Catena 165 (2018) 520–529

Contents lists available at ScienceDirect

Catena

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

Review on landslide susceptibility mapping using support vector machines

Yu Huang[a,b,*], Lu Zhao[a]

[a] Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China
[b] Key Laboratory of Geotechnical and Underground Engineering of the Ministry of Education, Tongji University, Shanghai 200092, China

ARTICLE INFO

Keywords:
Landslide susceptibility assessment
Landslide susceptibility mapping
Support vector machine
Integrated tool

ABSTRACT

Landslides are natural phenomena that can cause great loss of life and damage to property. A landslide susceptibility map is a useful tool to help with land management in landslide-prone areas. A support vector machine (SVM) is a machine learning algorithm that uses a small number of samples for prediction and has been widely used in recent years. This paper presents a review of landslide susceptibility mapping using SVM. It presents the basic concept of SVM and its application in landslide susceptibility assessment and mapping. Then it compares the SVM method with four other methods (analytic hierarchy process, logistic regression, artificial neural networks and random forests) used in landslide susceptibility mapping. The application of SVM in landslide susceptibility assessment and mapping is discussed and suggestions for future research are presented. Compared with some of the methods commonly used in landslide susceptibility assessment and mapping, SVM has its strengths and weaknesses owing to its unique theoretical basis. The combination of SVM and other techniques may yield better performance in landslide susceptibility assessment and mapping. A high-quality informative database is essential and classification of landslide types prior to landslide susceptibility assessment is important to help improve model performance.

1. Introduction

China contains a large number of plateaus and mountainous areas which tend to have complex geological conditions and poor climatic conditions. This can lead to hazardous environments, resulting in frequent landslides. According to data from the Ministry of Land and Resources P.R.C, landslides are the most common geological disaster in China (Fig. 1). Landslides can cause serious casualties and severe damage to property. The death toll caused by landslides worldwide is large (Flentje and Chowdhury, 2016). Petley (2012) showed that 32,322 lives lost had been recorded due to 2620 fatal non-seismic landslides over the period of 2004–2010. Landslides triggered by large-magnitude earthquakes can lead to enormous death toll (Flentje and Chowdhury, 2016). Nearly 60,000 individual landslides were triggered by the Wenchuan earthquake in Sichuan Province, China, on 12 May 2008 (Gorum et al., 2011), resulting in 20,000 deaths (Chigira et al., 2010). Four examples of these landslides in Sichuan province taken after the Wenchuan earthquake are shown in Fig. 2. Landslides were triggered by the northern Pakistan earthquake on 8 October 2005, leading to 26,500 deaths (Petley et al., 2006). To evaluate the risk of regional landslides under complex terrain conditions, the spatio-temporal motion and deformation of susceptible areas can be monitored using remote sensing techniques. Landslide susceptibility prediction shows the spatial probability of landslide occurrence and is a key technology in landslide monitoring, early warning and assessment (Wu et al., 2016).

Methods for assessing landslide susceptibility can be divided into three basic types: knowledge-based methods, physical methods and data-based methods (Corominas et al., 2014). The details of these three methods (Li et al., 2017) are listed in Table 1. There are both advantages and drawbacks to each method (Li et al., 2017) and there is no consensus about which method is the most effective. Generally, data-based methods are more suitable for regional landslide susceptibility assessment (Corominas et al., 2014).

In recent years, satellite-based earth observation technology and data-based hazard analysis methods have been increasingly used by scientists and disaster responders for global rapid assessment of disaster situations (Voigt et al., 2016). Assuming that a landslide is more likely to occur under conditions similar to those that caused past landslides (Tien Bui et al., 2012), data-based methods use a scientific model to predict the probability of landslide occurrence based on the spatial distribution of a series of landslide-influencing factors in the landslide-prone areas (Wu et al., 2016). A support vector machine (SVM), which is characterized by a small number of samples, nonlinearity and high dimension, is a machine learning algorithm which has become popular with the development of artificial intelligence and the widespread use

* Corresponding author at: Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China.
E-mail address: yhuang@tongji.edu.cn (Y. Huang).

https://doi.org/10.1016/j.catena.2018.03.003
Received 30 January 2018; Received in revised form 2 March 2018; Accepted 4 March 2018

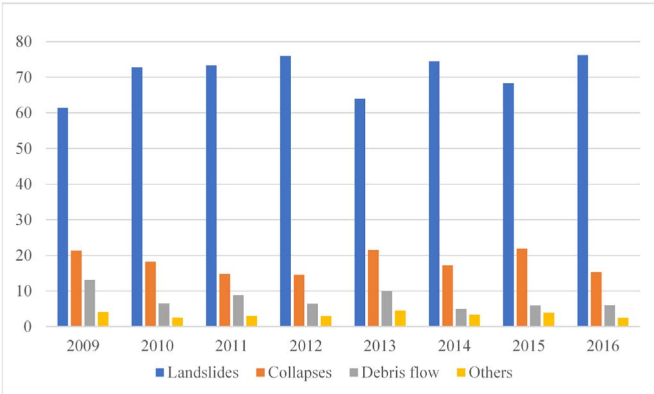

Fig. 1. Geological hazards in China from 2009 to 2016 (in percentage).

of geographic information system (GIS) and remote sensing (RS). In this paper, we discuss SVM-based landslide susceptibility assessment and mapping, its advantages and disadvantages compared with other assessment methods.

Most SVM-based studies of landslide susceptibility mapping focus on a specific procedure and compare maps generated by different methods, generating statistical data to evaluate the analysis methods without any insight into the essence of the methods (i.e., the algorithm, the complexity, ease of use, and suitability for larger datasets) (Pham et al., 2016a; Tien Bui et al., 2012; Xu et al., 2012a; Xu et al., 2012b). This paper presents an overview of landslide susceptibility mapping using SVM and compares different methods from the viewpoint of the algorithm used by each method. First, the application of SVM in assessing and mapping landslide susceptibility is introduced. Next, we compare SVM with four other landslide susceptibility assessment methods. Finally, based on its merits and demerits, some future work on SVM in landslide susceptibility mapping is proposed and conclusions drawn.

2. Application of SVM in landslide susceptibility mapping

2.1. Basic SVM theory

Based on statistical learning theory, SVM follows the principle of structural risk minimization and is used as a machine learning algorithm (Vapnik, 1995; Vapnik, 1999). The two main principles of SVM are: the optimal classification hyperplane and the use of a kernel function (Yao et al., 2008) (Fig. 3). In Fig. 3(a), the dots and squares represent two types of samples, H is the classification line between them, and H1 and H2 are lines parallel to H running through the sample points closest to the classification line (the support vectors). The distance between them is called the classification margin. The goal of the optimal classification hyperplane is to discriminate between the two types of samples correctly (though certain errors are allowed) while maximizing the classification margin. Fig. 3(b) illustrates the kernel function which helps to transform the input samples into a higher-dimensional space so that they can be classified linearly. Detailed

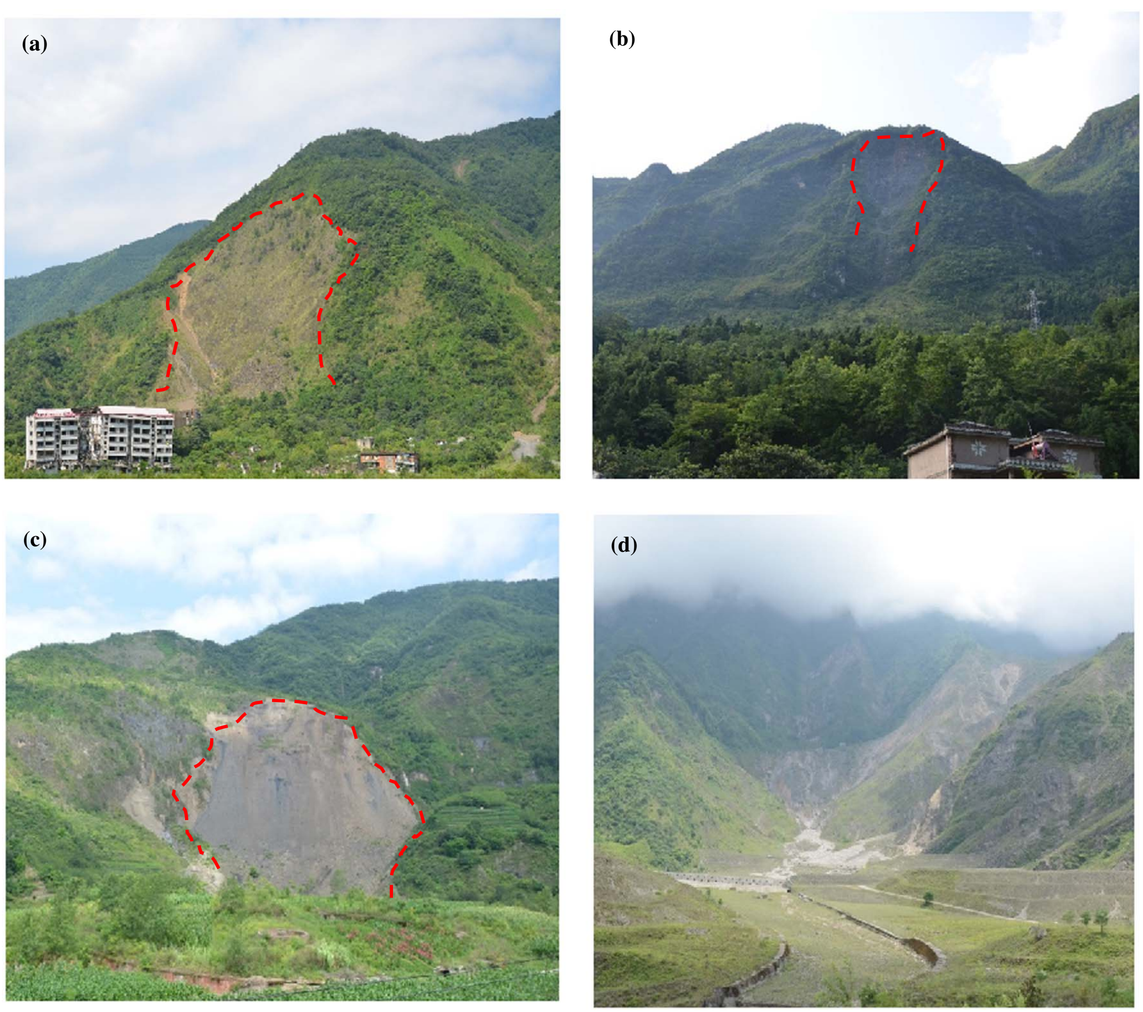

Fig. 2. Photos of four landslide sites in Sichuan taken after the 2008 Wenchuan earthquake.