Effects of different land use and land cover data on the landslide susceptibility zonation of road networks

Bruno M. Meneses, Susana Pereira, and Eusébio Reis
Centre for Geographical Studies, Institute of Geography and Spatial Planning, Universidade de Lisboa, Edif. IGOT, Rua Branca Edmée Marques, Lisbon, 1600-276, Portugal

Correspondence: Bruno M. Meneses (bmeneses@campus.ul.pt)

Received: 13 December 2017 – Discussion started: 18 December 2017
Revised: 12 February 2019 – Accepted: 19 February 2019 – Published: 11 March 2019

Abstract. This work evaluates the influence of land use and land cover (LUC) data with different properties on the landslide susceptibility zonation of the road network in the Zêzere watershed (Portugal). The information value method was used to assess the landslide susceptibility using two models: one including detailed LUC data (the Portuguese Land Cover Map – COS) and the other including more generalized LUC data (the CORINE Land Cover – CLC). A set of fixed independent layers was considered as landslide predisposing factors (slope angle, slope aspect, slope curvature, slope-over-area ratio, soil, and lithology) while COS and CLC were used to find the differences in the landslide susceptibility zonation. A landslide inventory was used as a dependent layer, including 259 shallow landslides obtained from the photointerpretation of orthophotos from 2005, and further validated in three sample areas. The landslide susceptibility maps were assigned to the road network data and resulted in two landslide susceptibility road network maps. The models’ performance was evaluated with prediction and success rate curves and the area under the curve (AUC). The landslide susceptibility results obtained in the two models present a high accuracy in terms of the AUC (>90 %), but the model with more detailed LUC data (COS) produces better results in the landslide susceptibility zonation on the road network with the highest landslide susceptibility.

1 Introduction

Landslides are natural processes that can constrain the free movement of people and goods when they directly or indirectly affect road networks (Bíl et al., 2014; 2015; Hilker et al., 2009; Meneses, 2011; Winter et al., 2013). The total or partial blockages of road networks have economic and societal impacts, particularly on the direct damage to the infrastructure (material damages), on the population (injuries and deaths) when driving on the affected infrastructure (Guillard and Zêzere, 2012; Pereira et al., 2014; 2017), or by causing indirect damages, such as delays, detours, material damage, and the rising prices of raw materials (Zêzere et al., 2008; Bíl et al., 2014; 2015; Jenelius and Mattsson, 2012; Winter et al., 2016).

Landslide susceptibility assessment is crucial to identifying locations with higher probabilities of landslide occurrence (Conforti et al., 2014; Guillard and Zêzere, 2012; Guzzetti et al., 2006; Pereira et al., 2014; van Westen et al., 2008). Landslide susceptibility is the likelihood of a landslide occurring in an determined area controlled by local terrain conditions; it may also include a description of the velocity and intensity of an existing or potential landslide (Fell et al., 2008; Günther et al., 2013; Guzzetti et al., 1999). Landslide susceptibility reflects the degree to which a terrain unit can be affected by future slope movements (Günther et al., 2013).

In general, the choice of landslide predisposing factors and the main details of the geographical information are not explained in a landslide susceptibility assessment based on statistical methods; rather, criteria defined in the literature (e.g., slope angle, slope aspect, slope curvature, soil, lithology, land use, and land cover) are used for this selection because they can explain the occurrence of slope movements in the study area (Blahut et al., 2010; Castella et al., 2007; Castellanos Abella, 2008; Guzzetti et al., 1999, 2006; Soeters et al., 2009; Meneses, 2011; Winter et al., 2013). The total or partial blockages of road networks have economic and societal impacts, particularly on the direct damage to the infrastructure (material damages), on the population (injuries and deaths) when driving on the affected infrastructure (Guillard and Zêzere, 2012; Pereira et al., 2014; 2017), or by causing indirect damages, such as delays, detours, material damage, and the rising prices of raw materials (Zêzere et al., 2008; Bíl et al., 2014; 2015; Jenelius and Mattsson, 2012; Winter et al., 2016).
and van Westen, 1996; van Westen et al., 2008; Zêzere et al., 2008, 2017).

Beyond the influence of different environmental factors (e.g., lithology, slope angle, slope morphology, topography, soils, and hydrology) on the spatial distribution of landslides, land use and land cover (LUC) dynamics are also an important factor on landslide susceptibility assessment (Guilllard and Zêzere, 2012). Certain land use and land cover changes (LUCCs) (e.g., deforestation, slope ruptures to road construction, steep slopes) increase the number of unstable slopes (Reichenbach et al., 2014), i.e., promoting the propensity for landslide occurrence, and can have an important impact on landslide activity (Beguería, 2006; Glade, 2003; Mugagga et al., 2012; Persichillo et al., 2017; van Westen et al., 2008).

The LUC, while a proxy variable, is very dynamic over time and is influenced by climate-driven changes and direct anthropogenic impacts (Promper et al., 2014). In this regard, it is an important predisposing factor to landslide susceptibility assessment, and Dymond et al. (2006) mention that importance: “the quality of the input land-cover map is important because the main purpose of the landslide susceptibility model is to identify where land cover needs to be changed.”

For instance, performing a landslide susceptibility analysis with a historical inventory over long periods (e.g., decades) demands the use of a permanent set of predisposing factors along the landslide inventory timeline. LUC can change over time; for this reason, it will be more accurate to use the LUC for different periods (Reichenbach et al., 2014) than using the most recent LUC map, to avoid spatial relations between past slope instability and incorrect LUC classes.

The scale of the predisposing factors directly influences the map elements’ representation and detail, as well as the choice of the scale of analysis of the final results (Leitner, 2004; Stoter et al., 2014). The choice in the level of detail will also constrain the modeling results. For example, Meneses et al. (2018b, c) obtained different LUCC results in Portugal due to the use of different LUC datasets, namely the CORINE Land Cover (CLC) and the official Land Cover Map of Portugal (Portuguese designation and acronym Carta de Ocupação do Solo, COS), with different properties concerning the scale (1 : 100 000 and 1 : 25 000, respectively), minimum mapping unit (25 and 1 ha, respectively), and generalization level (Table 1).

Due to the variation in the road network morphology (the length vs. width of the roads), the selection of appropriate data that integrate the analysis of road blockages caused by landslides requires a systematic assessment of the detailed properties of the landslide predisposing factors (Drobnjak et al., 2016; Imprialou and Quddus, 2017; Kazemi and Lim, 2005; Orongo, 2011) to obtain detailed landslide susceptibility results at the local scale (roads).

In this context, the main goal of this work is to evaluate the influence of the LUC data properties on the landslide susceptibility zonation of road networks. Two specific goals were defined: (i) to evaluate and quantify the landslide susceptibility results using two LUC datasets (CLC 2006 and COS 2007) with different properties (scale and minimum mapping unit) in two landslide susceptibility models; (ii) to use the output results of the two landslide susceptibility models to identify the sections of the main road network with the highest landslide susceptibility that will suffer future road blockages.

2 Materials and methods

2.1 Study area

This study was performed in the Zêzere watershed (5063.9 km²) located in the center region of mainland Portugal (Fig. 1). The north-northwest sector of this watershed is occupied by the Serra da Estrela, reaching a maximum elevation of 1993 m, where steep slopes can be found; in the central sector, the relief is less irregular when compared to the previous sector, but it still has steep slope areas (e.g., the vicinity of the Castelo de Bode and Cabril reservoirs); in the south-southwest sector, gentle slopes and flat areas are predominant.

The soils of the Zêzere watershed are very variable among the north-northwest, center, and southwest sectors. In the northwest sector, Cambisols predominate, with small areas of Fluvisols and eutric Lithosol along the Zêzere River. In the central area, Lithosols are dominant, with some areas of Cambisols. In the south-southwest sector, there are areas of Lithosols intercalated with Cambisols and Luvisols.

According to CLC 2006, the predominant types of LUC in the study area are forest and seminatural areas, which represent 72 % of the watershed area. Other LUC types are less representative, for example, agricultural land (25.5 %), artificialized land–urban areas (1.5 %), and water bodies (1 %), including an important freshwater reservoir, the Castelo de Bode dam (Meneses et al., 2015a). The LUC of this watershed is very dynamic, highlighting the LUCC in forest and agricultural areas derived from multiple socioeconomic driving forces (Meneses et al., 2017) and the degradation of vast forest areas by wildfires (Meneses et al., 2018a).

Due to the large extension of this watershed, three sample areas were selected according to the high density of landslides observed in these locations: the Serra da Estrela, Vila de Rei, and Ferreira do Zêzere municipalities (areas of 86.7, 191.5, and 190.4 km², respectively), where fieldwork was developed to validate part of the landslide inventory and the disruption of roads caused by landslides.

2.2 Data

The landslide predisposing factors used to model the landslide susceptibility in the Zêzere watershed were selected after reviewing the literature about the causal factors of landslide occurrence (Blahut et al., 2010; Castella et al., 2007;
Table 1. Properties of LUC data.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Land cover maps of Portugal</th>
<th>CORINE land cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acronym</td>
<td>COS</td>
<td>CLC</td>
</tr>
<tr>
<td>Scale</td>
<td>1 : 25,000</td>
<td>1 : 100,000</td>
</tr>
<tr>
<td>Minimum mapping unit</td>
<td>1 ha</td>
<td>25 ha</td>
</tr>
<tr>
<td>Data structure</td>
<td>Vector</td>
<td>Vector</td>
</tr>
<tr>
<td>Geometry</td>
<td>Polygons</td>
<td>Polygons</td>
</tr>
<tr>
<td>Minimum distance between lines</td>
<td>20 m</td>
<td>100 m</td>
</tr>
<tr>
<td>Base data</td>
<td>Orthophotos</td>
<td>Satellite images</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>0.5 m</td>
<td>20 m</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>Hierarchical (five levels)</td>
<td>Hierarchical (three levels)</td>
</tr>
<tr>
<td></td>
<td>225 classes</td>
<td>44 classes</td>
</tr>
<tr>
<td>Production method</td>
<td>Visual interpretation</td>
<td>Semiautomated production and visual interpretation</td>
</tr>
<tr>
<td>Date of production</td>
<td>2007</td>
<td>2006</td>
</tr>
</tbody>
</table>

Castellanos Abella, 2008; Guzzetti et al., 1999; Reichenbach et al., 2018; Soeters and van Westen, 1996; van Westen et al., 2008; Zêzere et al., 2008, 2017) (Fig. 2).

Six fixed landslide predisposing factors were considered: slope angle, slope aspect, slope curvature, slope-over-area ratio (SOAR), soil, and lithology. The LUC types of COS and CLC were used to find the differences in the landslide susceptibility zonation. The set of landslide predisposing factors and the corresponding classes (Fig. 2) were the same in all models, only changing the LUC data.

In general terms, an increasing slope angle promotes landslide occurrence and is a very good proxy of the shear stress (Zêzere et al., 2017). Slope instability is more frequent at the higher slope angles of the Serra da Estrela and throughout the Zêzere River valley. Also, in these areas, convex slope curvature is predominantly related to slope instability. The slope aspect is important in the spatial distribution of the different LUC types of the study area (Fig. 2) and in slope instability, especially in northwest-facing slopes (more exposed to rain and with higher humidity levels).

The SOAR is a proxy variable of the moisture retention, the soil water content, and the surface saturation zones (Zêzere et al., 2017), highlighting, in the Zêzere watershed, the upstream (very close to the Zêzere River) and southwest areas with a higher SOAR.

In the sample areas of the Vila de Rei and Ferreira do Zêzere municipalities, where a high landslide density was observed, schist and metasedimentary lithologies are predominant. Further, slope instability in the watershed is higher in the hortic Luvisols and in the LUC classes of forest and shrubland or herbaceous vegetation associations (Fig. 1).

The official LUC data available for the study area are CLC produced by the European Environment Agency (EEA) and COS produced by the General Directorate for Territorial Development (DGT) in Portugal. These LUC data (CLC and COS) have different properties and have been used in several studies about landslides in Portugal (e.g., Guillard and Zêzere, 2012; Meneses et al., 2015b; Piedade et al., 2011; Reis et al., 2003; Zêzere et al., 2017).

Table 1 describes the main properties of these LUC data (DGT, 2013; EEA, 2007; IGP, 2010). Among the differences between the two LUC datasets, the scale is highlighted because COS is the most detailed relative to CLC (proportion 1/4). However, the properties are not proportional between the two LUC datasets; while the COS features have a minimum mapping unit of 1 ha, CLC has a minimum mapping unit of 25 ha, and the minimum distance between lines is 20 m in COS, while in CLC it is 100 m.

To reduce possible discrepancies in the field, the LUC data were collected for near dates: CLC 2006 and COS 2007. The LUC data were developed with base information that matches in temporal terms, for example, the satellite images, orthophotos, and agricultural and forestry inventories used as auxiliary information. The nomenclature of these LUC data corresponds to the third level (see the official CLC nomenclature on the EEA website). In this study, the second level of the CLC nomenclature was used because it has a lower number of classes for the study area (12 of 31 classes).

The agreement among the LUC data is presented in Table 2. The forest class shows great differences between the two LUC datasets. For example, COS represents more forest area relative to CLC (34% and 26.9% of the study area,
respective) because a part of COS (approximately 10% of the study area) is classified as scrub and/or herbaceous vegetation associations in CLC. The reverse was also verified: approximately 5% of the study area is classified as scrub and/or herbaceous vegetation associations in COS, and this same area is represented by forest class in CLC. These discrepancies are derived from the LUC data properties because COS is more detailed and represents more degraded forest areas, especially where wildfires occurred. These events affected a large percentage of the watershed (Meneses et al., 2018a), especially the central sector, as a vast burned area culminated in a large transition of forest area to shrubland.

The forest, scrub and/or herbaceous vegetation associations and open spaces with little or no vegetation are the LUC types predominant in the hillsides with steep slopes (see Tables S1 and S2 in the Supplement). The remaining LUC classes present more area in the lower slopes (>10°).

The soil and lithology data were obtained from the environment atlas web platform published by the Portuguese Environment Agency (APA) at a 1:1000000 scale. A digital elevation model (DEM) was built using digital topographic maps at a 1:250000 scale (IGEOE), containing contour lines with 10 m equidistance.

Slope angle, slope aspect, slope curvature, and SOAR (topographic wetness index) layers were extracted from the DEM. Road network data (vector lines) were extracted from Portugal’s military cartography (itinerary maps, 1:500000 scale), available on the Portuguese Army Geospatial Information Center’s website. The road network was classified according to the roads’ width and their network hierarchy. Considering the road center line, a buffer of 5 m was defined for municipal roads, 10 m for complementary roads, and 20 m for superhighways. These distances were measured with geographic information systems (GIS) on the study area roads (directly on the orthophotos).

The landslide inventory was obtained using photointerpretation (orthophotos from 2005 and Google Earth images), a process supported by the ancillary topographic data and further fieldwork validation only performed in the sample areas (Fig. 1) due to the extension of the study area. A total of 128 landslides (predominantly shallow translational slides), with a total area of 74 042 m², were validated during fieldwork in the sample areas (49.4% of the total inventoried landslide cases). Among the landslides initially inventoried by photointerpretation in the sample areas, more than 90% of cases
Figure 2. Predisposing factors used in the landslide susceptibility assessment. Predisposing factor map legend. Curvature – Cv: convex; St: straight; Cc: concave. Lithology – A: alluvium; ACLD: arenites, conglomerates, limestones, dolomitic limestone; ACLM: arenites, conglomerates, limestones, dolomitic limestone and marl; ALSC: arenites, limestone, sand, stony banks and clay; CGA: clayey schist, graywackes and arenites; CALD: conglomerates, arenites, limestones, dolomitic limestone, marly limestone and marl; CALM: conglomerates, arenites, white limestone and red marl; G: gabbro; GD: glacial deposits; GS: granite and other stones; GP: granite porphyritic; LDM: limestones, dolomitic limestone, marly limestone and marl; Q: quartzite; RCMD: red sandstone, conglomerates, marl and dolomitic limestones; SG: sands and gravel; SRAC: sands, rocky, arenites and clay; SG: schists and graywackes; SGC: schist and graywacke complex; SAMQ: schists, amphibolite, mica schists, quartztite graywackes, carbonated stones and gneisses. Soil – HC: humic Cambisols; R: rankers; DC: dystric Cambisols; DF: dystric Fluvisols; EL: eutric Lithosol; CC: calcic Cambisols; CL: calcic Luvisols; HL: hortic Luvisols; Ch: Chromic Cambisols; EC: eutric Cambisols; CcC: calcic–chromic Cambisols; HP: hortic Podzols; EF: eutric Fluvisols. LUC – UF: urban fabric; ICT: industrial, commercial and transport units; MDC: mine, dump and construction sites; ANA: artificial, nonagricultural vegetated areas; AL: arable land; PC: permanent crops; P: pastures; HAA: heterogeneous agricultural areas; F: forests; SHV: scrub and/or herbaceous vegetation associations; OSV: open spaces with little or no vegetation; IW: inland waters.
Table 2. LUC data agreement (area in hectares) between CLC and COS classes.

<table>
<thead>
<tr>
<th>CLC Data</th>
<th>Urban Commercial</th>
<th>Industrial, Mine, dump, Artificial</th>
<th>Arable</th>
<th>Permanent Pastures</th>
<th>Heterogeneous Forests</th>
<th>Scrub</th>
<th>Open Inland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>COS</td>
<td>UF</td>
<td>ICT</td>
<td>MDC</td>
<td>ANA</td>
<td>SHV</td>
<td>OSV</td>
<td>IW</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1062</td>
<td>2688</td>
<td>1214</td>
<td>5640</td>
<td>3091</td>
<td>691</td>
<td></td>
<td>14249</td>
</tr>
<tr>
<td></td>
<td>4908</td>
<td>3698</td>
<td>1214</td>
<td>5640</td>
<td>3091</td>
<td>691</td>
<td></td>
<td>14249</td>
</tr>
<tr>
<td></td>
<td>7.3</td>
<td>10.6</td>
<td>1.7</td>
<td>15.6</td>
<td>8.1</td>
<td>1.2</td>
<td></td>
<td>31.3</td>
</tr>
<tr>
<td></td>
<td>101</td>
<td>202</td>
<td>1.9</td>
<td>0.5</td>
<td>12</td>
<td>0.1</td>
<td></td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>3160.2</td>
<td>439.8</td>
<td>77.3</td>
<td>100.8</td>
<td>207.7</td>
<td>502</td>
<td>15.7</td>
<td>337</td>
</tr>
<tr>
<td></td>
<td>4.6</td>
<td>650.4</td>
<td>83.0</td>
<td>9.5</td>
<td>33.4</td>
<td>27.4</td>
<td>9.0</td>
<td>1308</td>
</tr>
<tr>
<td></td>
<td>250.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>225.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>274.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>274.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>274.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>274.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Figure 3. Landslide size frequency distribution.

were confirmed. In these sample areas, road disruptions were also validated.

For the complete Zêzere watershed, 259 landslides were identified, predominantly of shallow type. Of the total, 32 landslides directly affected the road network (total or partial blockages by the material and seven cases with partial loss of infrastructure). The landslide inventory was randomly divided into two subsets (Fig. 1) (Chung and Fabbri, 2003): the landslide training group and the landslide test group (81.5% and 18.5% of the total landslide affected area, respectively). The statistical description of each landslide group is presented in Table 3.

The landslide size frequency distribution is different between the landslides that affected the road network and those that did not (Fig. 3). The area of the majority of landslides ranges between 101 and 200 m$^2$, while most of the landslides that affected the road network present a larger area (>1000 m$^2$).

All the predisposing factors and landslide inventory were converted to raster (resolution 10 m) to assess the landslide susceptibility. The selection of the predisposing factors' cell size was based on several geoinformation conversion tests in the Zêzere watershed previously performed by Meneses et al. (2016, 2018b).

2.3 Methods

The landslide susceptibility modeling was carried out using the information value (IV) method (Yan, 1988; Yin and Yan, 1988). The IV method is a bivariate statistical method that has been used in several studies and different areas with good results for landslide susceptibility assessment (e.g., Guillard and Zêzere, 2012; Oliveira et al., 2015a; Zêzere et al., 2017).
The IV of each class within each explanatory variable is given by Eq. (1) (Yan, 1988; Yin and Yan, 1988):

\[ IV_{x_i} = \ln \frac{S_i/N_i}{S/N}, \]

where \( IV_{x_i} \) is the IV of the variable \( x_i \), \( S_i \) is the number of terrain units with landslides and the presence of variable \( x_i \); \( N_i \) is the number of terrain units with variable \( x_i \), \( S \) is the total number of terrain units with landslides, and \( N \) is the total number of terrain units.

The IV method was applied in several landslide susceptibility zonation studies, providing good results (e.g., Che et al., 2012; Chen et al., 2016; Conforti et al., 2012) at the regional scale. This method was also applied in several studies conducted in Portugal, with good performance in susceptibility assessment (e.g., Guillard and Zêzere, 2012; Oliveira et al., 2015b; Pereira et al., 2014; Zêzere et al., 2017).

The a priori probability of finding a landslide unit in the study area \((S/N)\) and conditional probabilities for each class of the independent variables \((S_i/N_i)\) were calculated, obtaining the IV for these classes. However, the IV method presents constraints on obtaining the natural logarithm for negative results; in this case, the lower value calculated for each variable was assigned to classes when \( S_i \) was equal to zero.

The IVs of all the variables were combined to obtain the landslide susceptibility map \((LSM)\). For the final landslide susceptibility assessment, i.e., the integration of the IVs of all the independent variables, the following equation was considered:

\[ IV_j = \sum_{i=0}^{n} X_{ij} I_i, \]

where \( IV_j \) is the total IV of the cell \( j \), \( I_i \) is the information value of each cell of each independent variable, \( n \) is the number of variables, and \( X_{ij} \) assumes the value 1 or 0, depending on the presence or absence of the variable in the terrain unit.

Landslide susceptibility model performance was assessed using training landslides. Landslide areas in the test group were only used to perform an independent validation of the landslide susceptibility. Prediction rate curves \((PRCs)\) were computed for each final LSM (Chung and Fabbri, 1999, 2003) and also the area under the curve \((AUC)\). Success rate curves \((SRCs)\) were obtained for the landslide susceptibility road network maps using only the landslides that affected roads.

The importance of each independent variable in the landslide susceptibility assessment was also determined, so that the spatial influence of each predisposition factor in the models can be understood. The accountability \((A_I)\) and reliability \((R_I)\) indexes have been used in different contexts to assess the importance of each independent variable in the bivariate statistical methods (e.g., Blahut et al., 2010; Meneses et al., 2016). \( A_I \) explains how different classes of predisposition factors are relevant in the analysis because they contain the landslide area, while \( R_I \) depends on the average density of the landslide area in the predisposing factor classes that are more relevant to the development of this process. In this procedure, the \( A_I \) and \( R_I \) were determined using Eqs. (3) and (4), respectively (Blahut et al., 2010).

\[ A_I = \frac{\sum_{i=1}^{n} k}{N} 100 \]  

\[ R_I = \frac{\sum_{i=1}^{n} k}{\sum_{i=1}^{y} y} 100 \]

Here \( k \) is the landslide area in classes with the conditional probability values higher than a priori probability, \( N \) is the total landslide area, and \( y \) is the area of each class of independent variable with a conditional probability above the a priori probability.

Two landslide susceptibility models were built using the IV method (see results in Table S3), using the same set of predisposing factors, except the LUC data (Fig. 4): model 1 \((M1)\) was modeled with COS 2007 and resulted in landslide susceptibility map 1 \((LSM1)\); model 2 \((M2)\) was modeled with CLC 2006 and resulted in landslide susceptibility map 2 \((LSM2)\). LSM1 and LSM2 were correlated, and the corresponding spatial agreement was analyzed.
Information values of LSM1 and LSM2 were assigned to the road network (using GIS), resulting in a road network map with the landslide susceptibility location (landslide susceptibility of the road network – LSRN1 and LSRN2), where there is a higher spatial probability of road interruption or road interference caused by landslides. Different outputs of the two models (road network) were compared using the overall agreement and kappa coefficient (Congalton and Green, 2009), allowing the assessment of the consistency and agreement of the obtained results with different LUC datasets. The information of road disruptions caused by landslides was used to validate these results.

Landslide susceptibility maps were built and classified in 10 classes (deciles) containing an equal number of terrain units to allow visual comparison of the results.

3 Results

3.1 Landslide susceptibility

The landslide susceptibility results show spatial contrasts in the study area. Some areas in the center of the watershed (highlighting the vicinity of the Castelo de Bode reservoir) and the northern sectors (highlight the Serra da Estrela) present the highest landslide density and susceptibility (Fig. 5).

The results of the $A_1$ and $R_1$ indexes show important differences among the predisposing factors that have been integrated in the landslide susceptibility models (Table 4). The LUC predisposing factors (COS and CLC) registered the highest $A_1$ results, highlighting COS’s LUC types with a higher $A_1$. These results show the relevance of certain classes of COS in the predisposing factor dataset, by the number of landslide areas covered (emphasis on the forests, scrubland, and/or herbaceous vegetation associations, and open spaces with a scarcity or absence of vegetation).

The soil, SOAR, and slope angle present the highest values in the case of $R_1$, which shows that landslide density is concentrated in a reduced number of classes of each of the predisposing factor areas (e.g., hortic Luvisols, SOAR [22.5–25], and slope [between 25 and 45°]).

The landslide susceptibility model’s agreement test was performed using the landslide training inventory used to perform the outputs of each landslide susceptibility model, and these results were validated using the landslide test group.

<table>
<thead>
<tr>
<th>Factors</th>
<th>$A_1$</th>
<th>$R_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect</td>
<td>79.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Slope</td>
<td>76.1</td>
<td>0.6</td>
</tr>
<tr>
<td>SOAR</td>
<td>13.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Soil</td>
<td>62.4</td>
<td>1.0</td>
</tr>
<tr>
<td>Lithology</td>
<td>60.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Curvature</td>
<td>61.1</td>
<td>0.3</td>
</tr>
<tr>
<td>LUC (COS)</td>
<td>82.0</td>
<td>0.3</td>
</tr>
<tr>
<td>LUC (CLC)</td>
<td>76.0</td>
<td>0.3</td>
</tr>
</tbody>
</table>
The PRCs of each final susceptibility map (obtained from the results of the landslide training group) show slight variations (Fig. 6), but, in general terms, the curves are identical, demonstrating the high and similar performance of the models in the determination of landslide-susceptible areas.

The AUC of LSM1 and LSM2 that includes the same landslide information used to train the models is 94.1 % and 93.9 %, respectively. These results (landslide prediction) were considered to integrate the landslide susceptibility road network (LSRN1 and LSRN2) and the next analyses presented. Additionally, spatial differences were observed in the landslide susceptibility maps (Fig. 5), reflecting the differences of the influence of LUC properties.

When the two landslide susceptibility maps are reclassified into two classes (not susceptible IV \( \leq 0 \) and susceptible IV \( >0 \)), the susceptible area in LSM1 corresponds to 19.7 % and in LSM2 to 20.8 %. The CLC data provide IV results lower than the IV obtained with the COS data, but CLC is more generalized and justifies that the most susceptible area is observed in LSM2, compared to LSM1. The variation between the maximum and minimum IVs (3 and \(-3\) of \(\Delta IV\) in Fig. 5) show the landslide susceptibility differences derived from spatial representation of LUC classes of the two LUC datasets considered. The highest variations between LSM1 and LSM2 are found in places with reduced IVs (low and moderate susceptibility), marking the central sector of the study area. The areas with the highest IVs in LSM1 and LSM2 present a lower variation.

The PRCs of each final susceptibility map (obtained from the results of the landslide training group) show slight variations (Fig. 6), but, in general terms, the curves are identical, demonstrating the high and similar performance of the models in the determination of landslide-susceptible areas.

The AUC of LSM1 and LSM2 that includes the same landslide information used to train the models is 94.1 % and 93.9 %, respectively. These results (landslide prediction) were considered to integrate the landslide susceptibility road network (LSRN1 and LSRN2) and the next analyses presented. Additionally, spatial differences were observed in the landslide susceptibility maps (Fig. 5), reflecting the differences of the influence of LUC properties.

When the two landslide susceptibility maps are reclassified into two classes (not susceptible IV \( \leq 0 \) and susceptible IV \( >0 \)), the susceptible area in LSM1 corresponds to 19.7 % and in LSM2 to 20.8 %. The CLC data provide IV results lower than the IV obtained with the COS data, but CLC is more generalized and justifies that the most susceptible area is observed in LSM2, compared to LSM1. The variation between the maximum and minimum IVs (3 and \(-3\) of \(\Delta IV\) in Fig. 5) show the landslide susceptibility differences derived from spatial representation of LUC classes of the two LUC datasets considered. The highest variations between LSM1 and LSM2 are found in places with reduced IVs (low and moderate susceptibility), marking the central sector of the study area. The areas with the highest IVs in LSM1 and LSM2 present a lower variation.

### 3.2 Landslide susceptibility in the road network

Due to the width of the road network, in most cases, these infrastructures are not identified in the LUC data due to the properties or specifications (Table 1), namely, the minimum distance between lines considered in each LUC data in the research. The class “road and rail networks and associated land” (LUC nomenclature, level III) integrates the main class “industrial, commercial, and transport units” (level II); how-
ever, when a tabulation of the area of the road network used in this research and the LUC datasets was performed, the density of the roads was different in each LUC class between datasets (Fig. 7).

The IVs of LSM1 and LSM2 assigned to the road network differentiated the roads according to the landslide susceptibility, representing the highest IV where future landslides will occur and possibly rupture the road network or cause socioeconomic constraints due to total or partial blockages. In this case, the differences in the roads’ landslide susceptibilities were also analyzed.

The IVs assigned to the road network do not have spatial agreement between the two models. The difference between the maximum and minimum IVs of the LSRN1 and LSRN2 variations is significant, with IV variation of approximately 1. The interquartile range of the IV is greater in LSRN2 than in LSRN1 (Fig. 8). However, the IV average is similar in LSRN2 in comparison to LSRN1.

The landslide susceptibility map of the road network obtained by LSM1 (resulting in LSRN1) (Fig. 9) shows that it is spatially contrasted along the road network, highlighting the places where future landslides that may cause disturbances on the roads are most likely to occur. Conversely, in the landslide susceptibility map of the road networks obtained by LSM2 (resulting in LSRN2), the IV assigned to the road network is generally lower when compared to LSRN1, a result derived from the LUC generalization (CLC) used in the input of model 2.

LSRN1 includes 14.1% of the roads with a positive landslide susceptibility (IV ≥ 0), and the roads with high landslide susceptibility (IV > 10) represent only 0.1% of the total road network (Fig. 9). In LSRN2, the positive landslide susceptibility (IV ≥ 0) increases (compared with LSRN1) and comprises 14.7% of the total road network, where 0.1% of this network corresponds to a high landslide susceptibility (IV > 10).

LSRN2 does not show a high variation in short road distances; i.e., the IV tends to be extended within each polygon of the same class of the CLC’s LUC (larger polygons in comparison with the COS data), reducing the IV variation along the roads. The variation in the IV within each polygon of the LUC data is only explained by the remaining predisposing factors included in the model.

In LSRN2, the places with a high landslide susceptibility are not always identified as those where landslides effectively occurred (Fig. 10). The landslide susceptibility of the road network enhances the results obtained with COS (LSRN1) in areas with very high landslide susceptibility, precisely where landslides were validated in the fieldwork. These results show the importance of LUC data properties in the spatial differentiation of landslide susceptibility.

The spatial agreement and kappa coefficient between the LSRN1 and LSRN2 landslide susceptibility classes are 89.7 and 83.1%, respectively (Table 5). In general, the individual susceptibility classes present a high agreement (≥ 80%), except the high and very high classes of LSRN2 but with differences between the two models. For example, the landslide susceptibility class “very high” comprises 0.05% and 0.06% of the total road network in LSM1 and LSM2, respectively, but LSRN2 presents 20.4% of the omission differences in the same susceptibility class compared to the 3.8% commission differences of LSRN1. The intermediary susceptibility classes of the two models highlight the omission and commission differences.

Although variations exist between LSRN1 and LSRN2 landslide susceptibility, the relationship between the two models’ outputs is high, presenting a Pearson correlation co-
efficient of 0.98 (significance level $p<0.05$). The results of this correlation reflect the existence of an agreement on the spatial variation between LSRN1 and LSRN2; i.e., in general, when the IV of one output increases the other also increases, or vice versa, regardless of the discrepancy among the IVs of the same cells of each output.

The LSRN1 and LSRN2 results were crossed with all landslides that caused perturbations or disruptions of the road network, and the performance of models was assessed. Overall, the results were very good, with 89.5% and 89.3% AUC for LSRN1 and LSRN2, respectively. However, LSRN1 offers slightly better results when compared to LSRN2, as can be seen in the representation of the SRC (Fig. 11); i.e., up to 20% of the total area of the road network validates approximately 83% of the landslide susceptibility of LSRN1 and LSRN2. Nevertheless, LSRN2 shows a slightly better performance (to approximately 45% of the total area of the road network), but LSRN1 improves its validation performance at this point, being completely validated with 67% of the to-
Table 5. Spatial agreement between LSRN1 and LSRN2 (percentage of road network).

<table>
<thead>
<tr>
<th>LSRN1</th>
<th>Very low (IV &lt; −5)</th>
<th>Low (IV −5–0)</th>
<th>Moderate (IV 0–5)</th>
<th>High (IV 5–10)</th>
<th>Very high (IV &gt;10)</th>
<th>Total area (%)</th>
<th>Agreement (%)</th>
<th>Commission differ. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low (IV &lt; −5)</td>
<td>46.3</td>
<td>2.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>48.5</td>
<td>95.5</td>
<td>4.5</td>
</tr>
<tr>
<td>Low (IV −5–0)</td>
<td>3.5</td>
<td>31.7</td>
<td>2.3</td>
<td>0.0</td>
<td>0.0</td>
<td>37.4</td>
<td>84.6</td>
<td>15.4</td>
</tr>
<tr>
<td>Moderate (IV 0–5)</td>
<td>0.0</td>
<td>1.7</td>
<td>10.1</td>
<td>0.5</td>
<td>0.0</td>
<td>12.3</td>
<td>82.4</td>
<td>17.6</td>
</tr>
<tr>
<td>High (IV 5–10)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>1.5</td>
<td>0.01</td>
<td>1.7</td>
<td>88.1</td>
<td>11.9</td>
</tr>
<tr>
<td>Very high (IV &gt;10)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.05</td>
<td>0.05</td>
<td>96.2</td>
<td>3.8</td>
</tr>
<tr>
<td>Total area (%)</td>
<td>49.8</td>
<td>35.5</td>
<td>12.58</td>
<td>2.03</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreement (%)</td>
<td>93.0</td>
<td>89.2</td>
<td>80.3</td>
<td>75.0</td>
<td>79.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omission differ. (%)</td>
<td>7.0</td>
<td>10.8</td>
<td>19.7</td>
<td>25.0</td>
<td>20.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall agreement: 89.7 %
Kappa coefficient: 83.1 %

Figure 11. Success rate curves of LSRN1 and LSRN2 models.

4 Discussion

In landslide hazard and risk assessment, the LUC data integrate the controlling factor group and, in many evaluations, are directed by another factor input to the model. Usually LUC data are used as a landslide conditioning factor, which, in some cases, is scarce, generalized, and not very detailed. For example, Eeckhaut and Hervás (2012) verified that in different locations of Europe CLC is widely used for landslide assessment because it is the only LUC data available. Remote sensing and satellite images contributed to LUC data acquisition for landslide susceptibility assessment in different times (Guzzetti et al., 2012) and territories and minimize some problems of scarcity and detail (thematic and resolution). LUC is an important conditioning factor in landslide susceptibility (Pisano et al., 2017), and the high accountability index results prove this fact (Table 4).

There are several studies about the influence of land use cover changes on landslide susceptibility (e.g., Karsli et al., 2009; Mugagga et al., 2012; Promper et al., 2014; Reichenbach et al., 2014), although to the best of our knowledge there are no approaches that analyze the influence of different LUC datasets with different properties (date and base maps used on the production, spatial resolution, scale, minimum mapping unit, or others) on the landslide susceptibility results. When the landslide predisposing factors are collected, the LUC dataset must be selected according to its abovementioned properties and not only on the basis of its availability and free-of-charge conditions.

When different LUC datasets are available, the choice for the LUC dataset used in the landslide susceptibility assessment is not always clearly justified, and the results may vary according to LUC data properties selected. For Portugal different LUC datasets (with different properties) are available, but the use of each dataset can generate different conclusions; for example, different land use and land cover changes in the same period were observed by Meneses et al. (2018c).

This study highlights the landslide susceptibility differences derived exclusively from the LUC data properties because the other predisposing factor maps are the same in both models. Although if another method is used or the terrain mapping unit or another characteristic is changed, the results may vary, which has already been widely discussed (Chen et al., 2016; Den Eeckhaut et al., 2010; Guzzetti et al., 2006; Oliveira et al., 2015a; Zêzere et al., 2017).

Further, the data of soil and lithology were constrained and very generalized (1 : 100 000 and 1 : 50 000 scales, respectively), and this factor can influence the IV results if more detailed data were considered in the modeling process. The performance of the landslide susceptibility mapping and assessment is controlled by the quality of the available data not only the method (Pourghasemi et al., 2014).
Some research works refer to the quality of geoinformation (scale and precision) on the final result changes (e.g., Etter et al., 2006). In this case, the degree of completeness and the positional, geometric, and thematic agreement of the selected LUC data were evaluated by different proprietary institutions, with more than 80% accuracy, i.e., where the semantic inconsistency error was reduced, an important factor in reducing the error propagation and achieving a product with the best quality (Van Oort and Bregt, 2005; Regnauld, 2015).

The landslide inventory was obtained by photointerpretation, which is certainly not complete, especially in forest and agricultural areas, a fact that could have an impact on the landslide susceptibility zonation of the study area. This inventorying method does not allow for shallow or small landslide identification in forest areas, where the type, height, and density of the vegetation is important to landslide activity (Guzzetti et al., 2012), or in cultivated areas where agricultural practices erase the morphological and LUC signature of slope failures (Fiorucci et al., 2011). The quality and completeness of the landslide inventories can interfere with the quality of future landslide spatial occurrences (Galli et al., 2008; Guzzetti et al., 2012; Reichenbach et al., 2018). However, the landslide inventory is the same for both landslide models presented in this research, and the variation results depend exclusively on the LUC datasets that are integrated into each model.

The correlation between the outputs of each model is high, but there are spatial differences between them. The COS data are more detailed (1 : 25 000) than the CLC data (1 : 100 000), and the LUC classes are more differentiated in the territory, allowing greater detail and agreement in determining the areas with high landslide susceptibility, which were verified in LMS1. In model 2, CLC data are less detailed and contribute to IVs that are lower (low and very low landslide susceptibility in LSM2) compared with LSM1.

IV is more generalized along the road network at LSRN2 when compared with LSRN1, results derived exclusively from the input of LUC data with different properties in the models. These results highlight the importance of generalization and scale of LUC data selected in the landslide susceptibility assessment.

In the road network intersection with the LUC data, a high absence of road data was observed in the class industrial, commercial, and transport units, which is explained by the cartographic generalization due to the minimum mapping unit and minimum distance among lines of each LUC dataset. These factors exclude the road data due to the minimum requirements defined in the technical specifications of each LUC dataset creation. However, the distribution of the road network among the LUC classes is quite variable in both LUC datasets (COS and CLC), one of the factors that also justifies the variation in landslide susceptibility observed in different outputs.

The results of the PRC and AUC for LSM1 and LSM2 show a high quality and performance of both models in the landslide susceptibility area determination (Guzzetti et al., 2006), but LSM1 presents a slightly better performance. Nevertheless, the prediction landslide results were validated with the landslide test group and present good results to be assigned in the road network.

The LSRN1 and LSRN2 models’ validation results demonstrate that the models effectively identify the places where the landslides occurred and are more likely to occur in the future. In this case, the SRC and AUC note the high efficiency of the models (Guzzetti et al., 2006), with LSRN1 having a slightly higher efficiency, highlighting the properties of the LUC data.

Some roads in the study area were affected by landslides, a fact confirmed during the fieldwork developed to validate the landslide inventory (examples of some road blockage or damage in the Serra da Estrela and sample areas). In certain cases, the affected roads are important accesses points for the most isolated villages in the study area, and in some cases, a landslide can isolate the villages because part of the affected infrastructure is the only public access, a fact verified in the sample areas.

The results highlight the importance of LUC data properties in landslide assessment. More detailed LUC data (COS data) allow better landslide susceptibility results, a fact that was also described by Dymond et al. (2006), identifying some places where landslides occurred in the study area. Detailed predisposing factor data are recommended in landslide susceptibility assessment, a fact also mentioned in other studies. For example, Fressard et al. (2014) refer to the importance of detail in geomorphological variables to obtain high-quality results in landslide prediction.

This study was performed in a specific watershed, which highlights that landslide susceptibility changes according to the LUC data properties. It is recommended that the LUC data to be used as a predisposing factor of landslide susceptibility (e.g., in road networks) be as detailed as possible and small-scale LUC datasets be avoided (≥ 25 000). Further research is needed to test if these results change when the scale is different (e.g., national scale or very detailed scale).

In the analysis of the risk associated with road transportations, the higher the probability of a given event or incident, the greater the consequences (Berdica, 2002). In this context, the determination of the locales with the highest landslide susceptibility is very important, enabling prevention and minimizing these consequences or enabling better reactions when dealing with emergencies because road closures change the travel and reaction time (Meneses and Zézere, 2012).
5 Conclusions

Landslide susceptibility in the Zêzere watershed is spatially variable, highlighting some characteristics of the study area’s geo-factors in high landslide density in a specific location, for example, the highest slope angles and certain LUC types (e.g., forests and scrubland) and lithology.

The properties of the data that are integrated into the landslide susceptibility models are also an important issue to be considered since the variation in the properties of the same geo-factor provided different results, in this case LUC with different properties.

More detailed LUC data (COS) allow better landslide susceptibility results, while more generalized LUC data (CLC) resulted in the landslide susceptibility being more reduced, disallowing the identification of some places where landslides occurred. However, the results of the two susceptibility models showed good performance, a fact demonstrated by the validation of the models’ results.

The assignment of the landslide susceptibility results to the road network allowed the identification of the locations with the highest spatial probability for landslide occurrence. The LSRN1 map stands out with better results due to the integration of the COS dataset, showing the importance of LUC data detail in the identification of locations where landslides have occurred. The LSRN2 map does not have good performance in the identification of high landslide susceptibility in all road sections where landslides have occurred. In general, both LSRN1 and LSRN2 show the same trend in the spatial variation in landslide susceptibility of the study area’s road network, highlighting the high susceptibility on the slopes of the Serra da Estrela and near the Castelo de Bode reservoir.

Finally, LUC data properties were shown to be important in the variation in landslide susceptibility results. When the locations where landslides are likely to occur are known, alternative options can be created to avoid partial or complete isolation of certain localities, reduce the social and economic constraints of this population, and adopt preventive measures and alternative evacuation paths in case of landslide occurrence.

Data availability. The data used in this paper are accessible at http://www.dgterritorio.pt/, https://land.copernicus.eu/pan-european/corine-land-cover (LUC), https://sniamb.apambiente.pt/ (soil and lithology), and https://www.igeoe.pt (road network). All websites were last accessed on 10 June 2017.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/nhess-19-471-2019-supplement.

Author contributions. All authors contributed to the preparation of this paper.

References


Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “Landslide-road network interactions”. It is not associated with a conference.

Acknowledgements. This work was financed by national funds through the FCT-Portuguese Foundation for Science and Technology, I.P., under the framework of the project BeSafeSlide – Landslide early warning soft technology prototype to improve community resilience and adaptation to environmental change (PTDC/GES-AMB/30052/2017) and by Research Unit UID/GEO/00295/2013 (Centre for Geographical Studies). Bruno Meneses was financed through a grant of the Institute of Geography and Spatial Planning and Universidade de Lisboa, IGOT-UL (BD2015).

Edited by: Thomas Glade
Reviewed by: two anonymous referees


Orongo, N. D.: GIS Based, Cartographic Generalization in Multi-scale Environment: Lamu County, GPO, University of Nairobi, Kenya, 2011.


