Catena 135 (2015) 271–282

Contents lists available at ScienceDirect

Catena

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

Landslide susceptibility mapping in Mizunami City, Japan: A comparison between logistic regression, bivariate statistical analysis and multivariate adaptive regression spline models

Liang-Jie Wang [a,b,*], Min Guo [b], Kazuhide Sawada [b], Jie Lin [a,c], Jinchi Zhang [a,c]

[a] Collaborative Innovation Center of Sustainable Forestry in Southern China of Jiangsu Province, Nanjing Forestry University, Nanjing 210037, China
[b] Graduate School of Engineering, Gifu University, Gifu 5011193, Japan
[c] Jiangsu Provincial Key Lab of Soil Erosion and Ecological Restoration, Nanjing Forestry University, Nanjing 210037, China

ARTICLE INFO

Article history:
Received 29 April 2014
Received in revised form 31 July 2015
Accepted 13 August 2015
Available online 2 September 2015

Keywords:
Landslide susceptibility
Logistic regression (LR)
Bivariate statistical analysis (BS)
MARSpline
Mizunami City

ABSTRACT

Landslides are dangerous natural hazards. Because of their threat, a comprehensive landslide susceptibility map should be produced to reduce the possible damages to people and infrastructure. The quality of landslide susceptibility maps is influenced by many factors, such as the quality of input data and the selection of mathematical models. This study aimed to identify the optimal quantitative method for landslide susceptibility mapping in Mizunami City, Japan. Three mathematical methods, logistic regression (LR), bivariate statistical analysis (BS), and multivariate adaptive regression spline models (MARSplines), were used to create landslide-susceptibility maps by comparing the past landslide distribution and the conditioning factor thematic maps. A landslide inventory map with a total of 222 landslide locations was extracted from aerial photographs provided by NIED (National Research Institute for Earth Science and Disaster Prevention, Japan). Then, the landslide inventory was randomly divided into two datasets: 50% was used for training the models and the remaining 50% for validation purposes. The landslide inventory map provided by NIED and an area under the ROC curve were used to evaluate model performance. We found that the MARSpline method resulted in a better prediction rate (79%) when compared to LR (75%) and BS (77%). In addition, a higher percentage of landslide polygons were found in the high to very high classes using the MARSpline method. Therefore, we concluded that the MARSpline method was the most efficient method for landslide susceptibility mapping in this study area.

1. Introduction

In recent decades, landslides have attracted considerable attention because they are the most common disaster in the world in terms of human casualties and damage to social economies (Nefeslioglu et al., 2008; Shahabi et al., 2014). According to the formal definition, a landslide is the movement of a mass of rock, debris, or earth down a slope under the influence of gravity (Guzzetti, 2005; Varnes, 1978). To mitigate damage from landslides, the government must keep its citizens informed. However, the necessary field observations to produce this information may be time-consuming and costly, especially for larger areas. A tool for mapping and tracking landslides could help local governments to mitigate the associated economic losses. Many geomorphologists and engineering geologists have therefore proposed different techniques to evaluate these hazards, including landslide-susceptibility zoning.

Landslide susceptibility is the likelihood of a landslide occurring in an area given the local geo-environment (Brabb, 1984; Guzzetti, 2005; Guzzetti et al., 2006). The three main methodologies for assessing landslide susceptibility are qualitative, deterministic and statistical methods. In the qualitative methods, experts use their own knowledge to assign weights to conditioning factors to develop a susceptibility map (Regmi et al., 2010). Deterministic methods consider the slope angle, slope material strength, structure of rock discontinuities, rock and soil stratification, moisture content of the slope material, and depth of the groundwater table in a physics-based equation to determine an index of stability of a slope, e.g., the factor of safety (Regmi et al., 2010). Due to the rapid development of the geographical information system (GIS) technology, various quantitative or statistical methods have been proposed to assess landslide susceptibility, including the logistic regression (LR) model (Lee, 2007a; Tunusluoglu et al., 2008; Nandi and Shakoor, 2010; Das et al., 2010; Pradhan, 2010; Wang et al., 2013; Althuwaynee et al., 2014; Shahabi et al., 2014), fuzzy logic method (Ercanoglu and Gokceoglu, 2002; Kanungo et al., 2008; Lee, 2007b; Muthu et al., 2008; Pradhan, 2010), artificial neural network method (Bui et al., 2012; Chen et al., 2009; Conforti et al., 2014; Melchiorre et al., 2008; Poudyal et al., 2010; Pradhan and Lee, 2009, 2010a, 2010b, 2010c; Yilmaz, 2010), Bayes

* Corresponding author at: Collaborative Innovation Center of Sustainable Forestry in Southern China of Jiangsu Province, Nanjing Forestry University, Nanjing 210037, China.
E-mail address: ljwang0811@gmail.com (L.-J. Wang).

http://dx.doi.org/10.1016/j.catena.2015.08.007
0341-8162/

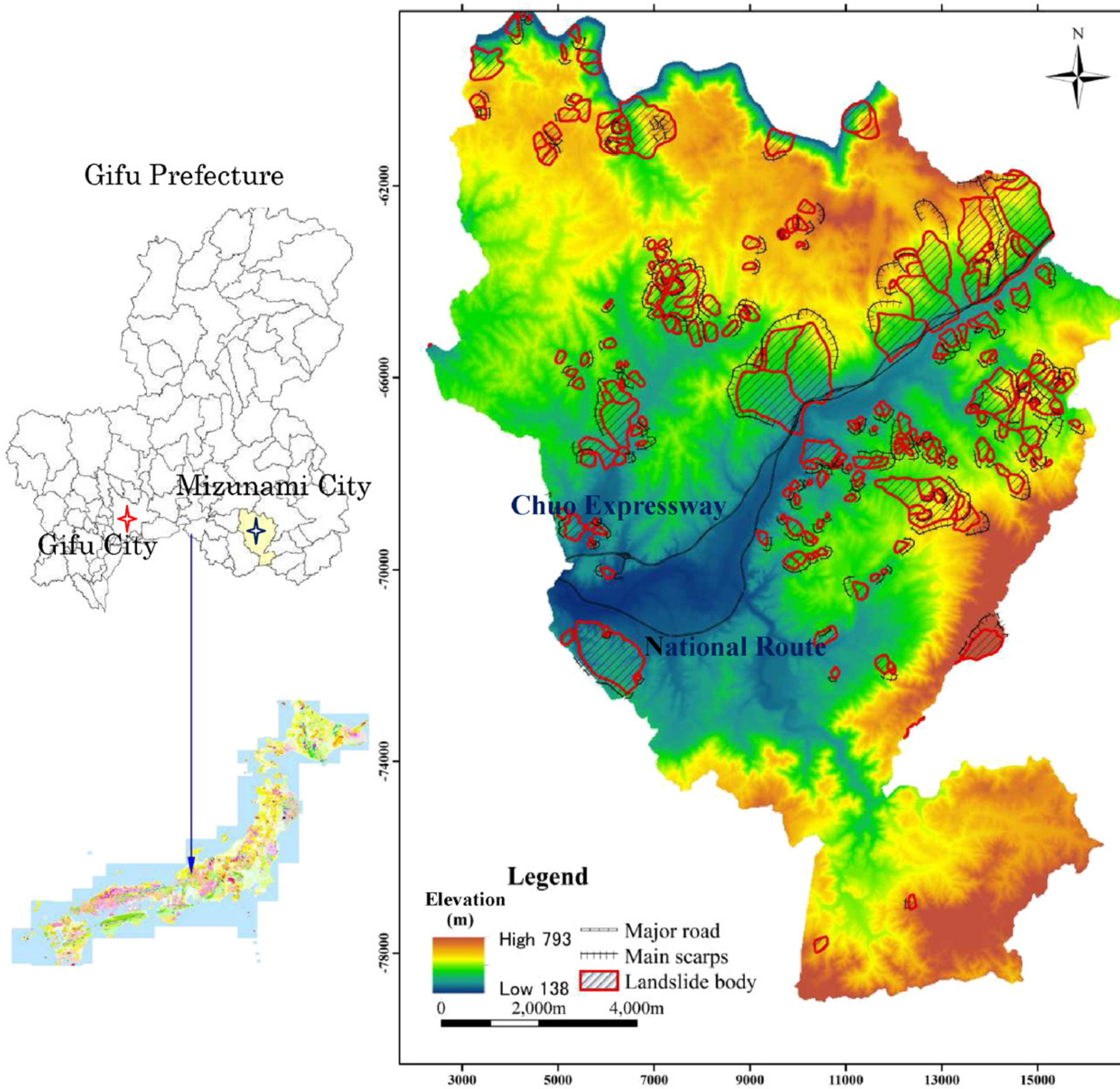

Fig. 1. Geographical location and distribution of landslide bodies in Mizunami City, Gifu Prefecture, Japan.

theorem based on weights of evidence (WOE) (Regmi et al., 2010), neural–fuzzy method (Oh and Pradhan, 2011; Vahidnia et al., 2010), support vector machines (SVMs) (Xu et al., 2012; Yao et al., 2008; Yilmaz, 2010), decision tree method (Pradhan, 2013; Saito et al., 2009; Yeon et al., 2010), bivariate statistical (BS) analysis (Nandi and Shakoor, 2010; Xu et al., 2012) and multivariate adaptive regression spline (MARSpline) model (Felicisimo et al., 2013).

Recent studies have compared different quantitative and statistical methods for predicting landslide susceptibility. Shahabi et al. (2014) compared logistic regression with frequency ratio and analytical hierarchy process, whereas Demir et al. (2014) compared logistic regression with the frequency ratio. Xu et al. (2012) compared logistic regression with bivariate statistics, artificial neural networks and support vector machines using three different kernel functions and found logistic regression to be the most efficient. Kavzoglu et al. (2014) compared multi-criteria decision analyses and support vector regression with logistic regression and found that these approaches outperformed the conventional logistic regression method in the mapping of shallow landslides. Nandi and Shakoor (2010) compared logistic regression with a bivariate statistics approach and found the logistic regression method to be the most accurate of these techniques.

Although LR and BS have been widely applied for landslide susceptibility analyses in different areas, the MARSpline method has only rarely been used to assess landslide susceptibility and was never compared to

Table 1
Lithologic components of the study area and their frequency ratio values.

Lithologic type	Description	Frequency ratio
J2-3ac	Triassic to Middle Jurassic chert block of Middle to Late Jurassic accretionary complex	2.322
J2-3ax	Melange matrix of Middle to Late Jurassic accretionary complex	0.0
K2gp	Late Cretaceous felsic plutonic rocks (Younger Ryoke Granite)	0.066
K2gr	Late Cretaceous granite (Younger Ryoke Granite)	0.890
K2gd	Late Cretaceous granodiorite (Younger Ryoke Granite)	0.0
M8tux	Ryoke metamorphic rocks (gneiss and schist)	0.0
J2-3as	Sandstone of Middle to Late Jurassic accretionary complex	1.066
Sn/Sr	Marine and non-marine sedimentary rock	0.957
K2vf/vi	Late Cretaceous non-alkaline felsic volcanic rocks/volcanic intrusive rocks	2.332

The definition of the frequency ratio is shown by formula 1.