An easy-to-use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm

A. Akgun, E.A. Sezer, H.A. Nefeslioglu, C. Gokceoglu, B. Pradhan

1. Introduction

According to a review of the international landslide literature (Gokceoglu and Sezer, 2009), the most frequently used terms in this field are "landslide susceptibility" and "Geographical Information Systems, GIS." Several possible reasons mainly include: (a) developments in computer techniques and Geographical Information Systems; (b) increasing awareness of the socio-economic significance of landslides; and (c) the complex nature of landslides. Due to the complex nature of landslides, several methods have been applied for the assessment of landslide susceptibility at a medium scale. In general, these methods can be classified into three main groups: statistical (e.g., Can et al., 2005; Gokceoglu et al., 2005; Lee, 2005; Clerici et al., 2006; Duman et al., 2006; Pradhan et al., 2006, 2008; Lee and Pradhan, 2006; Akgun and Bulut, 2007; Akgun et al., 2008; Tunusluoglu et al., 2008; Lamelas et al., 2008; Gokceoglu and Gokceoglu, 2001, 2002; Pradhan et al., 2006, 2008; Lee et al., 2006; Turrini and Visintainer, 1998, 1999; Gokceoglu and Aksoy, 1996; Gokceoglu and Aksoy, 1996; Turrini and Visintainer, 1998, 1999; Aynayev and Barbieri, 2005). In the literature of this field, it is possible to find attempts to assess landslide susceptibility based on expert opinion (Juang et al., 1992; Saboya et al., 2006; Ercanoglu et al., 2008) without constructing an inference system. However, all of the available methods for landslide assessment at a medium scale (~1:25,000) are associated with uncertainties arising from a lack of knowledge and high variability (Ercanoglu and Gokceoglu, 2002). For these reasons, certain generalizations...
and simplifications are applied during data-driven analyses for landslide susceptibility mapping at a medium scale. Additionally, these data-driven methods do not require an expert opinion. The data-driven methods include a training phase. After training, testing, and validation stages are employed for assessing the quality of the model. There are a number of disadvantages associated with this type of process. First, for the selection of the training data and the modeling the system uses different training sets. Second, the results produced are dependent on the dataset employed. Third, the methods used, such as decision trees, support vector machines, or artificial neural networks, are black box approaches, and the user is completely isolated from the learning stage. To increase the role of expert opinion in such methods and to provide new insight related to reusable model production for landslide susceptibility assessment, the Mamdani fuzzy inference system (FIS) appears to be a potentially suitable approach.

FIS are capable of reducing uncertainties when solving complex problems, and the advantages of FIS are summarized by Alvarez Grima (2000) as follows:

(a) FIS allows explicit expression of knowledge of the system via fuzzy “if–then” rules;
(b) FIS deals with the subjective uncertainty (fuzziness, vagueness, and imprecision) inherent in the way experts approach their problems;
(c) FIS can combine numerical and categorical data; and
(d) FIS provides a sound mathematical basis.

In the last decade, the Mamdani FIS has been used extensively to solve complex and nonlinear problems of engineering geology. An interesting and perhaps the most attractive characteristic of fuzzy models compared with other methods commonly used in geosciences, such as statistics, is that they are able to describe such problems in a transparent way (Setnes et al., 1998). A Mamdani FIS (Mamdani and Assilian, 1975) is a transparent and expert opinion-based system. Although several studies using this type of system (i.e., Den Hartog et al., 1997; Alvarez Grima and Babuska, 1999; Gokceoglu, 2002; Kayabasi et al., 2003; Sonmez et al., 2003, 2004; Nefeslioglu et al., 2003, 2006; Gokceoglu et al., 2009a; Yagiz and Gokceoglu, 2010) have been published related to various branches of engineering geology, a Mamdani FIS has been applied for a limited number of landslide susceptibility assessments to date (Lee and Juang, 1992; Juang et al., 1992; Saboya et al., 2006). Lee and Juang (1992) and Juang et al. (1992) proposed a quantitative assessment method to evaluate the slope failure potential. Using fuzzy sets and linguistic rules, they (Lee and Juang, 1992; Juang et al., 1992) assessed the slope failure potential of relatively large areas in Hong Kong. In 2006, Saboya et al. (2006) used the Mamdani FIS for assessment of failure susceptibility of soil slopes. Saboya et al. (2006) constructed a Mamdani FIS by considering the opinions of the experts. According to the authors (Saboya et al., 2006), the conception of intelligent models by means of fuzzy logic in geotechnical engineering is quite appropriate as long as in this expertise area, the judgement, induction, deduction, and a sense of importance are always present. Although useful for assessment of landslide susceptibility, the Mamdani FIS has not been employed commonly due to the difficulties when applied in large areas without some proper computer program. For this reason, the purposes of the present study are to develop a MATLAB program for the construction of a Mamdani FIS and to apply it for landslide susceptibility evaluation at a medium scale. The present study includes four main stages such as (i) development of MATLAB program, (ii) data production; (iii) construction of the FIS and landslide susceptibility mapping; and (iv) assessment of the results. The major benefits to be obtained from the study are as follows:

(a) The Mamdani FIS can be constructed completely based on expert opinion without an exhaustive data analysis process.
(b) An expert can reflect his/her opinions on landslides in a region in the model via fuzzy if–then rules.
(c) The computational load is low, and results can be obtained in a short time.
(d) Using the MATLAB program developed in this study, a Mamdani FIS can be constructed, and production of the degrees of landslide susceptibility can be easily obtained.

2. MATLAB program for Mamdani FIS

As noted above, several studies applying a Mamdani FIS have been performed using MATLAB and published by a number of researchers (Lee and Juang, 1992; Juang et al., 1992; Den Hartog et al., 1997; Alvarez Grima and Babuska, 1999; Gokceoglu, 2002; Kayabasi et al., 2003; Sonmez et al., 2004; Saboya et al., 2006; Nefeslioglu et al., 2006; Gokceoglu et al., 2009a; Yagiz and Gokceoglu, 2010) in the engineering geology literature. According to Setness et al. (1998), fuzzy models are described in the literature as transparent and interpretable, but in fact, in many engineering applications, they are used as a black-box tool. Various fuzzy inference systems have been proposed. The Mamdani FIS (Fig. 1) is one of these systems. The Mamdani FIS is perhaps the most appealing fuzzy method to employ for engineering geology problems (Alvarez Grima, 2000).

The general characteristics of a Mamdani FIS were described by Alvarez Grima (2000), and some general information on the Mamdani FIS is provided in this section. A Mamdani FIS is composed of the membership functions of input(s), fuzzy if–then linguistic rules, and output membership functions. The general if–then structure of the Mamdani algorithm is given as

\[ R_i : \text{if } x_i \text{ is } A_{i1} \text{ and } \ldots \text{then } y \text{ is } B_i (i = 1, 2, \ldots, k), \]

where \( k \) is the numbers of rules; \( x_i \) is the input variable (antecedent variable); and \( y \) is the output variable (consequent variable).

The details of the composition of fuzzy relationships were given by Ross (1995). Although many methods for the composition of fuzzy relations (e.g., min–max, max–max, min–min, max–min, etc.) are available in the literature, the max–min and max–product methods are the two most commonly used techniques (Ross, 1995). The basic form of a fuzzy composition process is given as

\[ B = A \circ R, \]

where \( A \) is the antecedent defined on universe \( X \); \( B \) is the consequent defined on universe \( Y \); and \( R \) is the fuzzy relation characterising the relation between specific inputs \( (X) \) and specific outputs \( (Y) \).

The calculation procedure of a Mamdani FIS can be given as follows (Alvarez Grima, 2000):

(1) Compute the degree of fulfilment \( (z_i) \) of the input for each rule \( i \) by considering the degree of membership \( (\mu) \)

\[ z_i = \min \left( \mu_{B_{i1}}(X_1), \mu_{B_{i2}}(X_2), \ldots, \mu_{B_{ik}}(X_k) \right). \]

(2) For each rule, derive the output fuzzy set \( B_i \) using the minimum \( t \)-norm

\[ \mu_{B_i}(Y) = z_i \land \mu_{B_i}(Y). \]
Aggregate the output fuzzy sets by taking the maximum:

$$
\mu_B = \max_{i=1,2,...,k} \left( \mu_{B_i}(y) \right),
$$

The final stage of the construction of a fuzzy inference system is to select the defuzzification method. Aggregation of two or more fuzzy output sets gives a new fuzzy set in the basic fuzzy algorithm. In most cases, the result in the form of a fuzzy set is converted into a crisp result by the defuzzification process (Berkan and Trubatch, 1997). Due to its common use in practice and computational simplicity, the center-of-gravity method (Eq. (6)) is considered for use in the defuzzification process.

$$
y = \frac{\int_{s} B(y) y dy}{\int_{s} B(y) dy},
$$

where \( s \) is the support for the fuzzy set \( B(y) \).

As stated previously, although the Mamdani FIS has been applied to various engineering geology problems, it has not been applied for landslide susceptibility assessment except for a few studies (Lee and Juang, 1992; Juang et al., 1992; Saboya et al., 2006). When considering the nature of landslide susceptibility mapping at a medium scale, the appropriateness of the Mamdani FIS for this problem is unargued. This is because landslide susceptibility assessment is carried out using a very large geographic dataset and several conditioning parameters. When the number of parameters increases, the number of fuzzy sets and if–then rules also increases. Additionally, obtaining results for each case is impossible. However, a simple and applicable tool for the elimination of these problems has not previously been reported. For this reason, the MATLAB program MamLand was developed and applied to the case of Sinop (Northern Turkey). Using MamLand, a Mamdani FIS can be generated, and a very large dataset can be processed easily.

MamLand was developed as a MATLAB 7.8 (R2009a) (MATLAB, 2009) application. Thus, it requires MATLAB to run, as it is developed for wrapping MATLAB functionalities by employing useful user interfaces for the purpose of landslide susceptibility map production. There are three MATLAB files in MamLand (Fig. 3), including the main, modeling, and inference interfaces. The main user interface enables the user to choose a modeling or inference section.

The modeling interface of the modeling module of MamLand is designed from the bottom up to form an FIS model (Fig. 2).

The main input for this module is an Excel format file including the planned model properties. Considering the commonness and readability of Excel files, a template is designed in Excel file format to carry the description of the model by the user (Fig. 3a). This template is thought of as a deliverable part of MamLand, but it can also be constructed easily by a user because of its simplicity. After completing the necessary information for the planned model in the Excel file template, it is supplied as the input to the modeling module using the names of the files that will be used for the storage of the model and the rules draft. The storage of the model is a process that is generally employed, but the storage of a rules draft is a special requirement for the purpose of MamLand. A number of parameters are used in the landslide susceptibility assessment process, and as a result, the number of rules can be very high. For example, 192 rules are used in this study, and there are only 2 membership functions used for most of the parameters. It is known that the increment in the number of fuzzy sets directly triggers the increment in the number of rules. Additionally, an expert can express his/her knowledge with linguistic terms that are used in the model. By considering all of these practices, the conditional portion of the rules is automatically produced in linguistic form and stored in a comma-separated text file for the purpose of compatibility with various applications (Fig. 3b). After an expert user constructs the consequent portions of the rules, he/she can import the rules, which are stored in an Excel file (Fig. 3c). Meanwhile, if an expert user wants to add, delete, or modify the conditional portions of the rules, he/she can do these activities easily because the produced matrix is fully compatible with the MATLAB rule matrix. After completing the rules for the model, all of the information related to the model is stored, and the model becomes ready to use. Thus, the first user can describe the model in an easy manner and transfer all relevant information to MATLAB with a minor effort. A second user (or the first user) has a comfortable environment in which to describe the rules because the main portion of the rules is automatically produced using the expert’s experience.

**Fig. 1.** A generalized scheme of the Mamdani FIS structure.
The expert fills in the consequent portions of the rules and uploads them easily, instead of constructing all of the rules individually from the MATLAB FIS GUI. When the numbers of rules increase, the importance of this ability increases.

The inference interface of the inference module of MamLand is highly simplified, as there is no complex process involved for the users. Additionally, the users of this module do not need to be experts on landslides. In the inference stage, the users select the model to be used for inference; then they select the data to be used; and finally, they specify the output file to be used for the storage of the degrees of landslide susceptibility. After completing these stages, the users can trigger the inference with the relevant button on the interface, and they receive a message about the status of the inference. After the inference is complete, if the users wish to observe the saturation of the output values, MamLand provides this information via a bar chart. The reason for displaying the saturation of the output values is to provide fast feedback for the users. In this graph, the output range is divided into the number of output fuzzy sets, and intervals are justified according to the ranges of these fuzzy sets. However, they are used only to divide the output range into meaningful intervals, and they do not directly symbolize the output fuzzy sets. In this chart, the users can see the number of the outputs located in each interval. This saturation can reflect information about the rule organization of the model or the number of cases included in the field data.

3. Study area

The susceptibility assessment in this study is performed in Sinop city and its near vicinity in the northernmost part of the Black Sea region of Turkey, located 42°06’00” and 41°45’00” in the north and 34°57’00” and 35°15’00” in the east (Fig. 4) and occupying an area of approximately 392 km².
Rainfall is the main source of water in the study area, and it is the most important reason for triggering landslides. According to meteorological data for the period 1970–2009, the study area receives an average annual rainfall of 672 mm. Most of the rainfall occurs during October, with a mean value of 91.3 mm (General Directory of Meteorological Services of Turkey, 2010). Topographical and morphological determinations using a digital elevation model (DEM) produced by interpolating the contour lines of a topographic sheet at a scale of 1:25,000 (Fig. 5) showed that topographical elevations range from 0 to 1040 m, and the slope angle reaches 88° in some locations in the area. Whereas the northern part of the area exhibits a gently sloping topography, the slopes in the southern region are steep, due to lithological characteristics.

The geological map used in this study had a scale of 1:100,000, and it was prepared by the General Directorate of Mineral Research and Explorations (2008) (Fig. 6). In the study region, various lithological units crop out ranging from the Campanian–lower Maastrichtian to the Quaternary. Since the late Campanian, the basin has been affected by compressional tectonic flysch-type sediments of the Cankurtaran formation and its volcanic component, which are the products of arc magmatism deposition (Gedik et al., 1981). Following the deposition of the Cankurtaran formation, during the late Maastrichtian–Paleocene, clayey limestone, marl (Akveren formation) in late Maastrichtian–Paleocene sandstone, marl (Atbasi formation) in Lutatian siltstone, and sandstone–claystone (Kusuri formation) were deposited (Gedik et al., 1981). This sedimentary sequence from bottom to top indicates regressive development in general. The sandstone beds concentrated at the lower levels of the Kusuri formation have been designated an Ayancik member. These tectonic and volcanic activities in the north have caused the stratigraphy around the Sinop Peninsula to differ somewhat from that of neighboring areas. In the late Campanian–Maastrichtian period, the Hamsaros formation, associated with volcanic activity, and the Karaada formation, consisting of shallow reefal limestone, were deposited (Gedik et al., 1981). After the Middle Eocene, the area became land. Around Sinop and Gerze, the Sinop formation from the Miocene, which exhibits various facies of shore, tidal, lagoonal, and deltaic environments, the Sarıkum formation of the late Pliocene–early Pleistocene, consisting of meandering river and floodplain sediments, and the Kale formation of the late Pleistocene, consisting of shore sediments, were deposited (Barka et al., 1985). The Holocene is represented by alluvial and beach sediments deposited in the northern parts of the study area.

The study area is one of the most landslide-prone areas of Turkey, and it has been subjected to a number of landslides. According to the classification proposed by Cruden and Varnes (1996), the type of these landslides is generally active deep-seated rotational, and they usually occur on slopes above rivers and creeks. Additionally, shoreline landslides are also common in the area.
The occurrence of landslides is mainly controlled by the lithology and slope gradient in the area. Surface-water erosion and undercutting of slopes by both surface water and cliff erosion are also considered as other important landslide-controlling and triggering factors. Due to the weathering of lithological units in the area, the occurrence of clay is common, and the presence of clays also contributes to the occurrence of landslides.

4. Data production

Among the parameters included in this model, a reliable landslide inventory describing the type, activity, and spatial distribution of all landslides is essential for landslide susceptibility assessment, as it is the fundamental component of the assessments. (Soeters and Van Westen, 1996). In Turkey, a landslide inventory project at a national scale has been conducted by the Geological Research Department of the General Directorate of Mineral Research and Exploration, MTA. When preparing the landslide inventory, to identify landslides, vertical black and white aerial photographs at a medium scale of 1:35,000 from 1955 to 1956 were used (Nefeslioglu et al., 2010). A landslide database of the study area is included in this national inventory data, and 351 landslide locations were taken into account in the study region. Because the modeling approach employed in this study is independent of landslide data, the landslide inventory was only used for the validation process at the end of the study. In previous studies, different sampling methodologies have been applied in using landslide data for landslide susceptibility mapping (Nefeslioglu et al., 2008a; Yilmaz, 2010a). In these studies, it was suggested that accumulation/deletion zones or rupture zones of the sliding masses should not be considered as an assessment parameter in either the modeling or the validation stages. For this purpose, initially considering a spatial resolution of $25 \times 25$ m$^2$, buffer zones were drawn over 25 m around the landslide polygons. Then, only the main scarp portions of these buffered polygons, which are recognized as the landslide occurrence area, were distinguished for use in the validation stage of the predicted results.

To produce a landslide susceptibility map for the study area using the Mamdani FIS, seven landslide conditioning parameters, including altitude, lithology, slope gradient, curvature, a normalized difference vegetation index (NDVI), a stream power index (SPI), and a topographical wetness index (TWI), were used (Fig. 7), and all of these data were produced in raster format with a pixel size of $25 \times 25$ m$^2$ to be compatible with the spatial resolution.

The lithology map was first digitized using the 1:100,000 scale published map, and then the obtained vector data were converted into raster data. After obtaining the digital lithology data, the relationships between the lithological units and the occurrence of landslides were examined. On close inspection of Table 1, it can clearly be seen that many of the landslides in the study region (76.5%) occurred in three formations: the Sinop formation, the Atbası formation, and the Akveren formation. The other lithological units associated with landslide occurrence, the Kusuri formation (13.66%) and the Cankurtaran formation (4.25%), were also evaluated as having landslide occurrence potential.

Topographical parameters including altitude, slope gradient, curvature, SPI, and TWI were produced from the DEM of the study area (Fig. 6). Initially, the DEM was constructed by implementation of 1:25,000 scale topographical map contours using ArcGIS 9.3 software (ArcGIS Version 9.3, 2008). Altitude was obtained from the DEM as grid-type data. Altitude was evaluated as a good indicator for landslide susceptibility and has been used by many researchers (Pachauri and Pant, 1992; Ercanoglu and Gokceoglu, 2002; Nefeslioglu et al., 2010). For this reason, in this study, altitude was considered as a landslide-controlling parameter. According to the relationship between altitude and landslide occurrence, the mean altitude value on the grid cells with landslides was found to be 954.95 m ($\pm$ 190.39) (Table 2).

It is clear that the slope gradient is one of the most important topographical parameters controlling landslide occurrence, and for this reason, these data are taken into consideration in almost all landslide susceptibility, hazard, and risk assessment studies (i.e., Atkinson and Massari, 1998; Can et al., 2005; Nefeslioglu et al., 2008a, 2010; Yilmaz, 2010b; Akgun and Turk, 2010; Pradhan et al., 2010), and the slope gradient is also considered as a landslide-controlling parameter in the present study. Based on the descriptive statistics given in Table 2, it can be seen that landslides in the area typically occur on gentle slopes. Topography is the main factor controlling the spatial variation of slope stabilities. It affects the spatial distribution of soil moisture, and groundwater flow often follows surface topography (Rodhe and Seibert, 1999; Zinko et al., 2005; Yilmaz, 2010b). Topographic indices have, therefore, been used to describe spatial soil moisture patterns (Moore et al., 1991). One of these indices is the TWI developed by Beven and Kirkby (1979) within a runoff model. To calculate TWI values, Moore et al. (1991) proposed the equation

$$\text{TWI} = \ln(A_s/\tan \beta), \quad (7)$$

where $A_s$ is the specific catchment area (m$^2$/m) and $\beta$ is the slope gradient (in degrees). According to Wood et al. (1990), the variation in topographical components is often much greater than the local variability in soil transmissivity, and Eq. (7) can be used to calculate the TWI. By comparing the obtained TWI values with
the landslide occurrence (Table 2), it can clearly be seen that the mean TWI values were commonly calculated on the landslide bodies. This can be assumed to be a result of the availability of lithological units with a relatively high permeability and low surface runoff. The other secondary derivative of the DEM is the SPI. This index is used to describe potential flow erosion and

![Fig. 7. Landslide conditioning parameters used in this study (A) altitude; (B) slope gradient; (C) curvature; (D) normalized difference vegetation index (NDVI); (E) stream power index (SPI); and (F) topographical wetness index (TWI).](image)

**Table 1**
The distribution of the geological formations with respect to landslide in the study area.

<table>
<thead>
<tr>
<th>Formation</th>
<th>Symbol</th>
<th>Grid cells with landslides</th>
<th>All grid cells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>% (a)</td>
<td>Frequency</td>
</tr>
<tr>
<td>Alluvium</td>
<td>Qal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Beach deposits</td>
<td>Qpc</td>
<td>94</td>
<td>0.10</td>
</tr>
<tr>
<td>Kale formation</td>
<td>Qpk</td>
<td>246</td>
<td>0.27</td>
</tr>
<tr>
<td>Sankum formation</td>
<td>TplQs</td>
<td>3607</td>
<td>4.00</td>
</tr>
<tr>
<td>Sinop formation</td>
<td>Tms</td>
<td>27373</td>
<td>30.31</td>
</tr>
<tr>
<td>Kusuri formation</td>
<td>Tek</td>
<td>12339</td>
<td>13.66</td>
</tr>
<tr>
<td>Ayancik member</td>
<td>Teka</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Atbas formation</td>
<td>Tpea</td>
<td>24326</td>
<td>26.94</td>
</tr>
<tr>
<td>Karaada limestone</td>
<td>Ktpk</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Akveren formation</td>
<td>Ktpa</td>
<td>17381</td>
<td>19.25</td>
</tr>
<tr>
<td>Hamsaros formation</td>
<td>Kh</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Cankurtaran formation</td>
<td>Kc</td>
<td>3841</td>
<td>4.25</td>
</tr>
</tbody>
</table>
related landscape processes (Moore et al., 1991). The SPI is calculated from the formula

\[ SPI = A_s \times \tan \beta, \]

where \( A_s \) is the specific catchment area (\( m^2/m \)) and \( \beta \) is the slope gradient in degrees. As the specific catchment area and gradient increase, the amount of water contributed by upslope areas and the velocity of water flow also increase.

Nefeslioglu et al. (2010) emphasized that maximum SPI values are calculated in drainage channels, and for this reason, the values of the SPI for grid cells with landslides should not be expected to be high. According to the statistical results given in Table 2, this expectation is also reasonable for this study. This is because while the differences between the maximum and the minimum SPI values for the grid cells with landslides are high, the mean values for the grid cells with landslides are relatively low. Thus, it can be concluded that the SPI parameter constitutes a vital landslide-conditioning parameter in the study area.

The next type of DEM-derived topographical data used in this study is for curvature. The curvature represents the morphology of the topography (Wilson and Gallant, 2000). Whereas a positive curvature indicates that the surface is upwardly concave at that terrain surface, a negative curvature indicates that the surface is upwardly concave at the terrain surface. A value of zero defines a straight terrain surface. The minimum, maximum, and mean curvature values obtained for the grid cells with landslides were \(-0.62, 0.54, \) and \(-0.18\), respectively (Table 2). According to these results, it can be concluded that many landslides occurred at concave terrain surfaces in the study area. The curvature, SPI, and TWI data were produced using DIGEM 2.0 software, which was developed for digital terrain analysis by Conrad (2002), and then exported into ArcGIS 9.3 software to be used for model construction.

In this study, the NDVI was used as an environmental parameter. The NDVI is a measure of surface reflectance and gives a quantitative estimate of vegetation growth and biomass (Hall et al., 1995). Very low values of the NDVI (0.1 and below) correspond to barren areas, sand, or snow. Moderate values (0.2–0.3) represent shrub and grasslands, whereas high values (0.6–0.8) indicate temperate and tropical rainforests (Weier and Herring, 2005). The NDVI was calculated from the formula

\[ NDVI = \frac{(IR - R)}{(IR + R)}, \]

where IR is the infrared portion of the electromagnetic spectrum and R is the red portion of the electromagnetic spectrum. Based on the statistical values given in Table 2, the minimum and maximum NDVI values for grid cells without landslides are \(-0.72 \) and 0.77, respectively, with a mean of 0.26. The mean NDVI value for grid cells with landslides is calculated as 0.22, and this value is very close to the mean value obtained for grid cells with landslides. Hence, it can be concluded that the landslides in the study area occur where the vegetation cover density is in the moderate class, representing shrub and grassland fields in the region.

5. Landslide susceptibility mapping, and validation of the model

In this study, a Mamdani FIS is constructed. The Mamdani FIS for the assessment of landslide susceptibility for Sinop (Northern Turkey) includes a total of 7 inputs: altitude, lithology, slope, curvature, NDVI, SPI, and TWI (Fig. 7). Except for lithology, all of the inputs are constructed using two membership functions (Fig. 8). Lithology is formed by three crisp membership functions. A total of 12 different lithology types are cropped out in the study area. These lithologies are reclassified under three classes in the present study. When applying these classifications, the landslide density and the behavior of the lithology type in opposing to landsliding are considered. The Sinop formation, Kusuri formation, Atbası formation, and Akveren formation are assigned to the “high” membership function because the landslide densities in these formations vary between 13.66% and 30.31%. The second class, “moderate,” includes two formations, the Sarikum formation and the Cankurtaran formation. The landslide densities of these formations are 4.00% and 4.25%, respectively. The other formations are not associated with landslides or have a negligible landslide density (\(< 1.00\%\)). To minimize the uncertainty, a 50% overlap is applied between the fuzzy sets for each input parameter, and triangular membership functions are used for each fuzzy set. To provide a generalization for the study area, the minimum number of the fuzzy sets is considered. This also affects the number of the if–then rules. The output includes five fuzzy sets in the form of triangular membership functions (Fig. 9). One of the main parts of a Mamdani FIS is the fuzzy if–then rules. In this study, a total 192 if–then rules are used. The if–then rules are described only using expert opinion. When describing the if–then rules, the following considerations obtained from general knowledge of landslides and the field observations are used:

(a) Landslides in the study area occur in association with four lithologies (classified as “high”). For this reason, if the rule includes “high” in the lithology input, the output is accepted as “high” or “very high”.

(b) The other landslide conditioning factors are considered equally. If three inputs are positive and three inputs are negative for landslide occurrence, the output is assigned as “low”.

(c) If four inputs are positive and two inputs are negative for landslide occurrence, the output is accepted as “moderate”.

<table>
<thead>
<tr>
<th>Data</th>
<th>Variable</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Altitude (m)</td>
<td>0.82</td>
<td>954.94</td>
<td>226.22</td>
<td>190.39</td>
</tr>
<tr>
<td></td>
<td>Slope gradient (°)</td>
<td>0.00</td>
<td>62.74</td>
<td>12.21</td>
<td>7.35</td>
</tr>
<tr>
<td></td>
<td>Curvature</td>
<td>-0.62</td>
<td>0.54</td>
<td>-0.18</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>NDVI</td>
<td>-0.59</td>
<td>0.72</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>SPI</td>
<td>0.00</td>
<td>37056.96</td>
<td>310.47</td>
<td>859.90</td>
</tr>
<tr>
<td></td>
<td>TWI</td>
<td>2.72</td>
<td>15.06</td>
<td>6.30</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>Altitude (m)</td>
<td>0.00</td>
<td>1040</td>
<td>204.22</td>
<td>210.51</td>
</tr>
<tr>
<td></td>
<td>Slope gradient (°)</td>
<td>0.00</td>
<td>87.94</td>
<td>10.60</td>
<td>9.31</td>
</tr>
<tr>
<td></td>
<td>Curvature</td>
<td>0.91</td>
<td>1.02</td>
<td>15 x 10^{-6}</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>NDVI</td>
<td>-0.72</td>
<td>0.77</td>
<td>0.26</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>SPI</td>
<td>0</td>
<td>41551.45</td>
<td>177.13</td>
<td>643.43</td>
</tr>
<tr>
<td></td>
<td>TWI</td>
<td>2.20</td>
<td>17.35</td>
<td>6.43</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Table 2: General descriptive statistics of topographical variables with respect to landslides.
(d) If five inputs are positive and one input is negative for landslide occurrence, the output is accepted as “high”.
(e) If six inputs are positive for landslide occurrence, the output is accepted as “very high”.
(f) If the lithology is “low,” the output is accepted as “very low”.

After modeling the Mamdani FIS and producing the susceptibility values using MamLand, the susceptibility values were stored as a text file. Then, this file was exported into ArcGIS 9.3 software as point data containing output database information. After exporting the data into point-type data, they were converted into a raster file type so that they could be evaluated, and a landslide susceptibility map was obtained (Fig. 9).

For visual interpretation of this map, the necessity of classifying data into categorical susceptibility classes arose. Four data classifiers for this purpose are reported in the literature: equal interval, standard deviation, natural break, and quantile classification (Ayalew et al., 2005; Akgun et al., 2008). When considering all of these classifiers, the distribution of the data should be taken into account because class intervals change based on the chosen classifier (Ayalew et al., 2005). For example, if the data distribution is close to normal, equal interval or standard deviation classifiers should be used. If the data distribution has a positive or negative skewness, the quantile or natural break distribution classifiers could be chosen. In this study, before choosing the best data classifier for the obtained data, the data distribution histogram was taken into consideration (Fig. 10). Then, all of the classifiers noted above were applied to the data. The equal interval and standard deviation classifiers were found to be generally unsuccessful due to the fact that the data distribution skewed positively. Additionally, the natural break classifier classified the data into categorical susceptibility classes that were inferior to the quantile classifier. Therefore, the quantile data classification approach was chosen, and the landslide susceptibility index map was classified into five susceptibility classes: very low, low, moderate, high, and very high (Fig. 11).
The validation of predicted data is one of the most important parts of a probability-based map production process (Remondo et al., 2003). In the landslide susceptibility literature, although several methods are available for this, the area under the curve obtained from the ROC (receiver operating characteristics) plot is the most preferred and applicable type of statistical assessment for this purpose (Lee, 2005; Yilmaz, 2010a, b; Akgun and Turk, 2010; Nefeslioglu et al., 2010). ROC (Zweig and Champbell, 1993) plots should be drawn. These curves are obtained by plotting all combinations of sensitivities (on the y-axis) and proportions of false-negatives (1-specificity; on the x-axis) that may be obtained by varying the decision threshold (Brenning, 2005). In the AUC method, the AUC, with values ranging from 0.5 to 1.0, is used to assess the accuracy of the constructed model. The AUC defines the quality of the probabilistic model by describing its ability to reliably predict an occurrence or nonoccurrence event (Remondo et al., 2003; Nandi and Shakoor, 2009). An ideal model presents an AUC value close to 1.0, whereas a value close to 0.5 indicates inaccuracy in the model (Fawcett, 2006). The AUC value of the ROC curve for the Mamdani FIS model was found to be 0.855 (Fig. 12).

6. Conclusions

A great number of landslide susceptibility assessment papers have been published in landslide literature. Several different approaches for the assessment of landslide susceptibility have been employed. Almost all of these approaches are data driven. Although the Mamdani FIS is an expert-based system, it had previously been applied to a limited number of landslide susceptibility assessment studies (Lee and Juang, 1992; Juang et al., 1992; Saboya et al., 2006). However, it is a suitable system for landslide susceptibility assessment because it has a considerable capacity to model complex and nonlinear systems. The main reason for its lack of use has been the absence of a suitable computer program in which to implement it. It is possible to construct a Mamdani FIS using MATLAB. However, it is not possible to run a model including a very large geographic dataset using existing MATLAB software. To eliminate this problem, the MATLAB program MamLand was developed in this study. This program will assist landslide experts in assessing regional landslide susceptibility by directly employing expert opinion. MamLand is designed to be easy to use.

Sinop (Northern Turkey), which is a landslide-prone area, was considered as a case study area. The landslides identified in the study area are generally of an active deep-seated rotational type, and they usually occur on gentle slopes. Shoreline landslides are also observed in the area. Considering the main characteristics of the landslides and the field observations, a total of 7 inputs were selected for the model: altitude, lithology, slope gradient, curvature, NDVI, SPI, and TWI. Except for lithology, all of the inputs are
formed by two membership functions, while lithology includes three input membership functions. To minimize the uncertainty, a 50% overlap is applied between the membership functions. The output, the landslide susceptibility degree, is formed by five membership functions (Fig. 9). The FIS constructed for the study area has a total of 192 if–then rules. The if–then rules are described only using expert opinion. The AUC value of the ROC curve for the Mamdani FIS model constructed in the present study was found to be 0.855. Finally, we conclude that this methodology will present new opportunities for landslide researchers, as they can employ their knowledge in this model via membership functions and if–then rules.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2011.04.012.

References

Akgun, A., Dag, S., Bulut, F., 2008. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood frequency ratio and weighted linear combination models. Environmental Geology 54 (6), 1127–1143.


Yilmaz, I., 2010a. The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability (CP) and artificial neural network (ANN). Environmental Earth Sciences 60, 505–519.

