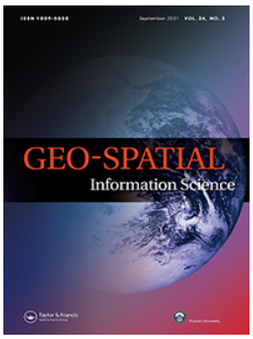

Geo-spatial Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgsi20

Landslide inventory validation and susceptibility mapping in the Gerecse Hills, Hungary

Dávid Gerzsenyi & Gáspár Albert

To cite this article: Dávid Gerzsenyi & Gáspár Albert (2021) Landslide inventory validation and susceptibility mapping in the Gerecse Hills, Hungary, Geo-spatial Information Science, 24:3, 498-508, DOI: 10.1080/10095020.2020.1870872

To link to this article: https://doi.org/10.1080/10095020.2020.1870872

GEO-SPATIAL INFORMATION SCIENCE
2021, VOL. 24, NO. 3, 498–508
https://doi.org/10.1080/10095020.2020.1870872

Landslide inventory validation and susceptibility mapping in the Gerecse Hills, Hungary

Dávid Gerzsenyi and Gáspár Albert

Department of Cartography and Geoinformatics, Eötvös Loránd University, Budapest, Hungary

ABSTRACT

Landslides pose a threat to property both in the populated and cultivated areas of the Gerecse Hills (Hungary). The currently available landslide inventory database holds the records from many sites in the area, but the database is out-of-date. Here we address the problem of revising the National Landslides Cadastre landslide inventory database by creating a landslide susceptibility map with a multivariate model based on likelihood ratio functions. The model is applied to the TanDEM-X DEM (0.4″ res.), the current landslide inventory of the area, and data acquired from geological maps. By comparing the distributions of four variables in the landslide and non-landslide area with grid computation methods, the model yields landslide susceptibility estimates for the study area. The estimations show to what extent a certain area is similar to the sample areas, therefore, its likelihood to be affected by landslides in the future. The accuracy of the model predictions was checked in the field and compared to the results of our previous study using the SRTM-1 DEM for a similar analysis. The model gave accurate estimates when certain correction measures were applied to the input datasets. The limitations of the model, the input datasets, and the suggested correction measures are also discussed.

ARTICLE HISTORY
Received 28 June 2018
Accepted 28 December 2020

KEYWORDS
Landslide susceptibility; landslide inventory; TanDEM-X; Gerecse Hills; landslides; geomorphometry

1. Introduction

Various slope movement processes are active today in the Gerecse Hills (Northern Central Hungary). These processes are posing problems both in the populated and the cultivated parts of the area (Lóczy, Balogh, and Ringer 1989; Szabó 2003), threatening built and agricultural property, especially in the north-western parts of the Gerecse Hills along the riverside bluffs of the Danube and in the steeper stream valleys descending toward the river (Schweitzer 1989). Dealing with the related problems requires knowledge of the location and the extent of the landslide-prone sites in the form of a landslide inventory database. An up-to-date landslide inventory and the thematic susceptibility and hazard maps derived from it can help landowners and local decision makers plan mitigation measures.

Landslides and other slope movements in the Gerecse area have been studied for a long time and many maps related to the topic were published (e.g. Gábris et al. 2018). However, these individual maps mainly focused only on a certain landslide site at a time and they cannot provide a comprehensive picture of the whole area regarding landslide susceptibility. The countrywide documenting of landslides in Hungary began in the 1970s under the framework of the National Landslides Cadastre (Pécsi, Juhász, and Schweitzer 1976). During the following years, the locations of the landslide sites were recorded on survey sheets along other information such as landslide-related damages and suggested mitigation measures (Kertész and Schweitzer 1991). The survey sheets were digitized in the 2000s and uploaded into the GIS database of the Hungarian National Landslide Cadastre. The landslide inventory database now holds records of many sites within the Gerecse area, but these records are outdated and inaccurate in many cases. Therefore, the revision of the inventory is clearly needed.

Modern remote sensing and GIS techniques and the broadening availability of higher resolution global DEM products (e.g. SRTM, TanDEM-X) allow the rapid morphometric analysis of landslide-related spatial phenomena. For identifying the location and extent of potential landslide areas, several morphometry-based susceptibility-mapping methods were developed (Guzzetti et al. 1999; Xie, Esaki, and Zhou 2004). Many of these methods require spatially continuous data providing quantitative (morphometric) and qualitative (geological) information about the processed area (Carrara et al. 1995, 1999; Davis, Chung, and Ohlmacher 2006). Davis, Chung, and Ohlmacher (2006) applied Chung's (2006) method on an area, which has similar geological and morphometric parameters as the Gerecse area in the present study. The model, developed originally by Chung and Fabbri (1993), yields landslide susceptibility estimations by analyzing the distribution of geomorphometric and geological

CONTACT Dávid Gerzsenyi beneriusgd@gmail.com

parameters as variables in the study area determining the characteristic features (according to the used parameters) of the local landslides. The authors used Chung's (2006) method for a previous landslide susceptibility analysis of Gerecse study area with the SRTM as base data (Gerzsenyi and Albert 2018).

The main goal of our research was to test the applicability of Chung's (2006) model for creating accurate landslide susceptibility maps of the study area with the materials currently available. The model compares the distribution of four variables (elevation, slope, aspect and surface geology) in the landslide-affected and the non-landslide-affected parts of the area and yields landslide susceptibility estimates. The second goal was to revisit the current landslide inventory database of the area based on the results and to suggest changes in the inventory where it is needed. The materials required for the analysis are a Digital Elevation Model (DEM) of the area for deriving the morphometric variables, geological maps of the area for creating the thematic geological data layer, and a landslide inventory marking the landslide-affected sites that are to be used as sample areas. The used data sources are the following:

- TanDEM-X DEM (0.4″, ~12 m resolution)
- National Landslides Cadastre, the Hungarian landslide inventory database
- Geological maps of the area.

In this study we explain the steps of creating a landslide susceptibility map for the Northwestern Gerecse, the most landslide-affected part of the hills, from the data preparation phase to the validation of the results in the field. The study also uses the results of a previous landslide susceptibility analysis about the whole Gerecse Hills area applying the same methods with the coarser resolution SRTM-1 DEM (Gerzsenyi and Albert 2018). Comparing the current results to the previous ones can provide useful information about the accuracy of the estimations. The suggested changes in the workflow and in the methodology are also discussed.

2. Geology of study area

The Gerecse Hills is in the northeastern part of the Transdanubian Mountain Range on the right side of the Danube River. Its northern border is the river itself, which flows to the east. The mean water level is approximately 107 m (a.s.l.). The areal extent of Gerecse Hills is around 850 km^2 with the maximum height of 633 m. Its central part is mostly forest covered, while on the lower parts, agricultural areas are more frequent. The surface is mainly composed of Pleistocene loess and Quaternary slope- and valley sediments, but at higher elevations, there are Oligocene, Eocene, and Mesozoic carbonate outcroppings (Gyalog and Síkhegyi 2005).

The oldest rocks on the surface are the late Triassic shallow-marine carbonates: Dachstein limestone and "Main" dolomite formations. Jurassic marine limestone, Cretaceous marine marl, and siliciclastic sediments of the continental slopes are also found in the area, mainly in the northern part of the hills (Császár et al. 2012). During the early Paleogene these carbonates were eroded, and they were buried under the various sequences of paralic coal-bearing layers, shallow-marine clayey-marls, and nummulitic limestones in the middle Eocene (Kercsmár and Fodor 2005). The re-exhumation of the Mesozoic limestones started during the Miocene, and as the result of multiphase tectonism, they now crop out on the horst areas. In the basins formed as structural grabens, Oligocene siliciclastic sediments (sand, marl, and conglomerate) are found. At the edges of these basins, littoral breccia, delta-sediments, and coal-bearing beds deposited during the Late Miocene to Early Pliocene (Magyar et al. 2013).

By the Late Pliocene, the sediment types of the paleo-Danube river became fluvial, and during the Pleistocene several terrace levels formed from incision processes (Pécsi 1971; Gábris 1994). Lacustrine limestones (travertines) covered large areas in the Northern regions of the Gerecse Hills in the Quaternary, and repeatedly deposited onto the fluvial sand and gravel (Pécsi 1971; Sierralta et al. 2010). During the glacial periods of the Quaternary, a periglacial semi-arid climate was present in the area; loess formed, burying all morphological features on the pediment areas (Albert et al. 2010). The now visible surface was formed by Quaternary slope and linear erosion, and Pleistocene wind erosion (Sebe et al. 2011). Most of the geological formations on the surface are the results of these processes: loess, deluvial and proluvial sediments.

The unconsolidated clastic sediments and the loess are important factors in landslides. Loess is the thickest on areas close to the Danube River, where it can exceed 15 m. On hillslopes, in the valleys, or on higher hills 5–10 m is the average thickness (Scharek et al. 2000). The combination of a basal clastic sediment layer and a thick loess cover create the most likely scenario for landslides, for hydraulic reasons (Kertész and Schweitzer 1991). The bigger the sediment load is on the unconsolidated terrace sediment, the more likely the landslide will occur. This and the proximity of the discharge points of the groundwater are the reasons why the high bluffs, close to the Danube River, are the most vulnerable to landslides in the Gerecse.

3. Landslide-related maps – clarification of terms

Before discussing landslide-related problems, the terms landslide susceptibility, hazard, and vulnerability must

be clarified, since these terms are often used with different meanings by different researchers. Based on the works of Varnes (1984), Carrara et al. (1995), and Parise (2001) we distinguish five main types of maps dealing with the former and future occurrence of landslides:

- Landslide inventory maps (or databases)
- Landslide activity maps
- Landslide susceptibility maps
- Landslide hazard maps
- Landslide vulnerability maps

Landslide inventory maps (or databases) mark the location of landslide-affected sites within a study area. Landslide inventories can either keep record of every known landslide in the studied area or only slope movements related to a certain event (e.g. heavy rainfall or earthquake). Besides their location, landslide inventory databases may also store information about the type and activity of the slope movements, along with suggested or already conducted mitigation measures, and potential risks.

Landslide activity maps depict the evolution of the landslides in an area within a given timeframe. They mostly focus only on the direction and speed of the moving material, and on the changes in the extent of the affected area, without providing information about the causes or the risks of the represented slope movement processes.

Landslide susceptibility maps show the likelihood of landslide occurrence in the study area. The level of landslide susceptibility is estimated by evaluating to what extent are certain factors causing landslides present in each areal unit. In susceptibility estimations, the temporal aspect of landslide occurrences is not always considered.

Landslide hazard maps provide information about the probability of landslide occurrence within a given timeframe in addition to the possible spatial distribution of future landslides.

Landslide vulnerability maps present the potential damage in real estate that can or that is likely to be caused by landslides estimated to occur over a given time. Such maps require contribution from experts outside of earth sciences to collect knowledge about the economic and social factors involved.

These landslide-related maps are interrelated: the landslide inventory database can be the base of a landslide susceptibility map for the area, which subsequently can be used to aid the revision works of the base inventory. Moreover, in the future by comparing the old and the new forms of landslide inventories, activity maps can be derived for certain sites. The susceptibility map can also serve as a basis for more complex landslide hazard or landslide vulnerability studies.

4. Data

The applied likelihood ratio function model (Chung 2006) analyzes the distribution of four variables (elevation, slope, aspect, and surface geology) on the landslide-affected and the non-landslide-affected parts of an area. These variables are considered here as the "analyzed variables". A fifth variable is used to mark the landslide-affected and the non-landslide-affected parts and is referred as the "area-defining variable".

The four analyzed variables and the landslide area-defining variable were acquired from various data sources. The morphometric parameters were derived from the TanDEM-X global coverage digital elevation model (approximate resolution: 12 m) with SAGA GIS routines. The digital form of the 1:100,000 Geological Map of Hungary (Gyalog and Síkhegyi 2005) was used as the base for the geological data layer. The landslide sites used as sample areas are from the Hungarian National Landslides Cadastre and from large-scale geological maps.

4.1. Geomorphometric parameters derived from DEM

The TanDEM-X DEM is a global coverage digital elevation model with the horizontal resolution of 0.4″ (approximately 12 m) and 2–4 m relative vertical accuracy. It was acquired between December 2010 and January 2015, and the DEM was derived from two complete coverage of the Earth. The model was corrected with the use of the water coverage mask that was provided with the DEM by masking the actual waterbodies. Since the DEM was acquired with X-band SAR interferometry the vegetation coverage was also present to a certain extent in the model (Wessel 2016; Gruber et al. 2016). The effect of the vegetation coverage was corrected by creating a forest coverage mask using publicly accessible aerial imagery taken in the same year as the elevation model acquisition. The elevation difference between the inner and outer borders of the forest polygons was computed, and the forest heights were interpolated from the estimated heights along the borders for the inner parts of the polygons. The forest height mask was then subtracted from the DEM. The ruggedness along the forest borders was smoothed with an averaging filter.

The corrected DEM became more suitable for the morphometric analysis, but the correction measures have to be taken into account when interpreting the results. The slope and aspect grids were derived from the corrected DEM using the Slope, Aspect, Curvature module (Zevenbergen and Thorne 1987) of the SAGA GIS (Conrad et al. 2015). The slope values are degrees from 0° to 90°. The aspect values are azimuth values between 0° and 360° counted clockwise from north.

4.2. Landslide inventory

The landslide inventory database of the Hungarian National Landslides Cadastre provided most of the landslide polygons used as sample data. The remaining part of the landslide marking data was acquired from recent geological maps. The survey of the Hungarian landslide inventory began in the 1970s, when not only the location of the landslide was recorded on the survey sheets but also information about their activity, type, geological and morphological characteristics, and suggested or already conducted hazard management procedures (Kertész and Schweitzer 1991). The survey sheets were later digitized. The Cadastre is available in spatial database form (containing approximately 1600 polygon and point features) and is accessible through WMS/WFS services. The database holds 28 records for the Northwestern Gerecse study area in polygon form but many of these polygons overlap each other to a certain extent. Since the current study only requires the spatial extent of the landslide as input data (and provides only the same as output) the overlapping parts were unified. In some cases, the boundaries of the landslide polygons were also adjusted to the actual topography based on geological maps of the area and previous field survey experiences of the team. At some places, larger-scale (1:50,000 and 1:25,000) manuscript maps of the local Geological Archives provided help to fine-tune the outline polygons of the landslides.

4.3. Geological data

Geological data were acquired from the 1:100,000 Geological Map of Hungary (Gyalog and Síkhegyi 2005). The dataset provides information about the superficial geology of the area. Based on the analogies of the work of Davis, Chung, and Ohlmacher (2006) and the generalized lithological categories of a smaller-scale map of the area (Budai and Gyalog 2010), the geological formations were sorted into six categories:

(1) Cemented sediments (Tertiary and older)
(2) Fluvial sediments
(3) Cemented carbonates (Quaternary)
(4) Loess
(5) Older carbonates
(6) Quaternary slope deposits

The polygons of the categorized geological features were then converted into grid form by assigning the appropriate category numbers to the grid cells overlapped by the polygons.

5. Method

We used Chung's (2006) model based on likelihood ratio functions to create the landslide susceptibility map (Figure 1). The here described version of the model needs two qualitative (thematic) and three quantitative data layers as input variables. The qualitative layers were the categorized geological features of the area and the landslide sites used as sample areas. Three geomorphometric parameters: elevation, slope, and aspect derived from a DEM were used as quantitative variables.

Since the model uses grid computation methods, the data layers containing the input parameters must be in the same grid system: layers must all have the same spatial resolution and the same projection. We chose the 12 m × 12 m resolution in the Hungarian National Grid (EOV, EPSG: 23700) as the preferred grid system. The elevation values of the DEM were already present in this raster grid system. The slope and aspect layers inherited the same system as they were derived from the elevation values. The qualitative layers were present as vector features at first. Therefore, each thematic category of the qualitative layers had to be assigned numeric values, these vector layers were converted into raster layers in the previously chosen 12 m resolution grid system. Category numbers from 1 to 6 were assigned to the categories of the geological thematic layer. The landslide polygons of the study area were converted into two grids. One

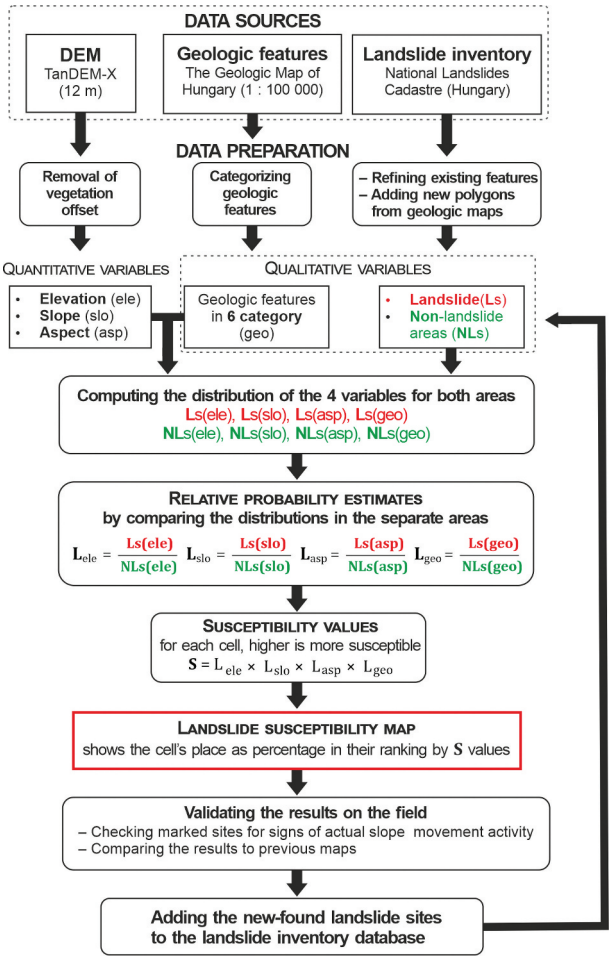

Figure 1. The workflow of the here used version of likelihood ratio functions model.

represented the landslide-affected parts by marking landslides with the value 1 and the remaining area with no-data. The non-landslide part was represented by the inverse of the latter grid.

After the data layers were converted into their appropriate grid form, the distributions of the elevation (**ele**), slope (**slo**), aspect (**asp**) and categorized geologic (**geo**) variables in both the landslide (**Ls**) and the non-landslide (**NLs**) parts of the area were computed. The first step is partitioning the grids of the four analyzed variables into a landslide and a non-landslide subarea. This masking operation is done by taking their cell-by-cell grid products with the corresponding area marking grid:

$$[Ls; NLs] \times [ele; slo; asp; geo] = Ls_{variable}; NLs_{variable} \quad (1)$$

Partitioning the grids of the analyzed variables using Equation (1) yielded 4–4 grids for the landslide and the non-landslide areas. Functions $Ls_{variable}$ and $NLs_{variable}$ represented the distribution functions of the four analyzed variables in landslide and the non-landslide area. The division of the $Ls_{variable}$ and $NLs_{variable}$ functions yielded relative probability estimates (likelihoods) regarding landslide occurrence for each value of the analyzed variables:

$$L_{variable} = Ls_{variable}/NLs_{variable} \quad (2)$$

These relative probability estimates are also conceivable as weights (from zero to infinity) assigned to each value of the variables representing their relative likelihood to be affected by landslides if only the given variable is considered. The computed weights were assigned to the values of the four analyzed grids.

Landslide susceptibility (**S**) values are derived by taking the grid product of the relative probability estimate grids of the analyzed variables:

$$\mathbf{S} = \mathbf{L_{ele}} \times \mathbf{L_{slo}} \times \mathbf{L_{asp}} \times \mathbf{L_{geo}} \quad (3)$$

The landslide susceptibility values are dimensionless numbers that show to what extent a certain grid cell is similar to the landslide pixels used as sample area. The higher the value, the more likely the cell will be affected by landslides in the future. On the landslide susceptibility map, the susceptibility values are shown as percentages, where n% means that the cell is more prone to landslides than n% of all cells (0%: least prone to landslides, 100% most prone to landslides).

6. Morphometric analysis of the area

6.1. Distributions of the analyzed variables

The data preparation phase yielded the six input grids with the same 12 m × 12 m resolution: two marking the landslide and the non-landslide areas and four containing the morphometric and geologic features of the study area. The landslide and the non-landslide parts of the analyzed variable grids were separated using the grid products for the marked areas and the analyzed grids (Equation (1)). The distributions of elevation (Figure 2), slope (Figure 3), aspect (Figure 4) and the categorized geologic features (Figure 5) were computed for these areas.

The elevation of the whole study area ranges from around 110 to 680 m in the DEM. The landslide-affected parts mostly fall into the range between 170 and 310 m (Figure 2). The average slope value in the non-landslide area is 9°, whereas in the landslide area the average slope is 11.8°. The curves of the slope values intersect at 7°, slopes steeper than 7° are relatively more likely in the landslide areas (Figure 3). The slopes are facing mostly southwest westward and northward in general, but slope descending toward southeast are relatively rare in the area. The landslide-affected slopes are mainly facing southwest westward, north, and northeastwards (Figure 4). The area is mainly covered with loess and fluvial sediments. The loess is even more

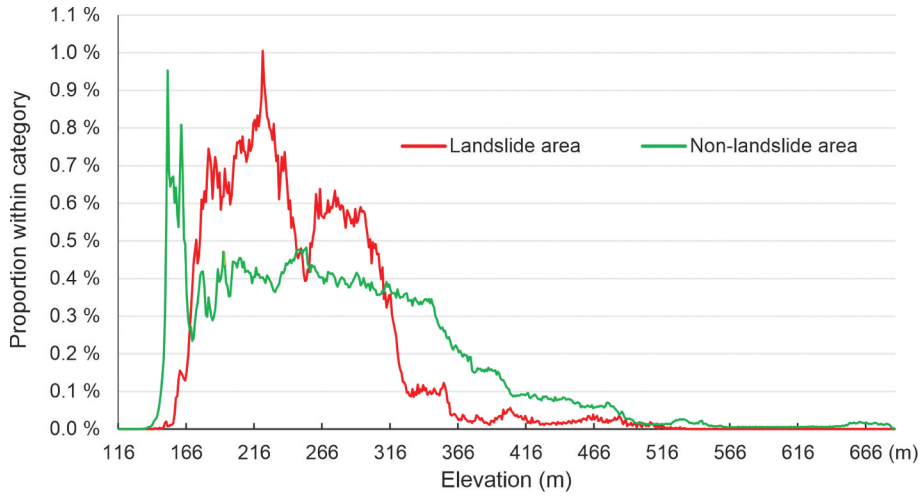

Figure 2. The distributions of elevation (m) values in the landslide and non-landslide area.

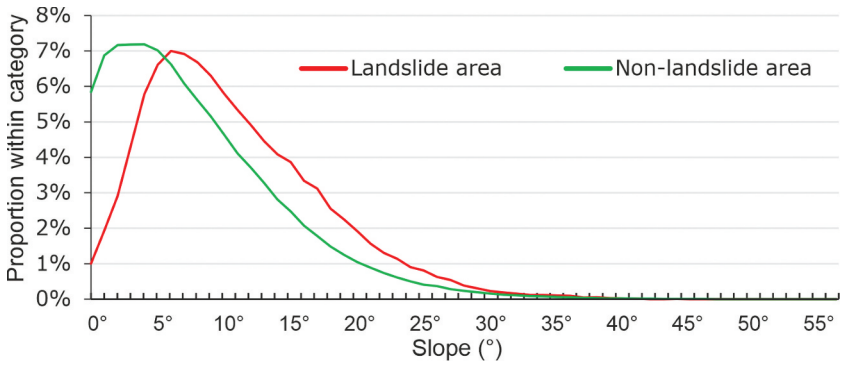

Figure 3. The distributions of slope (°) values in the landslide and non-landslide area.

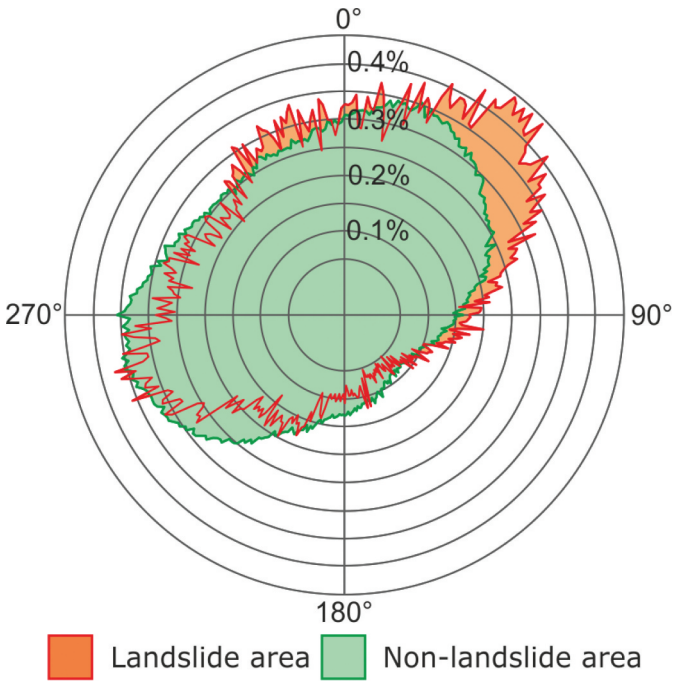

Figure 4. The distributions of aspect (°) values in the landslide and non-landslide area. Aspect directions clockwise: 0°: north, 90°: east, 180° south, 270°: west. Proportion within category: increasing outwards from the center.

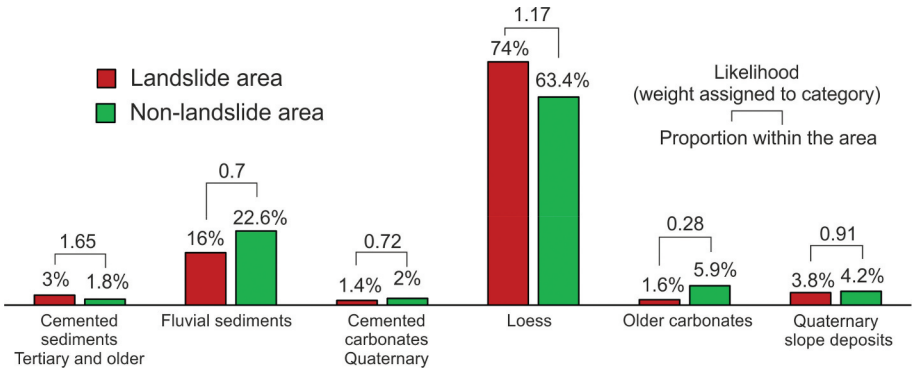

Figure 5. The distributions of categorized geologic features in the landslide and non-landslide area.

characteristic of the landslide-affected part than in general across the study area (Figure 5).

6.2. Relative probability estimates

After computing the distributions of the analyzed variables, weights were determined for each value of the variables by comparing the distributions of the two areas (Equation (2)). These weights were assigned to the four grids of the analyzed variables. Cells with weights less than 1 can be considered less likely to be affected by slope movements, and cells with greater weights are more likely according to the corresponding variable.

Elevation values had weights significantly above one in the 170–247 and 260–310 m ranges. These areas are more likely to be affected by landslides according to the model (Figure 6). Regarding slope values, weights above one were assigned to slopes steeper than 7°, and the slopes ranging from 16° to 30° got the highest relative likelihoods (Figure 7). Weights above one were assigned to slopes ranging from north to east and slopes facing in the northeast direction received the highest weights (Figure 8). It is notable that no aspect value had particularly low or high weights, making the aspect variable having less influence on the results. According to the model results, only cells in the Loess and the Cemented (older) sediments category were more prone to landslides (Figure 5). The Cemented (older) sediments category was assigned with the highest weight. However, only 3% of all cells fall into this category; therefore, unlike the Loess category, it does not have a significant impact on the results.

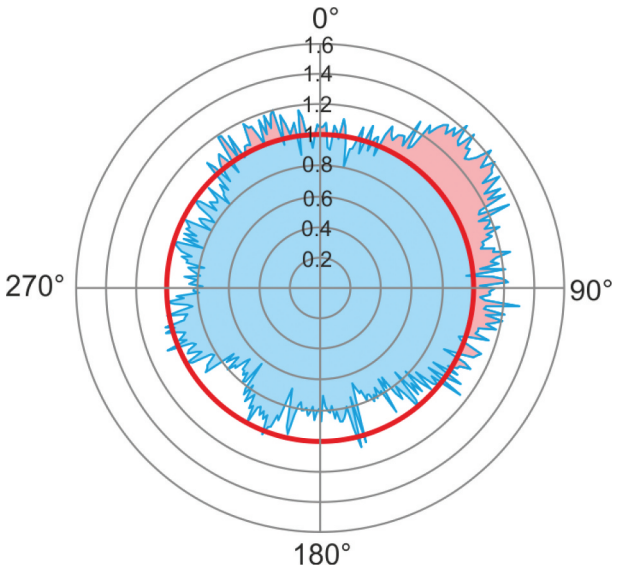

Figure 8. Weights assigned to the aspect (°) values. Aspect directions clockwise: 0°: north, 90°: east, 180° south, 270°: west. Weights: increasing outwards from the center.

6.3. Landslide susceptibility map

After assigning the estimated weights to the grid values, the next step was computing the susceptibility values by taking the grid product of the four grids containing the weights (Equation (3)). This yielded susceptibility values ranging from 0 to 9.54. Then, each value's normalized position in their ranking was assigned to the corresponding cells as percentages: 0% marked the cell with the lowest susceptibility value and 100% marked the cell with the highest susceptibility to landslides. The susceptibility values are presented in the percent form on the susceptibility map (Figure 9) as explained in the Method section.

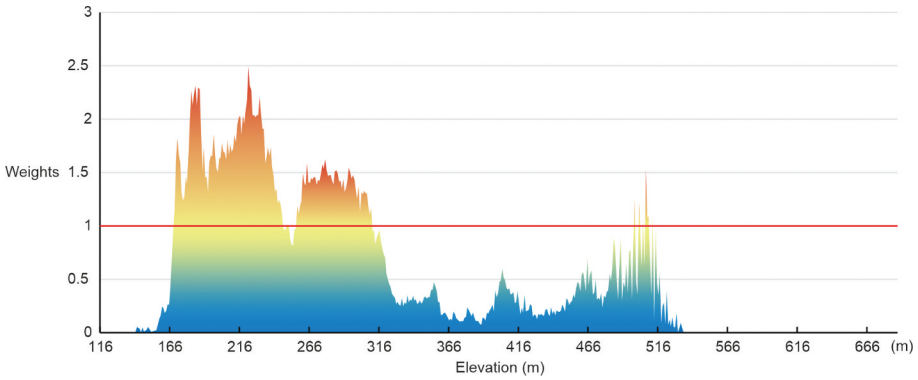

Figure 6. Weights assigned to the elevation (m) values.

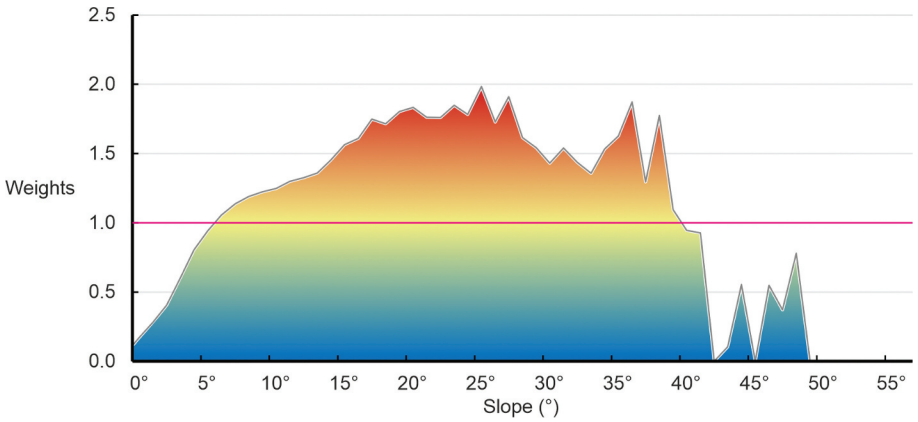

Figure 7. Weights assigned to the slope (°) values.

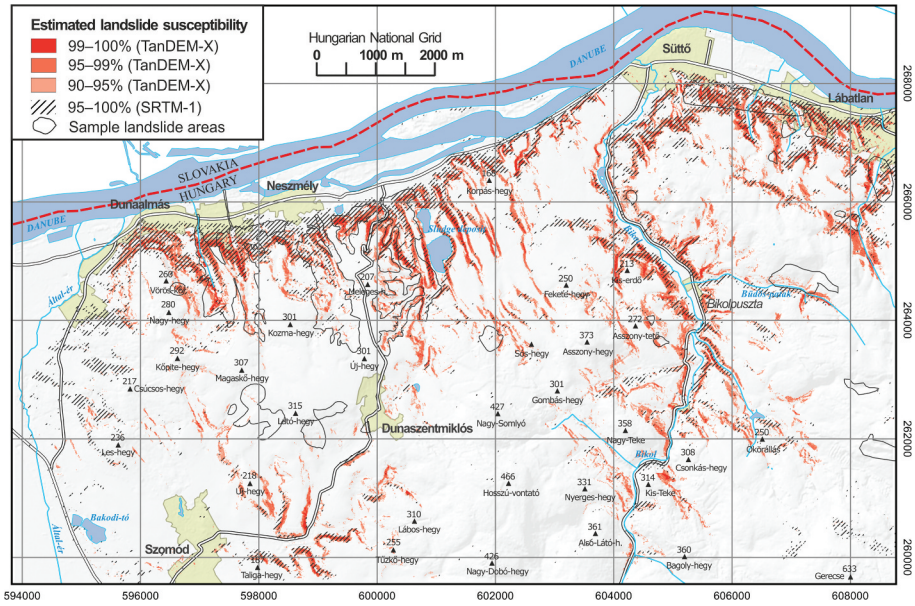

Figure 9. Landslide susceptibility map of the study area. Easting (X) and northing (Y) coordinates are in Hungarian National Grid projection (EPSG: 23700).

The landslide susceptibility map shows the location of the cells with 90% or higher susceptibility, one-tenth of the total study area considered the most prone to landslides. The map also marks the sample landslides and the areas marked as susceptible to landslides (95–100%) in our previous landslide susceptibility study of the area (Gerzsenyi and Albert 2018).

The areas marked as the most susceptible for landslides are the riverside bluffs along the Danube River and the steeper stream valleys descending toward the river. The susceptible areas also include the slopes alongside the valley of the Bikol Stream. The susceptible areas along the Danube overlap with the landslide areas used as sample areas in many cases. The sample areas near Dunaszentmiklós and Szomód villages have lower estimated susceptibility, possibly because they lie on gentler, mostly southward facing slopes.

7. Fieldwork

The main goal of the fieldwork was to evaluate the accuracy of the model predictions. The fieldwork also provided an opportunity to compare the accuracy of the susceptibility estimations to the results of our previous study (Gerzsenyi and Albert 2018). Due to the small number of records in the landslide inventory database, instead of using cross-validation methods, thorough ground-checks were carried out on the study area. The ground checking provided useful information about the accuracy of the model predictions. The fieldwork focused on two types of sites:

(1) Sites predicted to be susceptible to landslides by both the current and the previous model results.

(2) Sites predicted to be susceptible to landslides by only one of the model results.

In both cases, we were looking for signs of actual landslides or other slope movements in the surroundings of the sites marked with high susceptibility estimates. In the cases where only one of the results predicted high landslide susceptibility, our aim was to determine which model gave false high estimates. Besides the (fresh) landslide deposits, the signs of slope movements included soil creep, earth flow, debris flow, rock fall, topple, or the signs of fallen trees, and young new vegetation, or trees growing in odd, curved shapes. In the fieldwork, we used larger scale topographic maps with the susceptibility layer as an overlay because the contours of the topographic maps provided a more descriptive look at the overall topography than the contours derived from the DEM. We visited 28 sites during the field days. Table 1 lists the sites, as well as the actual state of slope movement activity on the sites, and the accuracy of the model predictions.

8. Results

The here presented work is the first demonstration of using the high-resolution DEM generated from the data of the TerraSAR-X and TanDEM-X satellites for environmental modeling in Hungary. The model was considered accurate where the observed state of the landslide location matched with the prediction of the model (Table 1). Only those model predictions were marked as active slope movements where more than a 75.0% relative probability of landslide activity was estimated. The current TanDEM-X-based model

Table 1. List of the visited sites that were marked by either of the models.

Site name	Slope movement activity	Model prediction (true if matches the activity)	
		SRTM-1	TanDEM-X
Old roman road (Szomód)	None	False	True
Old roman road II. (Szomód)	Present	True	True
Felső-nyárerdő forest (Dunaalmás)	None	True	False
Gully near Dunaalmás	Present	True, overestimated area	True
Kalács Hill (Szomód)	Present	True	True
Új Hill	Present to some extent	True	False
Borsós-diós (Szomód)	None (but marked on the geological maps)	True	True
Nyáros (Szomód)	Present to some extent	False	True
Dunaszentmiklós	Present, actual landslide in 2010	False	True
Bokros-dűlő (Dunaszentmiklós)	None (but marked on the geological maps)	True	True
Öreg Hill (Dunaszentmiklós)	None (but marked on the geological maps)	True	True
Bikol Creek Valley, upper	Present to some extent	True	True
Bikol Creek Valley (at Bikolpuszta)	Not characteristic	False	False
Forest near the Bikol Creek	None	False	False
Hajdú-temető (Büdös Creek)	None (but marked in the inventory)	False	False
Opposite side of Hajdú-temető	Present	True	True
Asszony-tető (Bikol Cr.)	Present	True	True
Leány-vár (Bikol Cr.)	Present	False	True
Szágodó-tető	Present in certain parts	False	True
Lábatlan Creek (valley)	Present to some extent	True	True
Izsány Valley	Present	True	True
Nyároska Valley	Present	True	True
Disznóskút Valley	Present	True	True
Total number of correct estimates		14.5	18
Total number of checked sites		24	24
Total rightness of model prediction		60.4%	75.0%

Scoring: false 0, overestimated area 0.5, true 1.

predictions were accurate in 75.0% of the cases, while the previous SRTM-based model only produced 60.4% accuracy on the checked sites. In the remaining cases, the estimated susceptibility was the result of inaccurate predictions. The lesser estimated susceptibility could also indicate that these areas were falsely marked as sites of slope movements in the inventory. We also found that in many cases, sites with high estimated susceptibility (over 90%) but without real signs of landslides were affected by other types of slope movement processes (mainly creeps).

One source of the inaccurate predictions were the artificial slopes in the DEMs on (mostly north or northeast facing) forest borders where the correction of the vegetation offset was not successful; however, this was significantly less characteristic in the case of the model predictions using the TanDEM-X. The inaccurate delineation of the actual landslide-affected areas was also identified as an error source from the input data sources in some cases. After identifying the source of these problems, the corresponding input datasets will be corrected in the future. Problems caused by anomalies in the DEM-based datasets can be reduced by using various filtering techniques on the elevation models. The errors caused by using inaccurate landslide polygons as input data should be corrected by editing the polygon boundaries to match the actual extent of the landslide-affected areas. Sites without any sign of slope movements should be ruled out from the input data. On sites where field checking is not possible the latest geothematic and topographic maps or remote sensing imagery can be used to search for signs of landslides.

9. Conclusions

After preparing the landslide susceptibility map of the Northwestern Gerecse Hills and validating the results on the field we came to the following conclusions:

Geomorphometric analyses need an accurate digital terrain model or a digital elevation model of approximately the same quality. Therefore, before using DEMs extracted from SRTM, and TanDEM-X data for geomorphometric analysis, the vegetation offset in models must be corrected. This task was carried out mostly successfully, making the TanDEM suitable for the analysis, but the correction measures have to be taken into account when interpreting the results.

To achieve accurate results, it was necessary to incorporate new landslide areas from the geological maps in addition to the landslide sites already marked in the National Landslides Cadastre. This measure provided the model with more comprehensive samples, and we were able to determine the characteristic features of the local landslide-affected areas more precisely. However, the ground checks made clear that in a few cases, the presence of slope movements is questionable, even on the sites marked on the recent the geological maps. Apart from the few questionable sample sites, the use of the additional sites and the correction of the landslide polygons enhanced the quality of the model predictions.

The landslide susceptibility estimations derived from the TanDEM-X DEM and other datasets can aid the revision of the current landslide inventory, as well as the discovery of the landslide-affected areas previously not recorded in the inventory. Keeping the inaccuracies

of the input datasets in mind, we concluded that the model is the most capable of giving accurate susceptibility estimates when it is used as a part of an iterative analyzing process. Thus, we suggest a workflow that includes the correction of the input landslide map: after the first run of the model, the boundaries of the input landslide sites should be corrected based on the validation works, the wrongfully marked areas should be deleted from the dataset, and the newly found sites should be recorded. The analysis should be conducted with the use of the new, corrected sample areas to provide accurate information about the characteristics of local landslide-affected areas.

Acknowledgments

The authors thank the anonymous reviewer and the editors for their work. The TanDEM-X data was provided by the German Aerospace Center (DLR), © DLR– 017.

Funding

The study was supported by the ÚNKP-17-2 New National Excellence Program of the Ministry of Human Capacities, Hungary [grant number ELTE/12421/65 (2017)]. This research was partly supported by the Thematic Excellence Programme, Industry and Digitization Subprogramme, NRDI Office [grant number ED_18-1-2019-0030].

Notes on contributors

Dávid Gerzsenyi is a PhD student of the Eötvös Loránd University (Budapest, Hungary). He holds an MSc degree in Cartography and a BSc degree in Earth Sciences with cartography and GIS specialization. His research topic is about the GIS-based landscape evolution modeling of the Gerecse Hills (Hungary).

Gáspár Albert is a geologist, cartographer and GIS specialist. He currently works at the Eötvös Loránd University (Budapest, Hungary) as a lecturer teaching basic- and geoscientific-mapping and modeling. He works on several research and exploration projects requiring geoinformatics.

ORCID

Dávid Gerzsenyi http://orcid.org/0000-0002-3453-6079
Gáspár Albert http://orcid.org/0000-0002-1723-8328

Data availability statement

The final product of the analysis (Landslide Susceptibility Map of the Northwestern Gerecse Hills (Hungary)) that supports the findings of this study is openly available in figshare (figshare.com) at https://figshare.com/articles/dataset/Landslide_Susceptibility_Map_of_the_Northwestern_Gerecse_Hills_Hungary_/12942686.

The TanDEM-X data (https://map.mbfsz.gov.hu/fdt100/) was provided by the German Aerospace Center (DLR) to D. Gerzsenyi for scientific use specified in proposal DEM_GEOL2269. © DLR–2017; The Hungarian National Landslides Cadastre [Országos Felszínmozgás Kataszter] landslide inventory is publicly available from the Mining and Geological Survey of Hungary (https://mbfsz.gov.hu/hatosagi-ugyek/nyilvantartasok/orszagos-felszinmozgas-kataszter); The Geological Map of Hungary (scale: 1:100,000, Gyalog and Síkhegyi 2005) is available from the Mining and Geological Survey of Hungary (webmap version: https://map.mbfsz.gov.hu/fdt100/). In addition, all the derived data based on these original data sources are available from the corresponding author upon reasonable request.

References

Albert, G., G. Csillag, L. Fodor, Z. Kercsmár, and I. Selmeczi. 2010. "Paleomorfológiai Megfigyelések a Gerecse Északkeleti Előterében [Paleomorphological Observations on the Northeastern Forelands of the Gerecse Hills]." In *Poster presented at Geográfia 2010 conference*, Pécs, Hungary. November 4–6. doi:10.13140/RG.2.2.13256.01288.

Budai, T., and L. Gyalog. 2010. *Geological Map of Hungary for Tourists*. Budapest: Geological Institute of Hungary.

Carrara, A., M. Cardinali, F. Guzzetti, and P. Reichenbach. 1995. "GIS Technology in Mapping Landslide Hazard." In *Geographical Information Systems in Assessing Natural Hazards*, edited by A. Carrara and F. Guzzetti, 135–175. Dordrecht: Springer. doi:10.1007/978-94-015-8404-3_8.

Carrara, A., F. Guzzetti, M. Cardinali, and P. Reichenbach. 1999. "Use of GIS Technology in the Prediction and Monitoring of Landslide Hazard." *Natural Hazards* 20 (2–3): 117–135. doi:10.1023/A:1008097111310.

Chung, C. 2006. "Using Likelihood Ratio Functions for Modeling the Conditional Probability of Occurrence of Future Landslides for Risk Assessment." *Computers & Geosciences* 32 (8): 1052–1068. doi:10.1016/j.cageo.2006.02.003.

Chung, C., and A. G. Fabbri. 1993. "The Representation of Geoscience Information for Data Integration." *Nonrenewable Resources* 2 (2): 122–139. doi:10.1007/BF02272809.

Conrad, O., B. Bechtel, M. Bock, H. Dietrich, E. Fischer, L. Gerlitz, J. Wehberg, V. Wichmann, and J. Böhner. 2015. "System for Automated Geoscientific Analyses (SAGA) V. 2.1.4." *Geoscientific Model Development* 8 (7): 1991–2007. doi:10.5194/gmd-8-1991-2015.

Császár, G., J. Haas, O. Sztanó, and B. Szinger. 2012. "From Late Triassic Passive to Early Cretaceous Active Continental Margin of Dominantly Carbonate Sediments in the Transdanubian Range, Western Tethys." *Journal of Alpine Geology* 54: 33–99.

Davis, J., C. Chung, and G. Ohlmacher. 2006. "Two Models for Evaluating Landslide Hazards." *Computers & Geosciences* 32 (8): 1120–1127. doi:10.1016/j.cageo.2006.02.006.

Gábris, G. 1994. "Pleistocene Evolution of the Danube in the Carpathian Basin." *Terra Nova* 6 (5): 495–501. doi:10.1111/j.1365-3121.1994.tb00893.x.

Gábris, G., M. Pécsi, F. Schweitzer, and T. Telbisz. 2018. "Relief." In *National Atlas of Hungary: Natural Environment*. edited by K. Kocsis, 42–57. Budapest: Geographical Institute MTA CSFK. http://www.nemzetiatlasz.hu/MNA/National-Atlas-of-Hungary_Vol2_Ch4.pdf

Gerzsenyi, D., and G. Albert. 2018. "Csuszamlásveszélyes Lejtők Helyének Becslése a Gerecsében Tapasztalati Változók Alapján." *[Landslide Susceptibility Estimations in the Gerecse Hills]. Földrajzi Közlemények* 142 (1):

21–35. https://www.foldrajzitarsasag.hu/downloads/foldrajzi_kozlemenyek_2018_142_evf_1_pp_021.pdf

Gruber, A., B. Wessel, M. Martone, and A. Roth. 2016. "The TanDEM-X DEM Mosaicking: Fusion of Multiple Acquisitions Using InSAR Quality Parameters." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (3): 1047–1057. doi:10.1109/JSTARS.2015.2421879.

Guzzetti, F., A. Carrara, M. Cardinali, and P. Reichenbach. 1999. "Landslide Hazard Evaluation: A Review of Current Techniques and Their Application in A Multi-Scale Study, Central Italy." *Geomorphology* 31 (1–4): 181–216. doi:10.1016/S0169-555X(99)00078-1.

Gyalog, L., and F. Síkhegyi. 2005. "Geological Map of Hungary." (Scale: 1:100000), Budapest, Hungary: Geological Institute of Hungary.

Kercsmár, Z., and L. Fodor. 2005. "Syn-Sedimentary Deformations in the Eocene Tatabánya Basin, Central Hungary." *Geolines* 19: 60–61. http://geolines.gli.cas.cz/fileadmin/volumes/volume19/G19-060.pdf

Kertész, Á., and F. Schweitzer. 1991. "Geomorphological Mapping of Landslides in Hungary with a Case Study on Mapping Danubian Bluffs." *Catena* 18 (5): 529–536. doi:10.1016/0341-8162(91)90053-Z.

Lóczy, D., J. Balogh, and Á. Ringer. 1989. "Landslide Hazard Induced by River Undercutting along the Danube ." *Geomorphological Hazards, Supplements of Geografia Fisica E Dinamica Quaternaria* 2: 5–11.

Magyar, I., D. Radivojević, O. Sztanó, S. Rastislav, K. Ujszászi, and M. Pócsik. 2013. "Progradation of the Paleo-Danube Shelf Margin across the Pannonian Basin during the Late Miocene and Early Pliocene ." *Global and Planetary Change* 103: 168–173. doi:10.1016/j.gloplacha.2012.06.007.

Parise, M. 2001. "Landslide Mapping Techniques and Their Use in the Assessment of the Landslide Hazard." *Physics and Chemistry of the Earth – Part C: Solar Terrestrial & Planetary Science* 26 (9): 697–703. doi:10.1016/S1464-1917(01)00069-1.

Pécsi, M. 1971. "The Development of the Hungarian Section of the Danube Valley." *Geoforum* 2 (2): 21–32. doi:10.1016/0016-7185(71)90003-0.

Pécsi, M., Á. Juhász, and F. Schweitzer. 1976. "Magyarországi Felszínmozgásos Területek Térképezése [Mapping of landslide-affected Areas in Hungary]." *Földrajzi Értesítő [Hungarian Geographical Bulletin]* 25 (2–4): 223–235. http://real-j.mtak.hu/5006/1/FoldrajziErtesito_1976.pdf

Scharek, P., P. Hermann, M. Kaiser, and J. Pristaš. 2000. "Map of Genetic Types and Thickness of Quaternary Sediments." In *Dunube Region Environmental Geology Programe*, edited by G. Császár, Jahrbuch der Geologischen Bundesanstalt. 142 (4): 447–455. Vienna, Austria: Geologischen Bundesanstalt. https://opac.geologie.ac.at/ais312/detail.aspx

Schweitzer, F. 1989. "Surface Movements of the Gerecse Mountain's Northern Part." In *Selected Environmental Studies*, edited by J. Galambos, 34–35. Budapest: Geographical Research Institute of HAS.

Sebe, K., G. Csillag, Z. Ruszkiczay-Rüdiger, L. Fodor, E. Thamó-Bozsó, P. Müller, and R. Braucher. 2011. "Wind Erosion under Cold Climate: A Fossil Periglacial Mega-Yardang System in Central Europe (Western Pannonian Basin, Hungary)." *Geomorphology* 134 (3–4): 470–482. doi:10.1016/j.geomorph.2011.08.003.

Sierralta, M., S. Kele, F. Melcher, U. Hambach, J. Reinders, R. van Geldern, and M. Frechen. 2010. "Characterisation and Uranium-series Dating of Travertine from Süttő in Hungary." *Quaternary International* 222 (1–2): 178–193. doi:10.1016/j.quaint.2009.04.004.

Szabó, J. 2003. "The Relationship between Landslide Activity and Weather: Examples from Hungary." *Natural Hazards and Earth System Science* 3 (1/2): 43–52. doi:10.5194/nhess-3-43-2003.

Varnes, D. J. 1984. *Landslide Hazard Zonation: A Review of Principles and Practice*. Paris: UNES.

Wessel, B. 2016. *TanDEM-X Ground Segment – DEM Products Specification Document*. Oberpfaffenhofen, Germany: EOC, DLR. https://elib.dlr.de/108014/

Xie, M., T. Esaki, and G. Zhou. 2004. "GIS-based Probabilistic Mapping of Landslide Hazard Using a Three-dimensional Deterministic Model." *Natural Hazard* 33 (2): 265–282. doi:10.1023/B:NHAZ.0000037036.01850.0d.

Zevenbergen, L. W., and C. R. Thorne. 1987. "Quantitative Analysis of Land Surface Topography." *Earth Surface Processes and Landforms* 12 (1): 47–56. doi:10.1002/esp.3290120107.