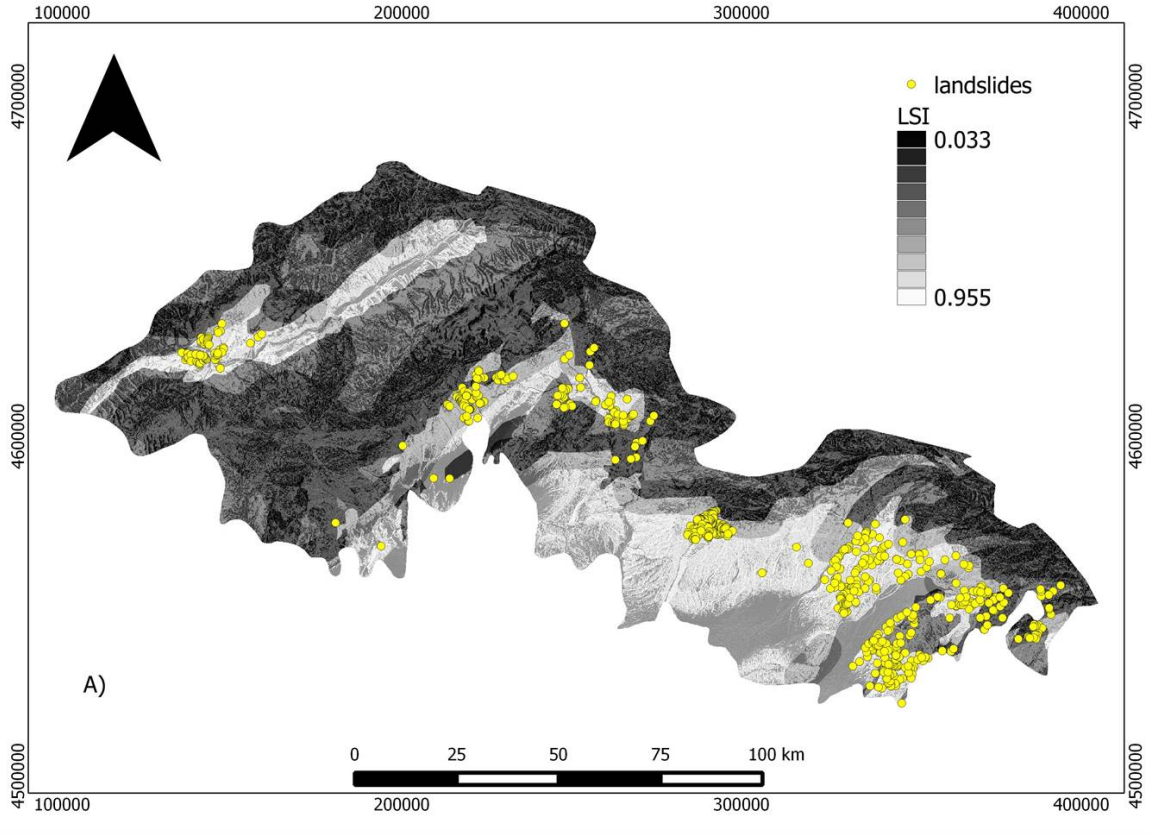
674


100000
200000
300000
400000
4700000
4600000
4500000
landslide training dataset
landslide test dataset
g)
0
25
50
75
100 km

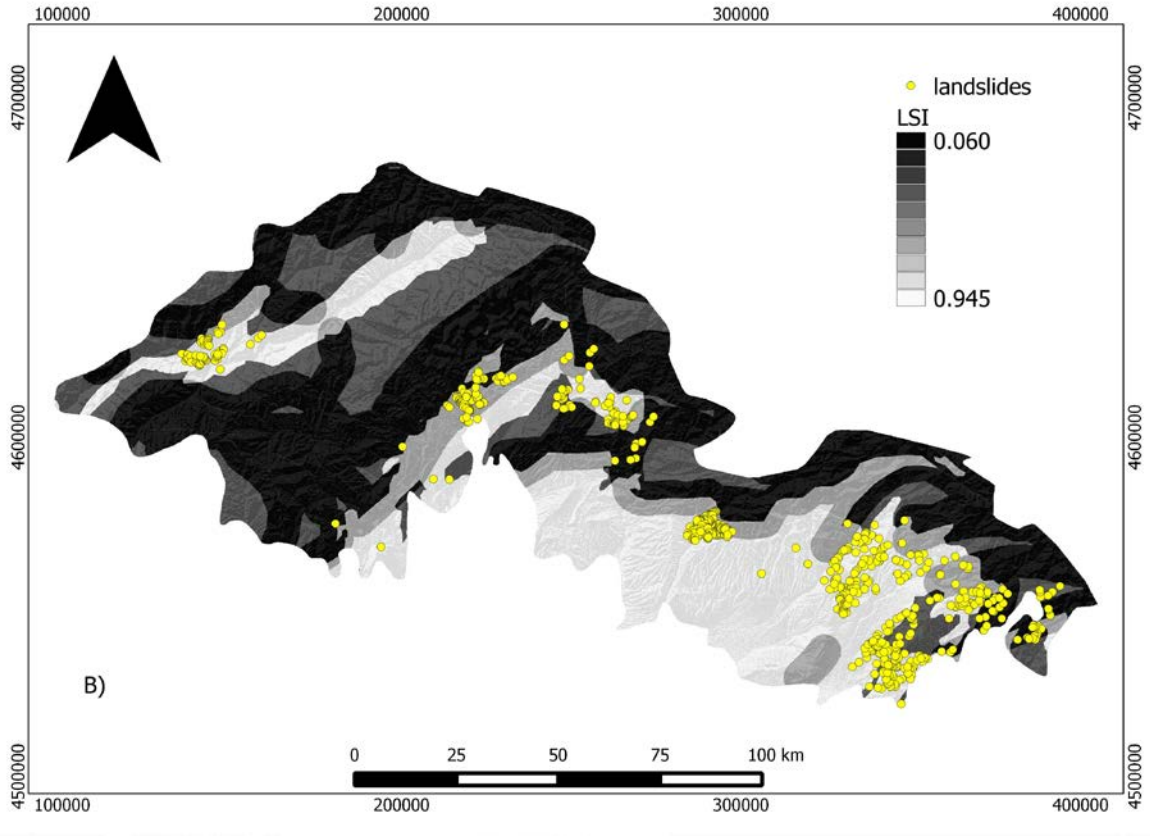
675

676

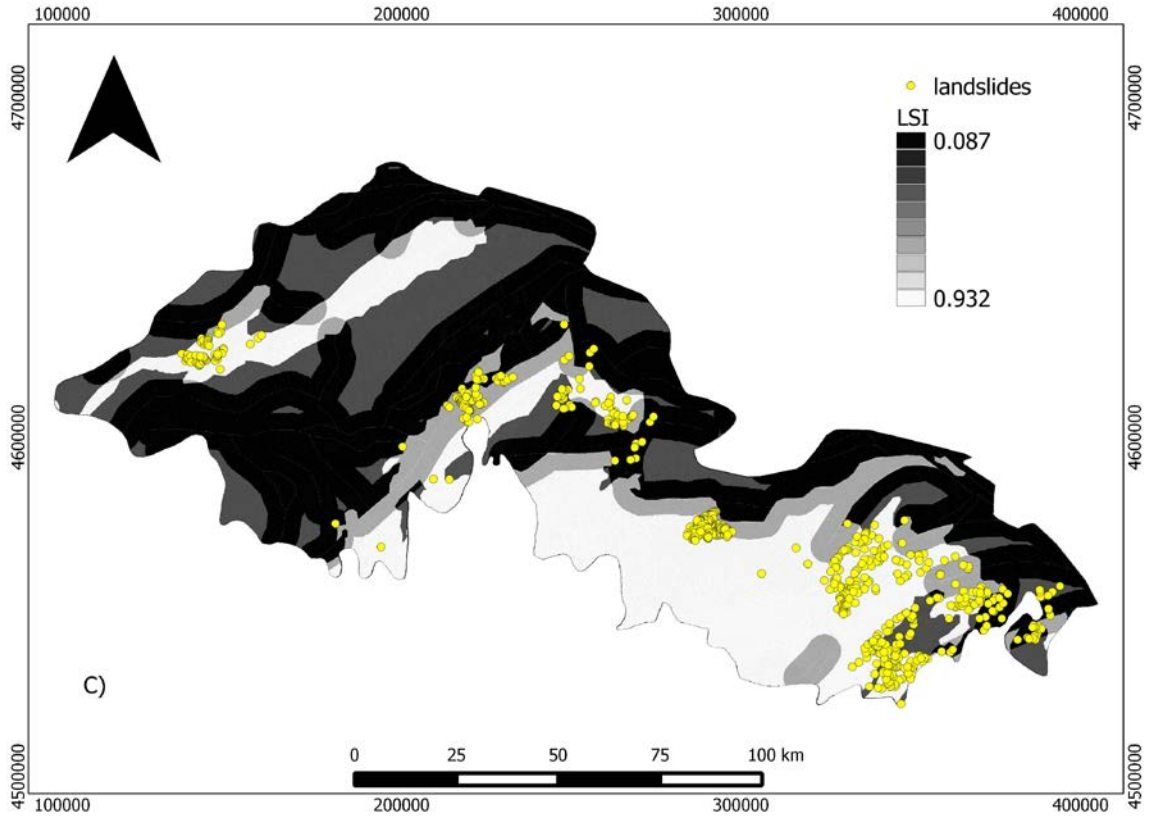
677 Fig4



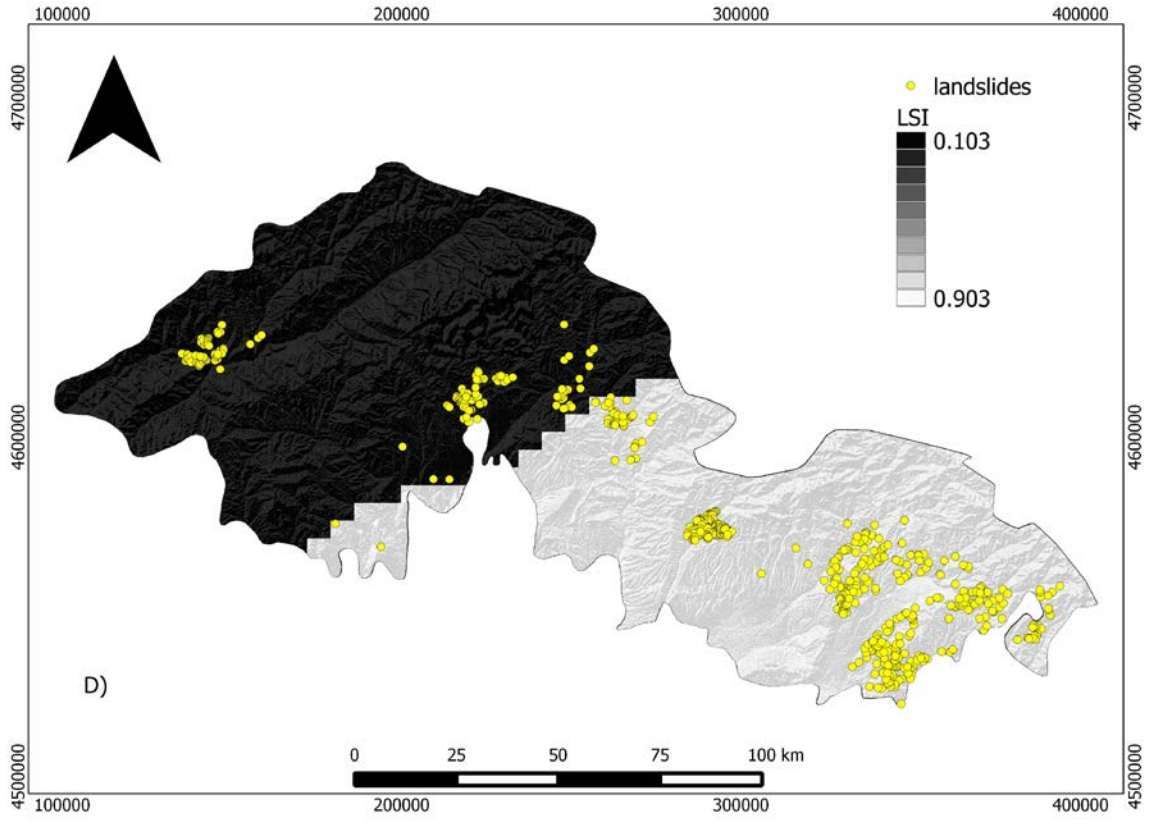
678



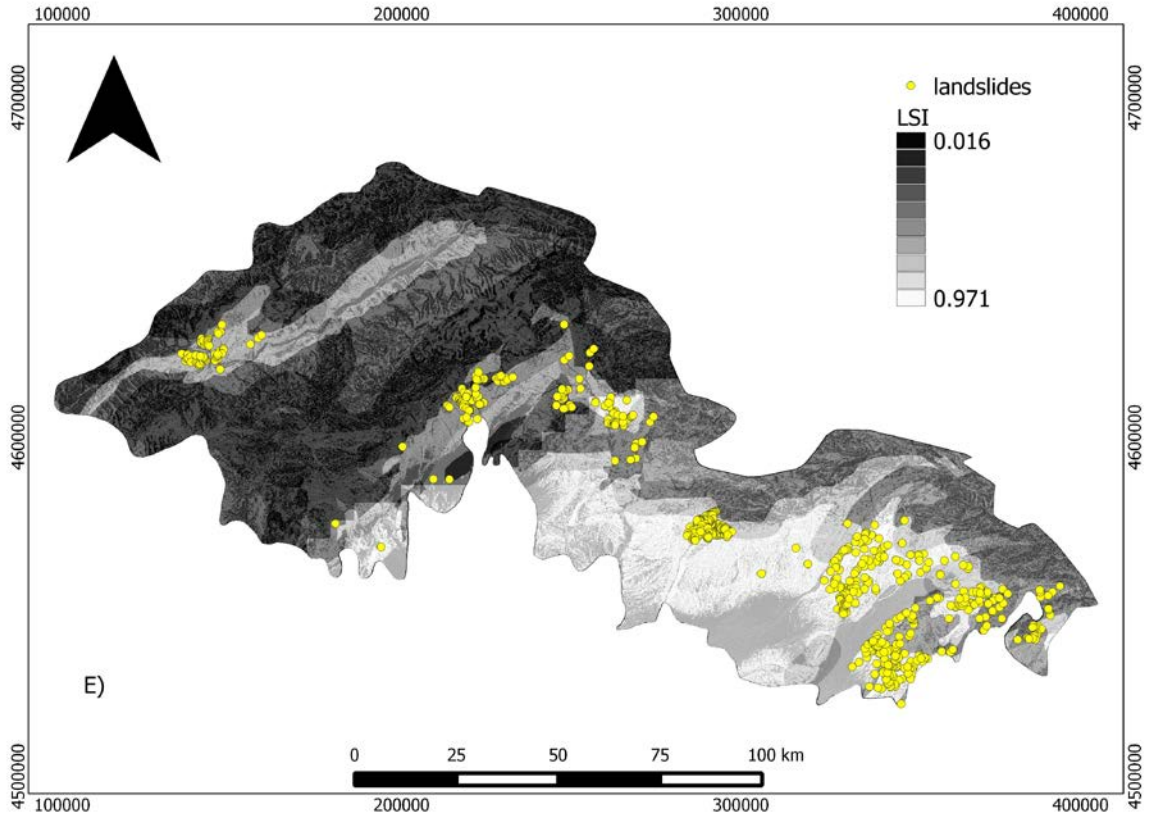
679


100000
200000
300000
400000
4700000
4600000
4500000
landslides
LSI
0.087
0.932
C)
0
25
50
75
100 km

680

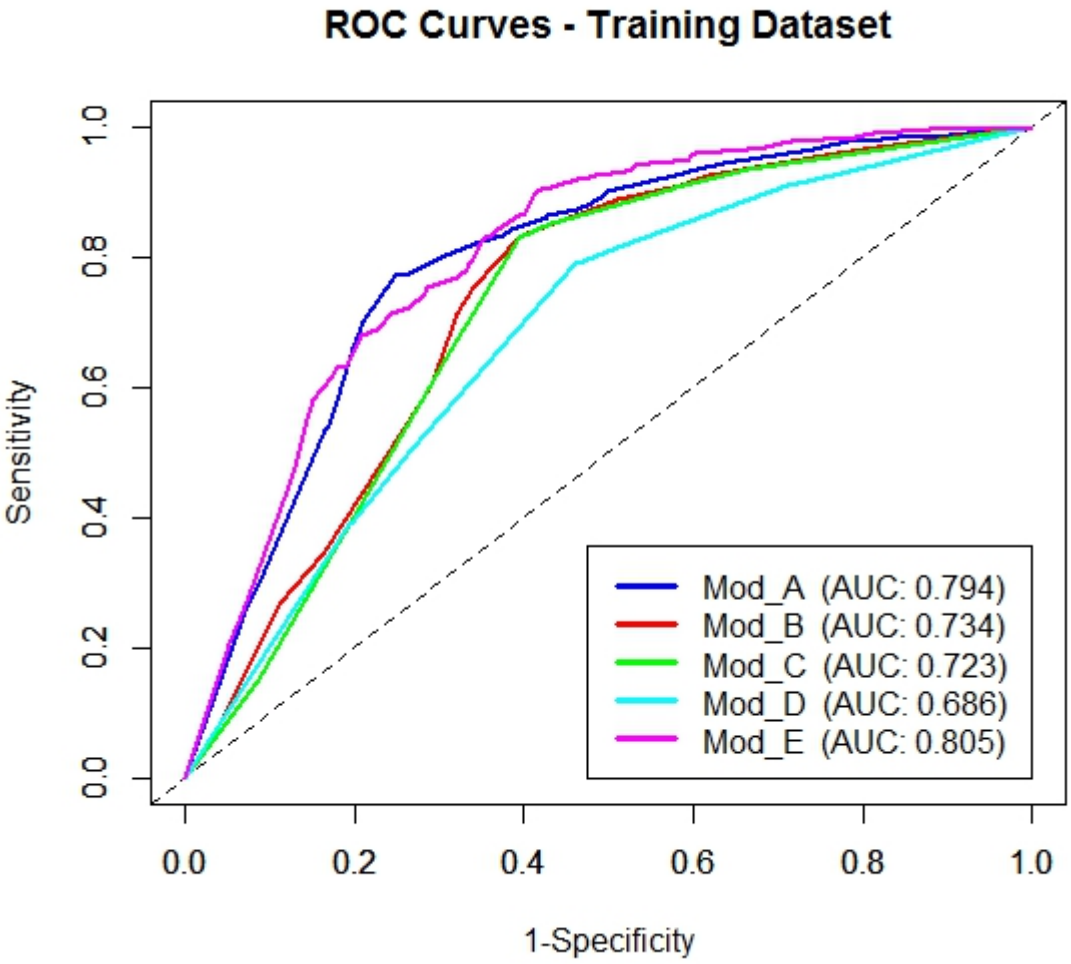

100000
200000
300000
400000
4700000
4600000
4500000
landslides
LSI
0.103
0.903
D)
0
25
50
75
100 km

681

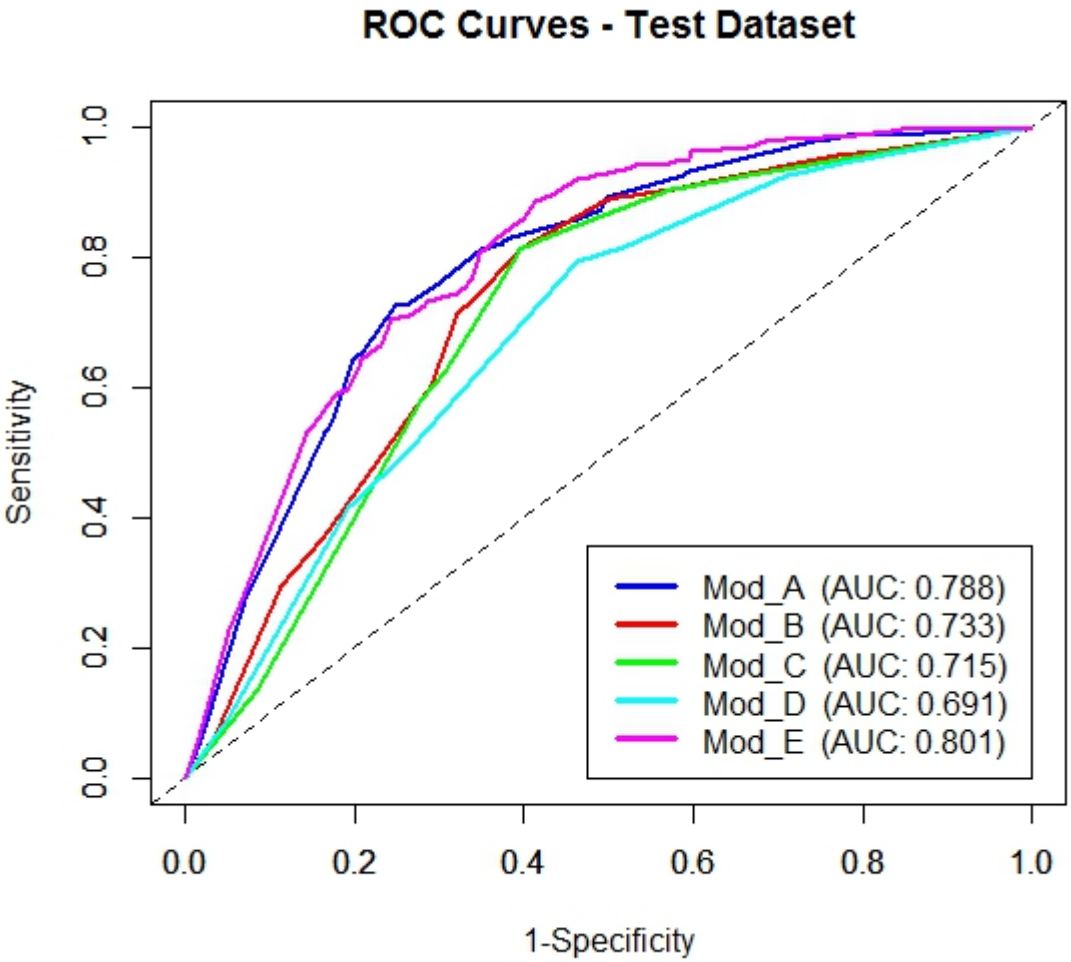

100000
200000
300000
400000
4700000
4600000
4500000
landslides
LSI
0.016
0.971
E)
0
25
50
75
100 km

682

683 Fig5


ROC Curves - Training Dataset
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
Mod_A (AUC: 0.794)
Mod_B (AUC: 0.734)
Mod_C (AUC: 0.723)
Mod_D (AUC: 0.686)
Mod_E (AUC: 0.805)

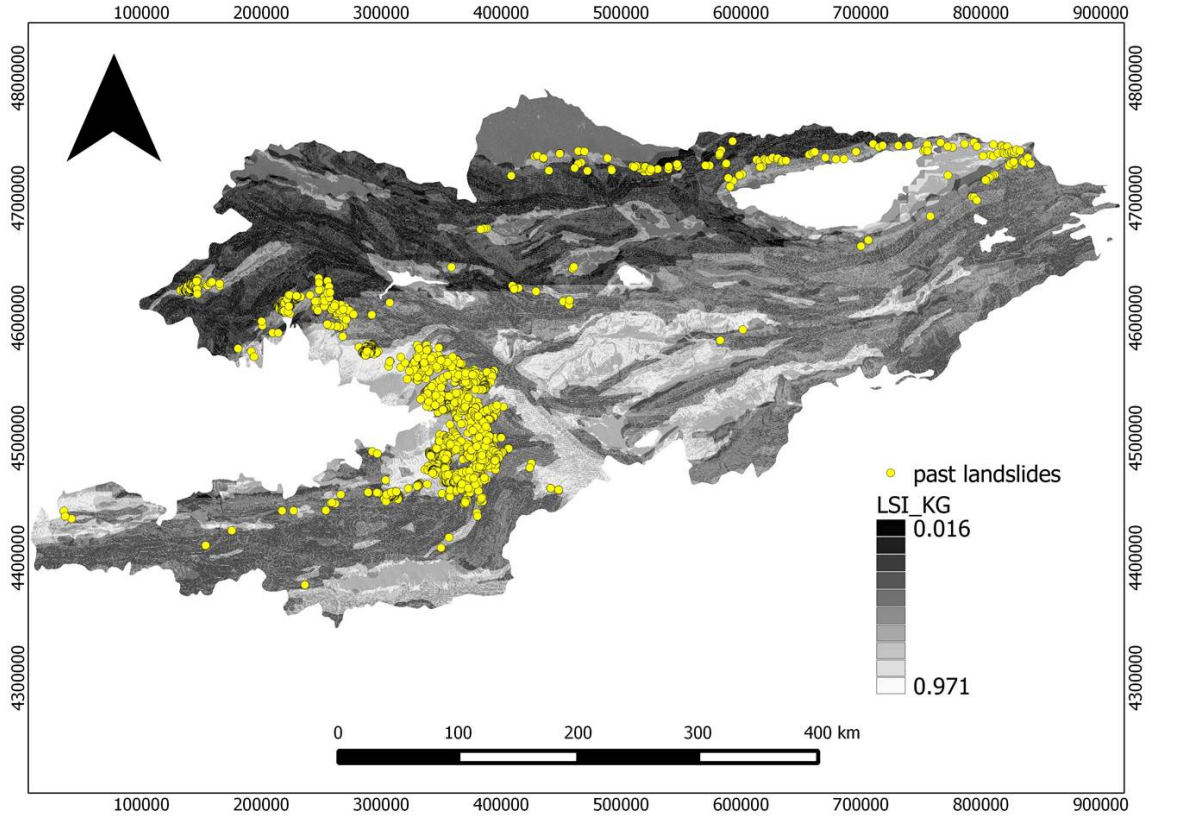
684


ROC Curves - Test Dataset
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
Mod_A (AUC: 0.788)
Mod_B (AUC: 0.733)
Mod_C (AUC: 0.715)
Mod_D (AUC: 0.691)
Mod_E (AUC: 0.801)

685

686

687 **Fig6**


100000
200000
300000
400000
500000
600000
700000
800000
900000
4800000
4700000
4600000
4500000
4400000
4300000
past landslides
LSI_KG
0.016
0.971
0
100
200
300
400 km

688

689 **Table 1** Contingency table for testing conditional independence between Factor 1 (F_1) and

690 Factor 2 (F_2)

	Factor 1Present	Factor 1 Absent	Totals
Factor 2 Present	$N\{ F_1 \cap F_2 \cap L\}$	$N\{ \overline{F}_1 \cap F_2 \cap L\}$	$N\{ F_2 \cap L\}$
Factor 2 Absent	$N\{ F_1 \cap \overline{F}_2 \cap L\}$	$N\{ \overline{F}_1 \cap \overline{F}_2 \cap L\}$	$N\{ \overline{F}_2 \cap L\}$
	$N\{ F_1 \cap L\}$	$N\{ \overline{F}_1 \cap L\}$	$N\{L\}$

691

692

693 **Table 2** Chi-square values for testing pair-wise conditional independency of all factors (99%

694 significance level). In bold conditional dependency is highlighted

	Slope	Aspect	Prof. Curv.	Geology	Dist. Faults	Seismic Intensity
Slope	-	2.943	0.064	6.310	0.600	**12.756**
Aspect		-	3.115	1.008	1.801	0.741
Prof. Curv.			-	1.466	0.003	0.817
Geology				-	0.320	**176.823**
Dist. Faults					-	**9.579**
Seismic Intensity						-

695

696

Table 3 Class, computed weights, variances and contrast values obtained from the application of the Weights-of-Evidence method to the Jalal-Abad study area

Factor / Class	total cells	landslide cells	free from landslides cells	W^+	$S^2(W^+)$	W^-	$S^2(W^-)$	C	$C / S(C)$
Slope gradient (°)									
0-6.6	560,923	36	560,887	-0.970	0.028	0.098	0.002	-1.069	-6.235
6.6-16.6	1,144,613	356	1,144,257	0.608	0.003	-0.420	0.003	1.028	13.240
16.6-27.5	1,193,560	250	1,193,310	0.213	0.004	-0.109	0.002	0.321	4.017
> 27.5	1,025,605	25	1,025,580	-1.938	0.040	0.263	0.002	-2.202	-10.801
Slope aspect (°)									
N (337.5 – 22.5)	353,555	84	353,471	0.339	0.012	-0.041	0.002	0.379	3.249
NE (22.5 – 67.5)	364,992	57	364,935	-0.081	0.018	0.008	0.002	-0.089	-0.642
E (67-5 – 112.5)	523,164	84	523,080	-0.053	0.012	0.008	0.002	-0.061	-0.524
SE (112.5 – 157.5)	535,069	84	534,985	-0.076	0.012	0.011	0.002	-0.087	-0.747
S (157.5 – 202.5)	561,148	67	561,081	-0.350	0.015	0.048	0.002	-0.397	-3.085
SW (202.5 – 247.5)	552,571	91	552,480	-0.028	0.011	0.004	0.002	-0.032	-0.287
W (247.5 – 292.5)	602,707	112	602,595	0.093	0.009	-0.018	0.002	0.111	1.068
NW (292.5 – 337.5)	426,817	88	426,729	0.197	0.011	-0.027	0.002	0.224	1.954
Profile curvature (1/m)									
-0.02507 − -0.00101	1,028,255	189	1,028,066	0.082	0.005	-0.031	0.002	0.113	1.310
-0.00101 − -0.00005	934,311	176	934,135	0.107	0.006	-0.036	0.002	0.142	1.617
-0.00005 − 0.00095	947,914	130	947,784	-0.211	0.008	0.059	0.002	-0.269	-2.757
0.00095 − 0.01891	1,028,232	172	1,028,060	-0.012	0.006	0.004	0.002	-0.017	-0.187
Geology(era)									
Cenozoic	1,161,722	203	1,161,519	0.031	0.005	-0.013	0.002	0.045	0.532
Mesozoic	391,100	351	390,749	1.668	0.003	-0.643	0.003	2.311	29.793
Paleozoic	2,382,883	113	2,382,770	-1.273	0.009	0.743	0.002	-2.016	-19.534
Distance from faults(km)									
< 1	522,534	31	522,503	-1.049	0.032	0.095	0.002	-1.144	-6.218
1 – 5	1,586,752	200	1,586,552	-0.295	0.005	0.159	0.002	-0.455	-5.378

5 – 10	970,330	205	970,125	0.221	0.005	-0.084	0.002	0.306	3.642
> 10	834,883	230	834,653	0.487	0.004	-0.185	0.002	0.671	8.241
Seismic Intensity (I)									
I = VII	1,145,958	55	1,145,903	-1.261	0.018	0.258	0.002	-1.519	-10.789
I = VIII	973,285	84	973,201	-0.674	0.012	0.149	0.002	-0.823	-7.055
I = IX	1,816,462	528	1,815,934	0.540	0.002	-0.950	0.007	1.490	15.633

699

700

Table 4 Four possible landslide susceptibility models of conditional independent factors, based on outcomes of chi-square test. Additionally, the model resulting by the combination of all factors is considered

Model A	Model B	Model C	Model D	Model E
Slope	Aspect	Prof. curvature	Aspect	Slope
Aspect	Prof. curvature	Geology	Prof. curvature	Aspect
Prof. curvature	Geology	Distance from faults	Seismic Intensity	Prof. curvature
Geology	Distance from faults			Geology
Distance from faults				Distance from Faults
				Seismic Intensity