



Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling

Wei Chen^a, Mahdi Panahi^b, Hamid Reza Pourghasemi^{c,*}

^a College of Geology & Environment, Xi'an University of Science and Technology, Xi'an 710054, China

^b Department of Geophysics, Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran, Iran

^c Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran

ARTICLE INFO

Keywords:

Landslide susceptibility
ANFIS
Genetic algorithm
Differential evolution
Particle swarm optimization
Hanyuan County

ABSTRACT

This paper presents GIS-based new ensemble data mining techniques that involve an adaptive neuro-fuzzy inference system (ANGIS) with genetic algorithm, differential evolution, and particle swarm optimization for landslide spatial modelling. This research was tested in Hanyuan County, which is a landslide-prone area in Sichuan Province, China. Different continuous and categorical landslide conditioning factors according to a literature review and data availability were selected, and their maps were digitized in a GIS environment. These layers are the slope angle, slope aspect, altitude, plan curvature, profile curvature, topographic wetness index, distance to faults, distance to rivers, distance to roads, lithology, normalized difference vegetation index and land use. According to historical information of individual landslide events, interpretation of the aerial photographs, and field surveys supported by the Sichuan Land Resources Bureau of China, 225 landslides were identified in the study area. The landslide locations were divided into two subsets, namely, training and validating (70/30), based on a random selection scheme. In this research, a probability certainty factor (PCF) model was used for the evaluation of the relationship between the landslides and conditioning factors. In the next step, three data mining techniques combined with the ANFIS model, including ANFIS-genetic algorithm (ANFIS-GA), ANFIS-differential evolution (ANFIS-DE), and ANFIS-particle swarm optimization (ANFIS-PSO), were used for the landslide spatial modelling and its zonation. Finally, the landslide susceptibility maps produced by the mentioned models were evaluated by the ROC curve. The results showed that the area under the curve (AUC) of all of the models was > 0.75. At the same time, the highest AUC value was for the ANFIS-DE model (0.844), followed by ANGIS-GA (0.821), and ANFIS-PSO (0.780). In general, the proposed ensemble data mining techniques can be applied for land use planning and management of landslide susceptibility and hazard in the study area and in other areas.

1. Introduction

Landslides are one of the most frequent and destructive natural hazards in the world and cause the loss of lives and property and damage to natural resources (Dagdelenler et al., 2016; Tsangaratos and Iliu, 2016). Therefore, to reduce these losses, it is valuable to assess the landslide susceptibility in a region (Moosavi and Niazi, 2016). Landslide susceptibility can be defined as the spatial probability of the landslide occurrences according to a set of geological and environmental conditions (Guzzetti et al., 2006). Landslide susceptibility maps can be helpful for decision makers and land use planning. In the past

two decades, with the development of GIS techniques, the configuration of more advanced qualitative and quantitative techniques has allowed the production of many studies that concern landslide susceptibility mapping (Van Westen et al., 2006). Several different methods and techniques for landslide susceptibility mapping have been proposed and tested in the world. However, no general agreement exists either on the methods or on the scope of producing the landslide susceptibility maps (Guzzetti et al., 1999; Pourghasemi et al., 2013a). Probabilistic models are more frequently used, and a large number of methodologies have been developed. The frequency ratio model, which is a simple and effective method, has been widely used for landslide susceptibility

* Corresponding author.

E-mail address: hr.pourghasemi@shirazu.ac.ir (H.R. Pourghasemi).

<http://dx.doi.org/10.1016/j.catena.2017.05.034>

Received 8 September 2016; Received in revised form 23 February 2017; Accepted 26 May 2017

Available online 01 June 2017

0341-8162/ © 2017 Elsevier B.V. All rights reserved.

mapping (Akgun, 2012; Chen et al., 2016b; Demir et al., 2013; Jaafari et al., 2014). The statistical index model, which is one of the bivariate models, was also used by some researchers (Chen et al., 2016a; Nasiri Aghdam et al., 2016; Tien Bui et al., 2011). Additionally, several other types of bivariate models, such as Dempster-Shafer (Chen et al., 2017a; Mohammady et al., 2012), certainty factor (Chen et al., 2016b; Devkota et al., 2013; Dou et al., 2014), weights of evidence (Chen et al., 2016b; Ozdemir and Altural, 2013; Regmi et al., 2014), and entropy (Constantin et al., 2011; Jaafari et al., 2014; Youssef et al., 2015a), have also been widely used for landslide spatial modelling.

In addition, some multivariate statistical approaches, such as logistic regression (Conoscenti et al., 2015; Erenner et al., 2016; Pourghasemi et al., 2013a), multiple adaptive regression splines (Conoscenti et al., 2015; Felicísimo et al., 2013; Wang et al., 2015), and generalized additive models (Brenning, 2008; Goetz et al., 2011; Park and Chi, 2008), were widely used and used by many researchers. The analytical hierarchy process (Chen et al., 2016b; Pourghasemi et al., 2013a; Youssef, 2015) and their combinations, such as multi-criteria decision analysis (MCDA), spatial multi-criteria evaluation (SMCE), and multi-criteria evaluation (MCE) (Erenner et al., 2016; Gorsevski and Jankowski, 2010; Pourghasemi et al., 2014), which are expert knowledge-based methods, have also been used in landslide susceptibility analysis.

Apart from the techniques mentioned above, some new soft computing techniques, such as artificial neural networks (Polykretis et al., 2015; Tien Bui et al., 2016c), fuzzy logic (Pourghasemi et al., 2012; Shahabi et al., 2015; Vakhshoori and Zare, 2016), support vector machines (Colkesen et al., 2016; Xu et al., 2012), adaptive neuro-fuzzy inference systems (ANFIS) (Dehnavi et al., 2015; Nasiri Aghdam et al., 2016; Tien Bui et al., 2012), kernel logistic regression (Hong et al., 2015; Tien Bui et al., 2016c), decision trees (Pradhan, 2013; Tsangaratos and Ilia, 2016; Wang et al., 2016), alternating decision trees (Hong et al., 2015; Pham et al., 2016), and random forest (Chen et al., 2017b; Trigila et al., 2015; Youssef et al., 2016) models have been used to evaluate landslide susceptibility because of their lower statistical limitations. However, based on an in-depth literature review, no research has been conducted that explores ensembles of ANFIS with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide susceptibility modelling.

Thus, the main objective of this research is to evaluate the effect of GIS-based ensembles of ANFIS with GA, DE, and PSO models for landslide susceptibility mapping in a landslide-prone area (Hanyuan County, China) and to evaluate and compare their performances. The main difference between the proposed methodology and the previously mentioned studies is that the ensembles of ANFIS with the GA, DE, and PSO models are for the first time implemented for landslide spatial modelling.

2. Materials and methods

In the present study, the methodology could be separated into five phases: (i) selection of a study area; (ii) data preparation, including construction of the landslide inventory map and conditioning factors; (iii) data correlation analysis using PCF (probability certainty factor) model; (iv) landslide spatial modelling using ANFIS-GA, ANFIS-DE, and ANFIS-PSO; and (v) validation and selection of the optimal models (Fig. 1).

2.1. Description of the study area

The Hanyuan County covers an area of approximately 2388 km², which is located in Sichuan Province, China. The region lies at a longitude of 102°16′ to 103°00′E and a latitude of 29°05′ to 29°43′N (Fig. 2). The minimum and maximum altitudes are 632 m and 3940 m above sea level, respectively. The climate of the study area is known as a monsoonal subtropical region with an annual average rainfall of

741.8 mm/year and an annual average temperature of 17.9 °C (<http://www.sxmb.gov.cn>). The average relative humidity is 68%, and the maximum wind speed is 15.3 m/s. There are five main types of land use, namely, farmland, residential areas, water bodies, bare land, and forest/grass. The slope cropland, urban and rural construction, road construction, hydropower projects, and mine exploitation are the main anthropogenic activity types in the study area. The Hanyuan County is located in the northern section of the Hengduan Mountains and the transition zone between the Western Sichuan Plateau and Sichuan Basin (Lu Tao et al., 2009). The study area is surrounded by mountains, and the terrain tilts from northwest to southeast. According to the characteristics and genetic types, the geomorphology can be divided into two types: valley erosion geomorphology and tectonic erosion geomorphology. The lithological units of the study area comprise several formations from the Proterozoic period to Quaternary sediments, except for those of Devonian and Cretaceous (Fig. 3). Therefore, the geology of this region is very complex. Mudstone, sandstone, shale, siltstone, and limestone are the most out-cropped lithological formations.

2.2. Data preparation

2.2.1. Landslide inventory map

In the present study, the landslide inventory map was prepared using a collection of historical information on individual landslide events, interpretation of the aerial photographs, and extensive field surveys supported by the Sichuan Land Resources Bureau of China. Finally, a total of 225 landslides were mapped (Fig. 2). Detailed information on landslides with a time span of > 40 years (since 1973) has been used. The smallest and largest sizes of the landslides identified are approximately 2×10^2 and 3×10^6 m², respectively. For landslide spatial modelling by the mentioned ensemble models, the landslide locations are divided into two subsets, namely, training and validating (70/30), based on a random selection scheme (Hong et al., 2015).

2.2.2. Dataset preparation for spatial modelling

In addition to the landslide inventory map, there is a variety of inter-related factors that affect the landslides. Twelve conditioning factors, namely, the slope angle, slope aspect, altitude, plan curvature, profile curvature, TWI (Topographic Wetness Index), distance to faults, distance to rivers, distance to roads, lithology, NDVI (Normalized Difference Vegetation Index), and land use are considered for the landslide spatial modelling and their mapping. These factors were selected according to the data availability and the literature review (Akgun, 2012; Nasiri Aghdam et al., 2016; Pourghasemi et al., 2013a). In this study, the ESRI ArcGIS 10.2 software was used to produce and display these data layers. All of the data layers were prepared in raster format with a spatial resolution of 30 m or a pixel size of 30*30 m². In the present study, the slope angle, slope aspect, altitude, plan curvature, profile curvature, and TWI factors were extracted from ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) Global DEM (Digital Elevation Model), with a 30 m * 30 m resolution. The distance to the faults was extracted from the geological map and was determined using the Euclidean distance method. The distances to the roads and the distances to the rivers were prepared from a topographic map at a 1:50,000 scale using the Euclidean distance method. The lithology map was also extracted from the geological map and converted into raster format. NDVI and land use maps were derived from Landsat 8 Operational Land Imager (OLI) images. From these, the NDVI values are defined as

$$NDVI = \frac{IR - R}{IR + R} \quad (1)$$

where *IR* is the infrared portion of the electromagnetic spectrum and *R* is the red portion of the electromagnetic spectrum (Pourghasemi et al.,

Download English Version:

<https://daneshyari.com/en/article/5769868>

Download Persian Version:

<https://daneshyari.com/article/5769868>

[Daneshyari.com](https://daneshyari.com)