Impact of Input Data on the Quality of the Landslide Susceptibility Large-Scale Maps: A Case Study from NW Croatia

Krkač, Martin; Bernat Gazibara, Sanja; Sinčić, Marko; Lukačić, Hrvoje; Šarić, Gabrijela; Mihalić Arbanas, Snježana

Impact of Input Data on the Quality of the Landslide Susceptibility Large-Scale Maps: A Case Study from NW Croatia

Martin Krkač, Sanja Bernat Gazibara, Marko Sinčić, Hrvoje Lukačić, Gabrijela Šarić, and Snježana Mihalić Arbanas

Abstract

The study presents analyses of input data impact on the quality of the landslide susceptibility large-scale maps. For comparison, two input data sets were used to produce two landslide susceptibility maps. The first input data set included free-available, small-scale data with low spatial accuracy, while the second set included high-resolution remote sensing data. The same nine types of landslide causal factors were derived and used for susceptibility analyses. Furthermore, LiDAR-based landslide inventory and bivariate statistical method, i.e. Information Value method, were used for susceptibility modelling. The resulting landslide susceptibility maps were compared with ROC curves. Success and prediction rates showed that the landslide susceptibility model based on causal factors derived from high-resolution remote sensing data is approximately 10% more accurate than the model based on causal factors derived from small-scale input data. Furthermore, based on the conducted research, it can be concluded that susceptibility modelling based on small-scale data and LiDAR-based inventories enables reliable landslide susceptibility assessments at the regional level.

M. Krkač (✉) · S. Bernat Gazibara · M. Sinčić · H. Lukačić · G. Šarić · S. Mihalić Arbanas
University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Zagreb, Croatia
e-mail: mkrkac@rgn.hr; sbernat@rgn.hr; msincic@rgn.hr; hlukacic@rgn.hr; gsaric@rgn.hr; smihalic@rgn.hr

1   Introduction

One of the first prerequisites for risk reduction measures and mitigation of landslide consequences is creating prognostic maps, such as landslide susceptibility maps. Reliable susceptibility maps on a large scale (≤1:5000) require good quality input data, i.e., detailed and complete landslide inventories (Glade 2001; van Westen et al. 2008) and appropriate resolution and spatial accuracy of geo-environmental and triggering factors (Guzzetti et al. 1999; van Westen et al. 2006). Reichenbach et al. (2018) pointed out that nowadays, landslide investigators are more interested in experimenting with different modelling techniques than acquiring good-quality input data and they often use landslide and geo-environmental information captured at different (in cases very different) cartographic scales even for the same study area.

The most usual mapping techniques for preparing large-scale landslide inventories are the interpretation of aerial and satellite images, interpretation of digital elevation models, and field mapping (Eeckhaut et al. 2007; Fiorucci et al. 2011; Bernat Gazibara 2019). The advantage of the LiDAR (Light Detection and Ranging) technique and data derived by ALS (Airborne Lidar Scanning) compared to other remote sensing techniques is ground detection in forested terrain and the possibility of creation of high-resolution bare-earth digital elevation models (DEMs), which enables the identification and mapping of small and shallow landslides in densely vegetated areas (Chigira et al. 2004; Eeckhaut et al. 2007; Razak et al. 2011; Đomlija 2018; Bernat Gazibara et al. 2019a, b).

For large-scale landslide susceptibility analyses, researchers mainly use geomorphological, hydrological, geological and land use data (Reichenbach et al. 2018). Geomorphological (i.e. slope, aspect, elevation) and some hydrological data (i.e. drainage network, wetness) can be derived from digital elevation models, which are generated from different input data in various scales, for example, from

topographic maps (Pellicani et al. 2014; Roşca et al. 2015; Zêzere et al. 2017; Vojteková and Vojtek 2020), aerial photographs, LiDAR data (Petschko et al. 2014; Yusof et al. 2015; Gaidzik and Ramírez-Herrera 2021), and satellite images in resolution ranged from 25 cm (Vojteková and Vojtek 2020) up to 15 or 30 m (Xing et al. 2021). Geological data, such as lithology or geological contacts, are usually available on smaller-scale maps, from 1:25,000 (Lee and Min 2001; Pellicani et al. 2014; Zêzere et al. 2017) to 1: 100,000 (Pellicani et al. 2014; Sinčić et al. 2022a). Land use data used in susceptibility analysis can be derived from satellite images, aerial photos, or official land use plans (Zêzere et al. 2017).

The objective of the paper is to analyse the impact of input data on the quality of landslide susceptibility large-scale maps. Therefore, two susceptibility maps were prepared based on the same thematic type of information but derived from input data which differ in spatial accuracy, resolution and scale. Furthermore, with conducted research, we wanted to address that the adequate scale of input data will minimise deviations from actual environmental conditions on final landslide susceptibility maps. The landslide inventory used for modelling and validating susceptibility maps is LiDAR-based (Krkač et al. 2022).

2 Study Area

The study area comprises 20.22 km^2 of hilly terrain, located in the Hrvatsko Zagorje region, NW Croatia (Fig. 1). Approximately 88% of the study area has slope angles greater than 5°, and approximately 57% of the area has slope angles between 12° and 32°, which makes the study area prone to sliding. In addition, the elevation ranges between 222 and 682 m a.s.l., whereas most of the study area (66%) is placed at elevations below 300 m a.s.l.

The lithological units of the study area are composed of Miocene (78%), Quaternary (14%) and Triassic sediments (7%). The Triassic sediments, located in the northeastern part of the study area, are composed of sandstones, shales, dolomites, limestones and dolomitised breccias (Šimunić et al. 1982). The Miocene sediments are composed of sandstones, marls, calcareous marls, biogenic and marly limestones, sands and tuffs. The Quaternary sediments, located in the valleys around streams and rivers, are composed of sands, silts and gravels.

According to the Corine Land Cover land use data (URL-1) the area is covered by forests (63%), agricultural areas (36%) and artificial areas (1%). According to official
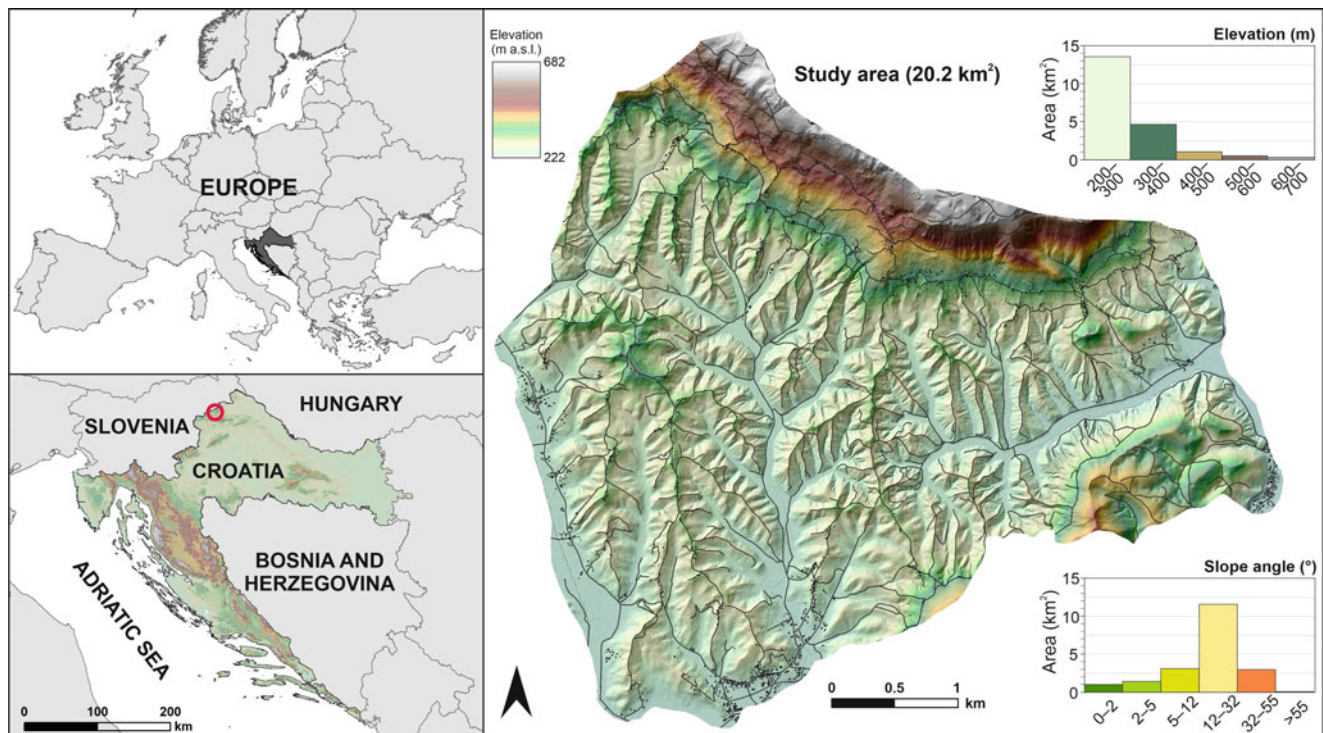

Fig. 1 Geomorphological conditions and location of the study area (NW Croatia)

spatial planning maps of the local municipalities (URL-2 and URL-3), 52% of the area is covered by forests, 40% by agricultural areas and pastures and 8% by artificial areas (Sinčić et al. 2022b).

Generally, precipitations and human activities are the primary triggers of landslides in NW Croatia (Bernat et al. 2014). The climate of the study area is continental, with a mild maritime influence. The mean annual precipitation for 1949–2020 is 873.7 mm (Zaninović et al. 2008), according to the data from the nearest meteorological station (30 km to the east).

3   Materials and Methods

3.1   Landslide Inventory

The landsides were interpreted from the detailed LiDAR digital terrain model (DTM) derivatives based on the recognition of landslide features (Krkač et al. 2022). The LiDAR DTM derivatives used to interpret the landslide morphology were hillshade maps, slope maps and contour lines. The mapping was performed on a large scale (1:100–1:500) to ensure the correct delineation of the landslide boundaries. In addition, orthophoto images from 2014 to 2016 were used to check the morphological forms along roads and houses, such as artificial fills and cuts, similar to landslides on DTM derivatives.

Totally, 912 landslides were mapped, with a total area of 0.408 km^2, or 2.02% of the study area. The mean landslide density is 45.1 slope failures per square kilometre. The size of the recorded landslides ranges from a minimum value of 3.3 m^2 to a maximum of 13,779 m^2, whereas the average area is 448 m^2 (median = 173 m^2, std. dev. = 880 m^2). The small size of the landslides is probably the result of geological conditions (mainly Miocene marls covered with residual soils) and geomorphological conditions, where the differences between the valley bottoms and the top of the hills are rarely higher than 100 m (Krkač et al. 2022). The

prevailing dominant types of landslides are shallow soil slides.

3.2   Input Data and Landslide Causal Factors

Several input data sources were used to acquire landslide causal factors (Table 1), considering the scale, the purpose of the study and availability. The input data for the geomorphological and hydrological causal factors were EU Digital Elevation Model (URL-1) and the LiDAR data. The EU Digital Elevation Model (EU-DEM) is available on a regional scale as a part of the Copernicus Land Monitoring Service (CLMS), providing elevation data in 25 m resolution. The average spacing of the LiDAR ground points, obtained by scanning the study area, was 0.28 m, providing sufficient density to create a DTM up to a 0.3 m resolution.

The input data for geological causal factor maps was Basic Geological Map (Šimunić et al. 1982; Aničić and Juriša 1984) on a scale of 1:100,000, which presents the only available source of geological information for the study area. Anthropogenic landslide input data were obtained from publicly available Corine Land Cover (CLC) (URL-1) and Open Street Map (OSM). CLC for the 2018 period was derived dominantly from Sentinel-2 satellite data and Landsat-8 for gap filling. OSM is a community-driven initiative providing map data for various applications about traffic infrastructure, buildings and other spatial information. Additional input data used to derive causal factor maps on a more detailed scale were digital orthophoto maps on a scale 1: 5000.

3.3   Information Value Method

The landslide susceptibility of the study area was assessed by the Information Value Method (IVM). IVM is a bivariate statistical approach based on the calculation of weights of landslide causal factors (Yin and Yan 1988; Jade and Sarkar

Table 1 Input data used for derivation of landslide causal factors

Data	TYPE OF DATA	Scale
EU-DEM	DTM	25 m resolution
LiDAR point cloud		5 m resolution
Croatian Basic Geological Map	Geological	1:100,000
Corine Land Cover	Land use	1:100,000
Open Street Map	Traffic infrastructure	Large
Digital Orthophoto Map	Orthophoto	1:5000

1993; Sarkar et al. 2013). In this method, landslide occurrence is considered a dependent variable, and each causal parameter is considered an independent variable (Farooq and Akram 2021).

The information value I_i of each variable i is given by Yin and Yan (1988):

$$I_i = log \frac{S_i/N_i}{S/N} \tag{1}$$

where S_i is the number of grid cells involving the parameter i and containing a landslide, N_i is the number of grid cells involving the parameter i, S is the number of grid cells with a landslide, and N is the total number of data points (grid cells). Negative values of I_i mean that the presence of the variable is not relevant in landslide development (Zêzere et al. 2017). Positive values of I_i indicate a relevant relationship between the presence of the variable and landslide development, whereas the higher value indicates a stronger relationship.

The total information value I_j for a grid cell j is given by the equation (Yin and Yan 1988):

$$I_j = \sum_{i=1}^{m} X_{ji} I_i \tag{2}$$

where m is the number of variables, X_{ji} is either 0 if the variable is not present in the grid cell j, or 1 if the variable is present. The obtained values are directly proportional to the susceptibility.

3.4   Validation of Landslide Susceptibility Maps

The performance of landslide susceptibility maps was evaluated by Receiver Operator Characteristic (ROC) curves (Lobo et al. 2008), i.e. a graphical representation of sensitivity to 1-specificity (false positive rate). ROC curves are one of the most used methods to describe the landslide susceptibility model performance (Rossi and Reichenbach 2016). The model's accuracy is considered higher if the ROC curve is closer to the upper left corner and lower if it is closer to the diagonal, representing a random test. If the ROC curve is derived from the set of landslides used for susceptibility modelling, it represents the success rate of the landslide susceptibility map. Likewise, the ROC curve derived with landslide inventory for model validation shows the prediction rate of the landslide susceptibility map. Out of 912 landslides in the LiDAR-based inventory, for the study area of 20.22 km^2, approximately 50% of the landslides were selected for landslide susceptibility analyses, while the remaining landslides were used to verify the results.

4   Landslide Causal Factors

From the available input data, two sets of nine landslide causal factors maps were derived for susceptibility assessment and to analyse the impact of input data on the quality of landslide susceptibility large-scale maps. The first set of causal factor maps, used in Scenario 1, is derived from less detailed input data, such as EU-DEM or CLC. This group of causal factor maps is derived in a raster resolution of 25 m. The second set of causal factor maps, used in Scenario 2, is derived from more detailed input data, such as LiDAR point cloud or digital orthophoto. This group of causal factor maps is derived in 5 m resolution. A list of landslide causal factor maps and input data is presented in Table 2, while the comparison of factor class area distributions is presented in Fig. 2.

The geomorphological factor maps derived from EU-DEM and LiDAR point clouds are elevation, slope and aspect. The elevation is a frequently used parameter in landslide susceptibility studies because landslides may form in specific relief ranges (Lee et al. 2001). Factor maps were created by reclassifying the elevations into classes of 50 m. The slope gradient is often considered to be the most important geomorphometric parameter used to analyse and describe relief (van Westen et al. 2008). Factor maps were created by reclassifying the slopes into classes of 5°. The aspect causal factor maps were classified into ten main classes (N, NE, E, SE, S, SW, W, NW and flat surface).

The hydrological factor maps, derived from EU-DEM and LiDAR DTM, were drainage network and topographic wetness. A drainage network is a set of all drainage systems in an area, i.e. a set of natural canals through which water constantly or temporarily flows and which are connected into a single stream and represent the smallest independent geomorphological component (Marković 1983). The drainage networks were derived with several tools in the Spatial Analyst (Hydrology Toolbox) in ArcGIS 10.8.1. Topographic wetness is a steady-state wetness index, and it is commonly used to quantify topographic control on hydrological processes (Sinčić et al. 2022b). The topographic wetness maps are derived by the Compound Topographic Index (CTI) tool in Geomorphometry and Gradient Metrics Toolbox (Evans et al. 2014). Lower values of CTI indicate lower wetness (hill tops), and higher CTI values indicate higher terrain wetness (plain areas).

Geological factor maps were obtained by digitising a geological map on a scale of 1:100,000, resulting in two causal factor maps, a stratigraphic unit map, and a distance to geological (lithological) contact map. However, field

Table 2 Landslide causal factors maps, derived from different input data, used in different susceptibility scenarios

		Input data	
Data type	Causal factor	Scenario 1	Scenario 2
Geomorphological	Elevation	EU-DEM	LiDAR point cloud
	Slope		
	Aspect		
Hydrological	Distance from the drainage network		
	Topographic wetness		
Geological	Lithology	Basic Geological Map	LiDAR DTM mapping, Basic Geological Map
	Distance from the geological contact		
Anthropogenic	Land use	Corine Land Cover	Digital orthophoto maps, LiDAR point cloud and DTM
	Distance from road	Open Street Map	

verification of Basic Geological Map showed significant deviations in the position of the geological contacts (Sinčić et al. 2022b). To obtain more precise data for Scenario 2, stratigraphic units (corresponding to engineering geological formations) were additionally mapped (by modifying the geological contacts and adding one more unit, i.e. superficial slope-wash and talus sediments) using LiDAR DTM derivatives according to the suggestions by Jagodnik et al. (2020a, b).

Land use causal factor map for Scenario 1 was derived from CLC. Considering large-scale landslide hazard assessment, all classification levels showed little difference regarding land use classes in the study area (Sinčić et al. 2022b). Therefore, the first hierarchy level was selected for a representative land use factor map. A more detailed land use map, for Scenario 2, was the most demanding to derive as required different input data sets (LiDAR data and digital orthophoto maps) and different processing methods. A detailed explanation of the land use map derivation can be found in Sinčić et al. (2022b). Distance from roads causal factor map was derived from OSM for Scenario 1 and by combining LiDAR data and digital orthophoto maps for a more detailed causal factor map in Scenario 2. The total length of all roads in Scenario 1 is 54 km and the total length of roads in Scenario 2 is 165 km.

5   Results

The GIS-based landslide susceptibility assessment for the study area in NW Croatia, for two scenarios, comprises the following steps: (1) the pairwise correlation analysis of input variables; (2) the calculation of information values for each class of each factor map based on landslide modelling set; (3) the calculation of total information values, i.e. calculation of susceptibility; (4) calculation of success rate based on landslide modelling set; (5) calculation of prediction rate based on the landslide validation set; (6) reclassification of susceptibility maps. The landslides used in the calculation of information values and for the calculation of success and prediction rates are firstly converted from polygons to points (centroids) and then to raster (25 m resolution in Scenario 1, and 5 m resolution in Scenario 2). In the process, some of the small landslides in Scenario 1, located near each other, were converted to the same pixel, resulting in a different number of landslide pixels in both scenarios (Table 3).

Before performing landslide susceptibility analysis of the study area, the pairwise conditional independence test of the input landslide causal factors was done using Pairwise correlation analysis in the LAND-SUITE V1.0 (Rossi and Reichenbach 2016; Rossi et al. 2021) i.e. LAND-SVA: LANDslide-Susceptibility Variable Analysis, which includes Pairwise correlation analysis and Multicollinearity test. Inspection of the correlation results confirmed no correlations within nine landslide causal factors of both scenarios, assuming a Pearsons' R absolute value of 0.5 as the threshold for detecting correlations.

The information values for each class of each factor map, for both scenarios, are presented in Table 3. According to the analyses of factor maps in Scenario 1, the highest influence on landslide development have slope classes with angles between 20 and 30° and aspect class with azimuth towards NE, while elevation and lithology do not influence landslide development. The analyses of factor maps in Scenario 2 show that the highest influence on landslide development have slope classes with angles between 30 and 45°, while elevation and lithology practically do not influence on landslide development.

The landslide susceptibility maps were derived by the calculation of the total information values according to the Eq. (2). Susceptibility maps for scenarios 1 and 2 are presented in Fig. 3. Figure 3 also shows details of reclassified susceptibility maps, classified by the quantile method. The performance of the IVM was evaluated using the ROC curve, i.e. the area under the curve (AUC). The success rate of the susceptibility model in Scenario 1, calculated on the training set, is AUC = 70.8%, and the predictive rate, based on the validation set, is AUC = 69.8% (Fig. 4a). The success rate of the susceptibility model in Scenario 2 is AUC = 80.3%, and the predictive rate is AUC = 78.8% (Fig. 4b).
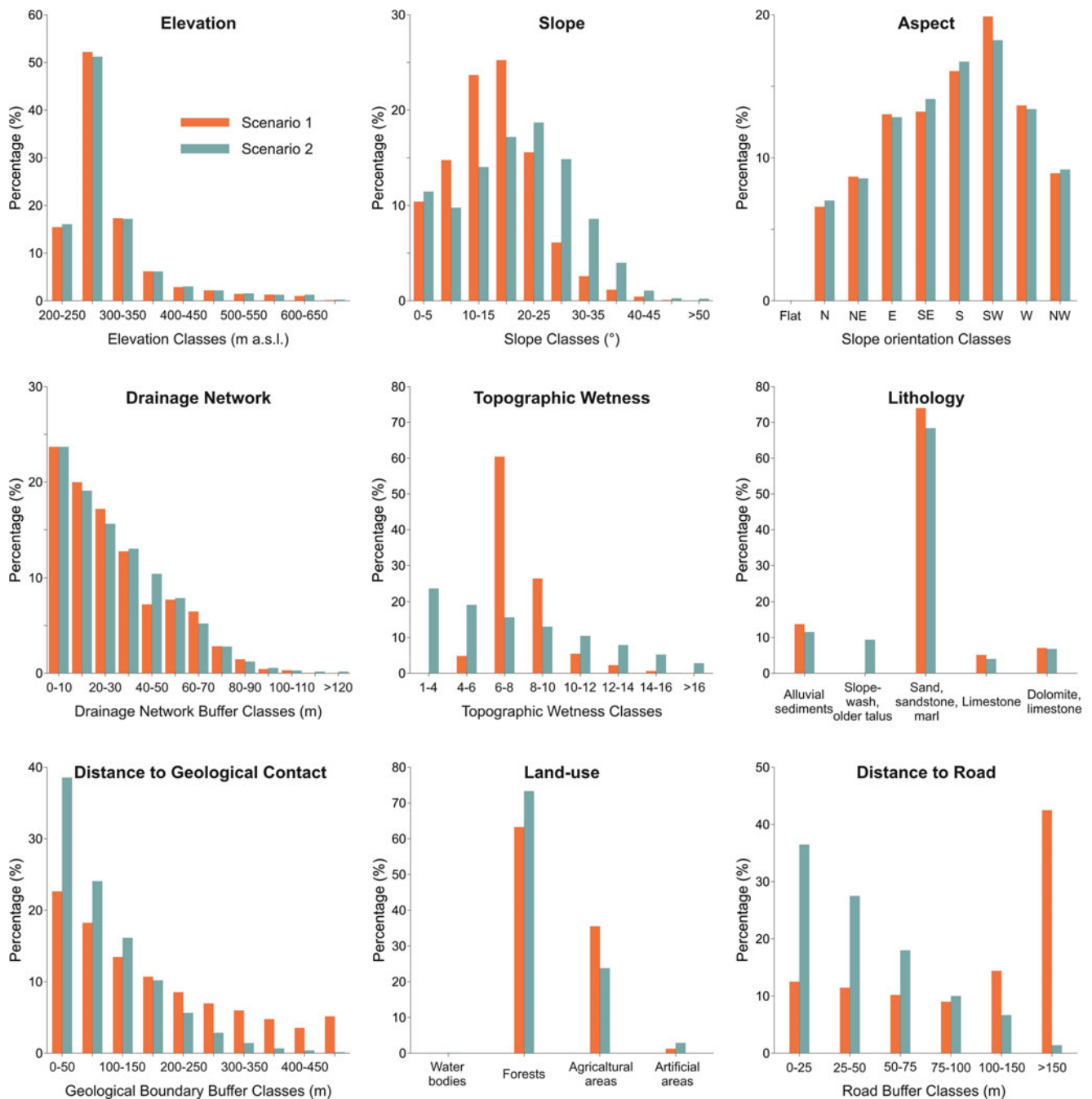
Fig. 2 Comparison of landslide causal factor classes derived from different types of input data

6   Discussion and Conclusions

Statistically-based susceptibility models are dependable on the type, abundance, quality and relevance of the available landslide and the geo-environmental information. In the presented research, we analysed the influence of input data quality and scale on the accuracy of landslide susceptibility models assessment on a large scale. For landslide susceptibility analysis, we chose the most commonly used thematic and environmental causal factors, i.e. elevation, slope, aspect, drainage network, topographic wetness, lithology, distance to geological contact, land use and distance to roads. For comparison, two different input data sets, i.e. scenarios, were used to derive landslide causal factor maps, while the landslide data used for susceptibility analysis was LiDAR-based inventory. Scenario 1 included free-available input data sets with a lower resolution and lover

Table 3 Calculated information values for landslide causal factors based on the Information Value Method

Causal factor map		Scenario 1				Scenario 2			
		Class description	Class (Npix)	Landslide (Npix)	IV	Class description	Class (Npix)	Landslide (Npix)	IV
Geomorphological landslide causal factors	Elevation m a.s.l.	200–250	4885	33	–0.3095	200–250	129,204	34	–0.3328
		250–300	16,514	309	0.1329	250–300	412,139	330	0.1504
		300–350	5475	71	–0.0262	300–350	138,447	68	–0.0618
		350–400	1954	10	–0.4300	350–400	49,249	11	–0.4040
		400–450	907	6	–0.3186	400–450	24,056	6	–0.3561
		450–500	690	4	–0.3759	450–500	17,346	2	–0.6912
		500–550	464	1	–0.8056	500–550	12,522	3	–0.3736
		550–600	404	1	–0.7455	550–600	10,249	1	–0.7637
		600–650	314	1	–0.6360	600–650	10,196	1	–0.7615
		650–700	39	0		650–700	1857	0	
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
	Slope angle (°)	0–5	3293	6	–0.8786	0–5	92,135	0	
		5–10	4670	41	–0.1957	5–10	78,518	13	–0.5341
		10–15	7488	84	–0.0892	10–15	112,905	38	–0.2260
		15–20	7983	141	0.10788	15–20	138,299	38	–0.3141
		20–25	4927	116	0.23270	20–25	150,428	69	–0.0915
		25–30	1936	41	0.18671	25–30	119,568	87	0.1089
		30–35	817	5	–0.3524	30–35	69,143	101	0.4115
		35–40	365	2	–0.4004	35–40	32,020	79	0.6392
		40–45	137	0		40–45	8562	29	0.7768
		45–50	28	0		45–50	2064	2	0.2333
		>50	2	0		>50	1623	0	
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
	Slope orientation	Flat terrain	2	0		Flat terrain	0	0	
		N	2077	42	0.1666	N	56,402	41	0.1085
		NE	2745	64	0.2284	NE	68,863	69	0.2478
		E	4127	71	0.0964	E	103,371	65	0.0455
		SE	4184	36	–0.2045	SE	113,655	40	–0.2066
		S	5082	46	–0.1824	S	134,594	46	–0.2193
		SW	6288	56	–0.1894	SW	146,682	64	–0.1132
		W	4320	71	0.07660	W	107,815	69	0.0531
		NW	2821	50	0.10939	NW	73,883	62	0.1708
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
Hydrological landslide causal factors	Proximity to the drainage network (m)	0–10	7493	129	0.0968	0–10	190,658	157	0.1626
		10–20	6320	102	0.0687	10–20	153,701	130	0.1742
		20–30	5433	95	0.1035	20–30	125,635	77	0.0344
		30–40	4024	44	–0.1004	30–40	104,686	48	–0.0917
		40–50	2282	29	–0.0351	40–50	83,849	31	–0.1852
		50–60	2440	17	–0.2961	50–60	63,336	3	–1.0776
		60–70	2048	15	–0.2744	60–70	41,879	6	–0.5969
		70–80	894	2	–0.7895	70–80	22,499	2	–0.8042
		80–90	466	1	–0.8076	80–90	9708	1	–0.7402
		90–100	138	1	–0.2790	90–100	4414	0	
		100–110	95	0		100–110	2260	0	
		110–120	8	1	0.9577	110–120	1282	0	
		>120	5	0		>120	1358	1	0.1141
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
	Topographic wetness					1–4	80,088	50	0.0424
		4–6	1522	12	–0.2424	4–6	432,300	256	0.0194
		6–8	19,113	267	0.0060	6–8	202,915	115	0.0004
		8–10	8382	124	0.0309	8–10	58,028	25	–0.1187

(continued)

Table 3 (continued)

Causal factor map		Scenario 1				Scenario 2			
		Class description	Class (Npix)	Landslide (Npix)	IV	Class description	Class (Npix)	Landslide (Npix)	IV
		10–12	1712	26	0.0423	10–12	19,900	6	–0.2737
		12–14	712	6	–0.2135	12–14	8439	2	–0.3783
		14–16	190	1	–0.4179	14–16	2892	2	0.0868
		>16	15	0		>16	703	0	
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
Geological landslide causal factors	Lithology	Q_2	4335	25	–0.3782	M, S, Gal	92,563	1	–1.7195
		M_2^2	1615	8	–0.4443	MS, Cb_d	75,261	42	–0.0063
		M_1^2	23,125	393	0.0911	S, Ss, Ms	550,716	407	0.1156
		Θ	290	3	–0.1244	Ls	32,231	1	–1.2613
		E_3	45	0		D, Ls	54,494	5	–0.7904
		T_2	1839	5	–0.7048				
		T_1	397	2	–0.4369				
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
	Proximity to the geological contact (m)	0–50	7162	90	–0.0400	0–50	310,330	177	0.0031
		50–100	5770	73	–0.0370	50–100	193,700	130	0.0738
		100–150	4259	58	–0.0050	100–150	129,881	83	0.0525
		150–200	3387	56	0.0792	150–200	82,099	29	–0.2050
		200–250	2698	38	0.0096	200–250	45,286	19	–0.1302
		250–300	2206	31	0.0086	250–300	22,988	10	–0.1145
		300–350	1898	25	–0.0195	300–350	11,361	5	–0.1095
		350–400	1512	30	0.1584	350–400	5298	3	0.0000
		400–450	1124	20	0.1111	400–450	3184	0	
		>450	1630	15	–0.1753	>450	1138	0	
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
Anthropogenic landslide causal factors	Land cover	Forest	20,015	320	0.0646	Forest	590,194	425	0.1044
		Agricultural	11,235	116	–0.1253	Agricultural	191,383	28	–0.5878
		Artificial	396	0		Artificial	23,039	2	–0.8145
						Water bodies	649	1	0.4347
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	
	Proximity to roads (m)	0–25	3958	26	–0.3217	0–25	293,600	128	–0.1136
		15–50	3621	50	0.0010	15–50	221,333	155	0.0923
		50–75	3222	53	0.0770	50–75	144,854	105	0.1072
		75–100	2847	42	0.0297	75–100	80,428	47	0.0137
		100–150	4556	61	–0.0124	100–150	53,707	20	–0.1820
		>150	13,442	204	0.0420	>150	11,343	1	–0.8078
		SUM	**31,646**	**436**		**SUM**	**805,265**	**456**	

spatial accuracy, such as EU-DEM or Corine Land Cover. In contrast, Scenario 2 included a more detailed input data set, such as high-resolution remote sensing data, i.e. LiDAR point cloud and digital orthophoto imagery in resolution 0.5 m.

The difference between the causal factor maps derived from different input data sets is visible from the class area distribution (Fig. 2). The highest distinction in the class area distribution of geomorphological and hydrological factor maps is among slope maps and topographic wetness maps. Slope map derived from LiDAR DTM generally have higher total area values in classes with a lower slope angle than slope

maps derived from EU-DEM. This observation is similar to previous research, such as Grohmann (2015), who concluded that the slope parameter is prone to changes as the resolution is coarsened, with a strong decrease in maximum values, mean values and standard deviation. Therefore, it can be concluded that high-resolution LiDAR DTM enables a more accurate presentation of environmental conditions in the case of slope and topographic wetness. Contrarily, the differences in the class area distribution of elevation, aspect and drainage network maps derived from EU-DEM and LiDAR DTM are generally low.
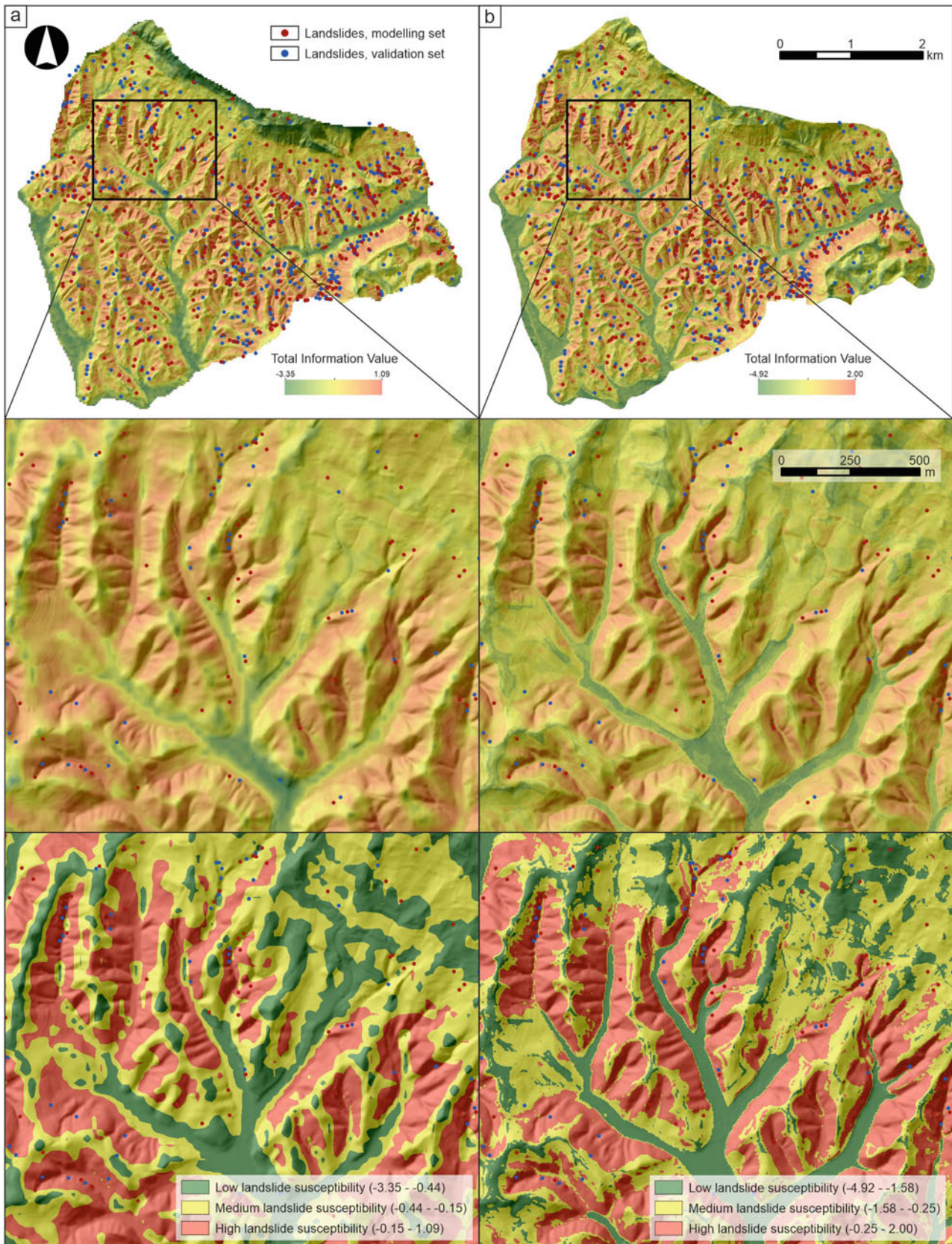
Fig. 3 Comparison of landslide susceptibility map classes derived from different types of input data: (**a**) Scenario 1; (**b**) Scenario 2
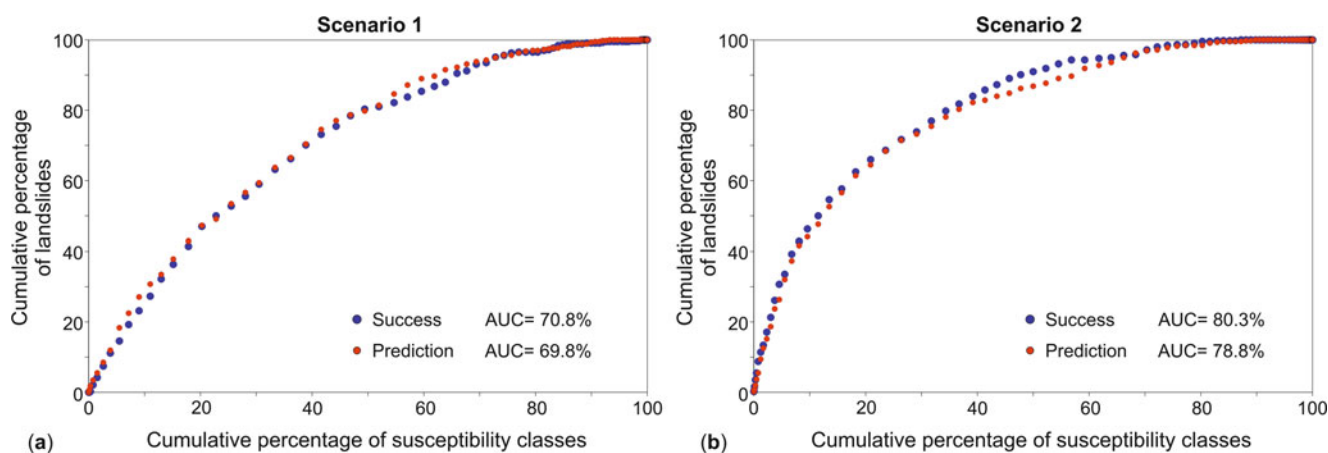
Fig. 4 ROC curve results for the rate of success and the rate of prediction for the landslide susceptibility maps

The lithological factor map for Scenario 1, derived from the Basic Geological Map in scale 1:100,000, has a similar distribution of lithological units as the engineering geological unit map interpreted based on LiDAR DTM derivatives. An exception is the presence of an additional engineering geological unit interpreted on LiDAR DTM derivatives, i.e. slopewash and older talus, which significantly affected the class area distribution of distance to the geological contact map in Scenario 2 compared to the class area distribution of factor map derived from Basic Geological Map 1:100,000 in Scenario 1.

The land use map for Scenario 1, derived from Corine Land Cover, has similar class area distribution as the land use map derived based on the automatic classification of high-resolution orthophoto imagery in Scenario 2. However, the spatial distribution of the land use classes between the scenarios is significant (Sinčić et al. 2022b), i.e. degree of cartographic matching between the two land use maps would be low and probably have a significant impact on the accuracy of the final landslide susceptibility maps. For example, the distance to roads map in Scenario 2, derived based on high-resolution remote sensing data, resulted in the total length of roads for almost 300% more than the length of roads derived only from the OSM input data in Scenario 1, and consequently, a significantly higher area percentage of the class 0–75 m in Scenario 2.

Generally, landslide susceptibility analysis and information values showed that the same factor classes in both scenarios showed an influence on landslide occurrence. Only factor classes in topographic wetness and proximity to geological contact maps have the opposite effect on landslide susceptibility. The most significant factor classes for the landslide occurrence are slope class 20–25° and aspect class NE for Scenario 1, and slope classes 35–45° for Scenario 2. Furthermore, the most significant factor classes in Scenario 2 have 2–3 times higher information values than the most significant factor classes in Scenario 1. High information

values in Scenario 2 resulted in different ranges of total susceptibility values in the two landslide susceptibility maps.

The resulting large-scale landslide susceptibility maps were compared with ROC curves, i.e. the success and predictive rates of AUC values for Scenario 1 (derived based on available small-scale input data) and Scenario 2 (derived based on high-resolution remote sensing data). The success and predictive rates for Scenario 1 are approximately 10% lower than the AUC values for Scenario 2. On first look, unclassified landslide susceptibility maps may look relatively similar, although the most significant difference between the two scenarios is the percentage of low landslide susceptibility values in valleys. Considering the stated, the landslide susceptibility map based on causal factors derived from high-resolution remote sensing data (Scenario 2) has minor deviations from actual environmental conditions. i.e. spatial accuracy of resulting susceptibility assessment is higher. However, after the classification of landslide susceptibility maps into three classes based on the quantile method, on the Scenario 1 map, it is possible to observe that the low susceptibility class is spread from the flat valleys to the hilly slopes. Contrary, on the Scenario 2 map, the low susceptibility class is confined only to flat areas. In the end, it can be concluded that different input data resolutions and scales can provide similar information regarding large-scale landslide susceptibility assessment, although high-resolution input data provide more spatially accurate information. Also, large-scale landslide susceptibility modelling based on small-scale input data would show significantly different results if we did not use sustainable complete, LiDAR-based landslide inventory in the analysis. Furthermore, it follows that it is possible to conduct regional landslide susceptibility assessment in small pilot areas (10–30 km^2) with high-resolution remote sensing data sets and apply them to larger research areas (>100 km^2) with similar geomorphological and geological conditions and only small-scale available input data.

References

Aničić B, Juriša M (1984) Basic geological map, scale 1:100,000, Rogatec, Sheet 33-68

Bernat Gazibara S (2019) Methodology for landslide mapping using high resolution digital elevation model in the Podsljeme area (City of Zagreb). Doctoral Thesis, Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb

Bernat Gazibara S, Krkač M, Mihalić Arbanas S (2019a) Verification of historical landslide inventory maps for the Podsljeme area in the City of Zagreb using LiDAR-based landslide inventory. Rudarsko-geološko-naftni zbornik 34(1):45–48. https://doi.org/10.17794/rgn.2019.1.5

Bernat Gazibara S, Krkač M, Mihalić Arbanas S (2019b) Landslide inventory mapping using LiDAR data in the City of Zagreb (Croatia). Journal of Maps 15:773–779. https://doi.org/10.1080/17445647.2019.1671906

Bernat S, Mihalić Arbanas S, Krkač M (2014) Inventory of precipitation triggered landslides in the winter of 2013 in Zagreb (Croatia, Europe). In: Sassa K, Canuti P, Yin Y (eds) Landslide science for a safer geoenvironment. Springer International, Cham, pp 829–835

Chigira M, Duan F, Yagi H, Furuya T (2004) Using an airborne laser scanner for the identification of shallow landslides and susceptibility assessment in an area of ignimbrite overlain by permeable pyroclastics. Landslides 1:203–209. https://doi.org/10.1007/s10346-004-0029-x

Đomlija P (2018) Identification and classification of landslides and erosion phenomena using the visual interpretation of the Vinodol Valley digital elevation model. Doctoral Thesis, Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb

Eeckhaut MVD, Poesen J, Verstraeten G, Vanacker V, Nyssen J, Moeyersons J, van Beek LPH, Vandekerckhove L (2007) Use of LIDAR-derived images for mapping old landslides under forest. Earth Surf Process Landforms 32:754–769. https://doi.org/10.1002/esp.1417

Evans JS, Oakleaf J, Cushman SA, Theobald D (2014) An ArcGIS toolbox for surface gradient and geomorphometric modeling, version 2.0-0

Farooq S, Akram MS (2021) Landslide susceptibility mapping using information value method in Jhelum Valley of the Himalayas. Arab J Geosci 14:824. https://doi.org/10.1007/s12517-021-07147-7

Fiorucci F, Cardinali M, Carlà R, Rossi M, Mondini AC, Santurri L, Ardizzone F, Guzzetti F (2011) Seasonal landslide mapping and estimation of landslide mobilisation rates using aerial and satellite images. Geomorphology 129:59–70. https://doi.org/10.1016/j.geomorph.2011.01.013

Gaidzik K, Ramírez-Herrera MT (2021) The importance of input data on landslide susceptibility mapping. Sci Rep 11:19334. https://doi.org/10.1038/s41598-021-98830-y

Glade T (2001) Landslide hazard assessment and historical landslide data — an inseparable couple. In: Glade T, Albini P, Francés F (eds) The use of historical data in natural hazard assessments. Springer, Dordrecht, pp 153–168

Grohmann CH (2015) Effects of spatial resolution on slope and aspect derivation for regional-scale analysis. Comput Geosci 77:111–117. https://doi.org/10.1016/j.cageo.2015.02.003

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31:181–216. https://doi.org/10.1016/S0169-555X(99)00078-1

Jade S, Sarkar S (1993) Statistical models for slope instability classification. Eng Geol 36:91–98. https://doi.org/10.1016/0013-7952(93)90021-4

Jagodnik P, Bernat Gazibara S, Arbanas Ž, Mihalić Arbanas S (2020a) Engineering geological mapping using airborne LiDAR datasets – an example from the Vinodol Valley, Croatia. J Maps 16:855–866. https://doi.org/10.1080/17445647.2020.1831980

Jagodnik P, Bernat Gazibara S, Jagodnik V (2020b) Types and distribution of quaternary deposits originating from carbonate rock slopes in the Vinodol valley, Croatia – new insight using airborne lidar data. Mgpb 35:57–77. https://doi.org/10.17794/rgn.2020.4.6

Krkač M, Bernat Gazibara S, Sinčić M, Lukačić H, Mihalić Arbanas S (2022) Landslide inventory mapping based on LiDAR data: a case study from Hrvatsko Zagorje (Croatia). In: Proceedings of the 5th Regional Symposium on Landslides in the Adriatic – Balkan Region. Faculty of Civil Engineering, University of Rijeka and Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb, pp 81–86

Lee S, Min K (2001) Statistical analysis of landslide susceptibility at Yongin, Korea. Environ Geol 40:1095–1113. https://doi.org/10.1007/s002540100310

Lee CF, Li J, Xu ZW, Dai FC (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. Environ Geol 40:381–391. https://doi.org/10.1007/s002540000163

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecol Biogeogr 17:145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Marković M (1983) Osnove primjenjene geomorfologije (Basics of applied geomorphology). Geoinstitut, Beograd

Pellicani R, Van Westen CJ, Spilotro G (2014) Assessing landslide exposure in areas with limited landslide information. Landslides 11:463–480. https://doi.org/10.1007/s10346-013-0386-4

Petschko H, Brenning A, Bell R, Goetz J, Glade T (2014) Assessing the quality of landslide susceptibility maps – case study Lower Austria. Nat Hazards Earth Syst Sci 14:95–118. https://doi.org/10.5194/nhess-14-95-2014

Razak KA, Straatsma MW, van Westen CJ, Malet J-P, de Jong SM (2011) Airborne laser scanning of forested landslides characterisation: terrain model quality and visualisation. Geomorphology 126:186–200. https://doi.org/10.1016/j.geomorph.2010.11.003

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. Earth-Sci Rev 180:60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Roşca S, Bilaşco Ş, Petrea D, Fodorean I, Vescan I, Filip S, Magut F-L (2015) Large scale landslide susceptibility assessment using the statistical methods of logistic regression and BSA – study case: the sub-basin of the small Niraj (Transylvania Depression, Romania). Landslides and Debris Flows Hazards

Rossi M, Reichenbach P (2016) LAND-SE: a software for statistically based landslide susceptibility zonation, version 1.0. Geosci Model Dev 9:3533–3543. https://doi.org/10.5194/gmd-9-3533-2016

Rossi M, Bornaetxea T, Reichenbach P (2021) LAND-SUITE V1.0: a suite of tools for statistically-based landslide susceptibility zonation. Earth and space science informatics

Sarkar S, Roy AK, Martha TR (2013) Landslide susceptibility assessment using Information Value Method in parts of the Darjeeling Himalayas. J Geol Soc India 82:351–362. https://doi.org/10.1007/s12594-013-0162-z

Šimunić A, Pikija M, Hečimović I (1982) Basic geological map, scale 1:100,000, Varaždin, Sheet 33-69

Sinčić M, Bernat Gazibara S, Krkač M, Mihalić Arbanas S (2022a) Landslide susceptibility assessment of the City of Karlovac using the

bivariate statistical analysis. Rudarsko-geološko-naftni zbornik 34(1):45–48. https://doi.org/10.17794/rgn.2022.2.13

Sinčić M, Bernat Gazibara S, Krkač M, Lukačić H, Mihalić Arbanas S (2022b) The use of high-resolution remote sensing data in preparation of input data for large-scale landslide hazard assessments. Land 11:1360. https://doi.org/10.3390/land11081360

URL-1: Https://Land.Copernicus.Eu/Pan-European/Corine-Land-Cover/Clc2018?Tab=download (CLC2018_CLC2012_V2018 12_20b2.Gdb). Accessed 15 June 2022

URL-2: Http://Arhiva.Vzz.Hr/Images/Stories/Prostorni-Plan/ LEPOGLAVA_ID/LEPOGLAVA_ID2/ K1_KORISTENJE_I_NAMJENA_POVRSINA_HTRS.Pdf. Accessed 15 June 2022

URL-3: Http://Arhiva.Vzz.Hr/Images/Stories/Prostorni-Plan/BEDNJA/ 2-ID-PPUO-Bednja-05-2017/K1_KORISTENJE_I NAMJENA_POVRSINA_ID2.Pdf. Accessed 15 June 2022

van Westen CJ, van Asch TWJ, Soeters R (2006) Landslide hazard and risk zonation—why is it still so difficult? Bull Eng Geol Environ 65: 167–184. https://doi.org/10.1007/s10064-005-0023-0

van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Eng Geol 102:112–131. https://doi.org/10.1016/j. enggeo.2008.03.010

Vojteková J, Vojtek M (2020) Assessment of landslide susceptibility at a local spatial scale applying the multi-criteria analysis and GIS: a case study from Slovakia. Geomatics Nat Hazards Risk 11:131–148. https://doi.org/10.1080/19475705.2020.1713233

Xing Y, Yue J, Guo Z et al (2021) Large-scale landslide susceptibility mapping using an integrated machine learning model: a case study in the Lvliang Mountains of China. Front Earth Sci 9:722491. https:// doi.org/10.3389/feart.2021.722491

Yin KL, Yan TZ (1988) Statistical prediction model for slope instability of methamorphosed rocks. In: Proceedings of Fifth International Symposium on Landslides. Lausanne, pp 1269–1272

Yusof NM, Pradhan B, Shafri HZM, Jebur MN, Yusof Z (2015) Spatial landslide hazard assessment along the Jelapang Corridor of the North-South Expressway in Malaysia using high resolution airborne LiDAR data. Arab J Geosci 8:9789–9800. https://doi.org/10.1007/ s12517-015-1937-x

Zaninović K, Gajić-Čapka M, Perčec Tadić M, Vučetić M, Milković J, Bajić A, Cindrić K, Cindrić L, Katušin Z, Kaučić D (2008) Climate atlas of Croatia 1961. – 1990., 1971. – 2000.

Zêzere JL, Pereira S, Melo R et al (2017) Mapping landslide suscepti-bility using data-driven methods. Science of The Total Environment 589:250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188