Computers & Geosciences 36 (2010) 1101–1114

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

A GIS-based neuro-fuzzy procedure for integrating knowledge and data in landslide susceptibility mapping

Mohammad H. Vahidnia*, Ali A. Alesheikh, Abbas Alimohammadi, Farhad Hosseinali

Faculty of Geodesy and Geomatics Eng. K.N. Toosi University of Technology, No. 1346, Mirdamad Cross, Valiasr Street, Tehran 19967-15433, Iran

ARTICLE INFO

Article history:
Received 15 April 2009
Received in revised form
19 April 2010
Accepted 29 April 2010

Keywords:
Geographic information system (GIS)
Landslide susceptibility map (LSM)
Fuzzy inference system (FIS)
Artificial neural network (ANN)
Back-propagation algorithm

ABSTRACT

A significant portion of the Mazandaran Province in Iran is prone to landslides due to climatic conditions, excessive rain, geology, and geomorphologic characteristics. These landslides cause damage to property and pose a threat to human lives. Numerous solutions have been proposed to assess landslide susceptibility over regions such as this one. This study proposes an indirect assessment strategy that shares in the advantages of quantitative and qualitative assessment methods. It employs a fuzzy inference system (FIS) to model expert knowledge, and an artificial neural network (ANN) to identify non-linear behavior and generalize historical data to the entire region. The results of the FIS are averaged with the intensity values of existing landslides, and then used as outputs to train the ANN. The input patterns include both physical landscape characteristics (criterion maps) and landslide inventory maps. The ANN is trained with a modified back-propagation algorithm. As part of this study, the strategy is implemented as a GIS extension using ArcGIS®. This tool was used to create a four-domain landslide susceptibility map of the Mazandaran province. The overall accuracy of the LSM is estimated at 90.5%.

1. Introduction

Landslides are significant geologic hazards and a natural source of devastation around the world. Iran, the subject of this study, is no exception. Entire landscapes can be dramatically reshaped by landslides, overwhelming more gradual erosion processes. The expansion of urban and man-made structures into potentially hazardous, landslide-prone areas leads to extensive property and infrastructure damage, and incurs some loss of life every year.

Since the early 1970's, many scientists have attempted to produce susceptibility maps portraying the spatial distribution of landslide hazards, often applying GIS-based techniques. Their efforts have demonstrated that such maps are effective tools for planners and decision makers (Nefeslioglu et al., 2008).

A hazard map, more accurately referred to as a landslide susceptibility map (LSM), aims at predicting where slope failures are most likely to occur. In this context, "susceptibility" is defined as the likelihood that a landslide will occur if temporal factors or triggers such as rainfalls and earthquakes are not considered (Dai et al., 2002). Landslide susceptibility assessment is the process of estimating this likelihood on the basis of physical terrain characteristics such as slope, landuse, and lithology. In addition, researchers often consider spatial correlations between important terrain characteristics and the past landslide distribution (Crozier and Glade, 2005; Soriso Valvo, 2002).

The most important performance attributes of a LSM are the reliability and accuracy of future probabilities (Soeters and van Westen, 1996). The degree of accuracy mainly depends upon the amount and quality of available data, the working scale of the map, and the methodologies of analysis and modeling (Yalcin, 2008; Ayalew and Yamagishi, 2005).

Several models and methods have been proposed to produce LSMs using geographic information systems (GIS). However, until recently there has been no general agreement on which class of methods is most suitable. LSM generation procedures can be classified as either direct or indirect (Carrara and Guzzetti, 1995), and as qualitative (Barredo et al., 2000) or quantitative (Agnesi et al., 2003).

Direct methods are based on a detailed geomorphological map and map the degree of hazard directly onto this field. Although direct methods have several advantages, they are time-consuming and depend heavily upon the expertise of the geomorphologist (Barredo et al., 2000).

All indirect methods, according to Clerici et al. (2002) and Süzen and Doyuran (2004), have certain steps in common:

1) Mapping past landslides in the target region.
2) Mapping a set of geological/geomorphological factors that are supposed to be directly or indirectly correlated with slope instability.

* Corresponding author. Tel.: +98 21 8878 6212; fax: +98 21 8878 6213.
E-mail addresses: vahidnia84@gmail.com (M.H. Vahidnia), alesheikh@kntu.ac.ir (A.A. Alesheikh), alimoh_abb@yahoo.com (A. Alimohammadi), frdhal@gmail.com (F. Hosseinali).

0098-3004/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2010.04.004

3) Estimating the correlations of these factors with slope instability.
4) Dividing the land area into domains of differing landslide susceptibility on the basis of these correlations (hazard zoning).

Quantitative approaches employ mathematical methodologies to estimate the susceptibility under rigorous constraints. Qualitative approaches, on the other hand, rely on the expert opinion of an individual or group (Neaupane and Piantanakulchai, 2006). Qualitative methods are subjective and define the hazard zones in explanatory terms. Quantitative methods estimate the probability of landslide occurrence numerically at every point in the region, so define hazard zones on a continuous scale (Guzzetti et al., 1999). An up-to-date landslide inventory map, with complete information on past mass movements, is required to accurately estimate these probabilities. The quantitative methods are also less subjective than any qualitative approaches and have been attempted more recently by academic and research institutions (Ermini et al., 2005).

The literature on landslide susceptibility assessment proposes numerous quantitative approaches. The bivariate statistical models by Süzen and Doyuran (2004) and Thiery et al. (2007) see widespread use. Other researchers prefer multivariate statistical techniques such as discriminant analysis (Carrara et al., 2003) or linear and logistic regression (Dai and Lee, 2002, 2003; Ayalew and Yamagishi, 2005; Yesilnacar and Topal, 2005; Greco et al., 2007) or non-linear methods such as artificial neural networks (Lee et al., 2004; Arora et al., 2004; Gomez and Kavzoglu, 2005; Ermini et al., 2005; Yesilnacar and Topal, 2005; Kanungo et al., 2006). Examples of qualitative or semi-quantitative approaches include the fuzzy logic theory of Zadeh (1965), employed by Saboya et al. (2006); the analytic hierarchy process (AHP) of Saaty (1980), employed by Yalcin (2008); the analytic network process (ANP) of Saaty (1999), employed by Neaupane and Piantanakulchai (2006); and the weighted linear combination (WLC) method by Ayalew et al. (2004).

The neuro-fuzzy model of Kanungo et al. (2005) is an indirect, completely quantitative technique for assessing landslide susceptibility. This method, while not directly related to the neuro-fuzzy system proposed in this paper, nonetheless produced remarkable results. It assigns membership degrees to landscape factors based on their contributions to landslide occurrence. The method employs an ANN to solve the regression and classification problem.

According to the literature, landslide susceptibility mapping typically suffers from two issues. If a qualitative technique is being used, then un-modeled phenomena or incomplete knowledge weakens the expert decisions. Quantitative methods suffer a similar impact from inaccurate or low-precision data. A method combining expert knowledge with objective data, or to put it differently mixing the quantitative and qualitative approaches, may lead to a more reliable outcome. Expert knowledge can make up for deficiencies in physical data, while quantitative analysis strips some of the subjectivity from individual opinions.

A fuzzy inference system (FIS) is a flexible and non-linear model that incorporates expert knowledge in the style of human thinking. It is appropriate for constructing a framework of easy inferences. On the other hand, the learning capabilities and outcomes of neural networks make them a natural match for fuzzy systems. ANNs can automate or support the process of developing a fuzzy system for a given task.

This paper proposes a new strategy for integrating expert knowledge and existing landslide data into a reliable LSM. Furthermore, we develop a GIS-based extension using the ArcGIS® software that implements the procedure. The first step integrates the outputs of a fuzzy inference system with actual landslide intensity values. (It should be noted that there are several possible definitions of intensity, based on physical parameters such as landslide dimensions, the volume of displaced material, and depth. This study uses the depth, normalized to the interval [0, 1].) Second, a neural network is trained to solve the regression problem between landslide susceptibility and selected physical landscape characteristics. Finally, the model is generalized and tested for the entire study area.

This paper does not attempt to modify known procedures for constructing a fuzzy rule base. Rather, it employs the power of neural networks to identify and generalize associations between physical inputs (landscape parameters) and the adjusted outputs of the fuzzy inference system.

1.1. Study area and materials used

The Mazandaran province is a mountainous and forested region located in the central part of the Alborz Mountains, which produce most of the landslides occurring in Iran. The study area is contained within the limits 52°31′ and 53°27′ E, 35°52′ and 36°30′ N and has a surface area of about 3440 km^2. The climate of the region is Mediterranean, with relatively high annual rainfall (about 1000 mm/yr). The study area lies near the Firouzkouh region, which has relatively high seismicity. In addition, an active branch of the Alborz Mountains Fault Zone (AMFZ) passes through the study area. Excessive rainfall, an extensive drainage network, and the generally low resistance of soil and rocks to pressure and variations of slope are widely recognized as the most important factors responsible for landslides and slope failures in this region (www.ngdir.ir).

For the purposes of the present study, the most important geomorphological units are the slope zones contributing to landslide occurrence. These zones are discontinuous, so we began by preparing an accurate landslide inventory map. A total of 151 active and dormant landslides were collected (Fig. 1), as well as instability-related parameters such as slope, aspect, curvature, landuse, lithology, distance from rivers, and distance from faults (Fig. 2).

Only translational landslides, the most common type, are used for this study. These landslides are well distributed over the study region. The slope, aspect, and curvature are derived from a digital elevation model (DEM). Slope is considered one of the most important factors, particularly when the angle is large. Aspect has an indirect influence on slope instability, through its relationship to the prevailing wind. Lithology (composition and structure) is related to landslide susceptibility because stronger rocks are more resistant to the driving forces. Landuse also plays an important role, since the incidence of landslides is inversely related to vegetation density (barren lands are more prone to landslides than thick forests). Proximity to faults contributes to slope instability and susceptibility, affecting not just surface structures but also terrain permeability. Finally, the erosion associated with rivers in hilly regions is a direct cause of many landslides. While other causal factors may be present in the data, these seven were selected for the strength of their correlation with landsides, their representation over the study area, their wide spatial variation, ease of measurement, and non-redundancy (Yalcin, 2008).

All preliminary thematic maps (the landslide inventory, a digital elevation model (DEM) with 1 m height precision and 80 m resolution, rivers, faults, lithology, and landuse) were obtained on a fine scale (1:1,000,000) from the National Geosciences Database of Iran (NGDIR) and the Forest Range & Watershed management organization (FRW).