Computers & Geosciences 56 (2013) 23–31

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

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

Integration of dynamic rainfall data with environmental factors to forecast debris flow using an improved GMDH model

Hui Zhang [a,b], Xiangnan Liu [a,*], Erli Cai [a], Gang Huang [c], Chao Ding [a]

[a] School of Information Engineering, China University of Geosciences, Beijing 100083, China
[b] School of Computer and Information Engineering, Tianjin Institute of Urban Construction, Tianjin 300384, China
[c] Information center, Unit 91917, Beijing 100841, China

ARTICLE INFO

Article history:
Received 6 December 2012
Received in revised form
30 January 2013
Accepted 5 February 2013
Available online 17 February 2013

Keywords:
GMDH model
Dynamic rainfall data
Environmental factors
Simulated annealing algorithm
Genetic algorithm

ABSTRACT

The objective of this study was to apply an improved Group Method of Data Handling (GMDH) network model for prediction of debris flow by integrating dynamic rainfall data and environmental factors. The rainfall data were collected from weather information, and the environmental data were extracted from RS, GIS, drilling data, and geophysical data. The input variables used in the SAGA-GMDH model were derived from six variables acquired by Kernel Linear Discriminant Analysis (KLDA). The results showed that the GMDH for prediction of debris flow performed well using the training, validation, and testing sets (R^2 above 0.80 and ARE below 3.54%). The SAGA-GMDH model was subsequently compared with a back-propagation (BP) neural network model and adaptive network fuzzy interference system (ANFIS). The accuracies of the SAGA-GMDH model prediction were slightly better than those of other two models, which demonstrated that the SAGA-GMDH model was more suitable for prediction of debris flow.

1. Introduction

Rainfall-triggered debris flow is a serious type of natural hazard that occurs in certain places in the world. In the past few decades, rainfall-triggered debris flow has caused serious loss of human lives and extensive damage to property (Larsen and Simon, 1993; Crosta, 2004; DeGraff et al., 2011). A number of methods for forecasting of rainfall-triggered debris flow have been demonstrated (Tsaparas et al. 2002; Santoso et al. 2011). The general approaches can be classified into two types: the field-based qualitative approach or the data-driven quantitative approach (Nandi and Shakoor, 2010; Xu et al.,2012; Kung et al., 2012). Traditionally, the assessment of a debris flow disaster has been carried out using physical model experiments (Cui, 1991) and long-term outdoor observation. With the development of RS and GIS, the environmental data are easy to obtain. Because remote sensing has many advantages over physical model experiments or long-term outdoor observations, i.e., lower cost, faster data acquisition, better spatial and temporal continuity, great progress has been facilitated by remote sensing (Ou et al., 2006; Vahidnia et al., 2010; Oh and Pradhan, 2011). However, the existing work on prediction of debris flow only considers such rainfall factors as accumulated rainfall capacity, rainfall intensity, and rainfall capacity. In addition to the precipitation, the geological environmental factors are also important conditions in the process of debris flow forecasting. If the influence of the geological environmental factors is included in the debris flow prediction model, the forecast result will result in higher objectivity and credibility. Through a combination of geological environmental factors and rainfall factors, an accurate and valid forecast for debris flow in single gully can be achieved. Therefore, it is necessary to develop a significant method that couples the environmental factors with rainfall factors to predict the debris flow in a single gully.

Due to the uncertainties and complexities of the factors involved in the causes of debris flow, it is generally difficult to quantitatively analyze these influences and to predict the occurrence of debris flow. System identification techniques are applied in many fields in order to model and predict the behaviors of unknown and/or very complex systems based on given input–output data. Soft computing methods, which concern computation in an imprecise environment, have gained significant attention. Among these methodologies, the Group Method of Data Handling (GMDH) is an effective method for establishing a mathematical model of a complex system using a heuristic self-organization approach (Ivakhnenko, 1976). One characteristic of the GMDH is that it provides an automatic modeling mechanism (Zhu et al., 2012). This automatic modeling mechanism has been

* Corresponding author at: China University of Geosciences, School of Information Engineering, Beijing's haidian district xueyua, Beijing 100083, China.
Tel.: +86 1082321796; fax: +86 1082322095.
E-mail addresses: zh1809@sina.com, hust-zhanghui@163.com (H. Zhang), liuxncugb@163.com (X. Liu).

0098-3004/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2013.02.003

successfully applied to build Bayesian networks (Xiao et al., 2009) and Mamdani-type fuzzy models (Mueller and Lemke, 2000), among others. Another desirable characteristic of the GMDH is its immunity to noise. It is well known that when the data contain noise, the most dangerous effect is over-fitting (Schittenkopf et al., 1997), which produces that models tend to be excessively complex and characterized by poor generalization. A detailed discussion of noise immunity in the GMDH can be found in (Madala and Ivakhnenko, 1994). The objective of this research is to apply the GMDH model to predict the debris flow in a single gully using a combination of environmental factors and rainfall factors. This is to say, a kind of system identification technique is applied to predict debris flow.

2. Materials and methods

2.1. Study areas

This research focuses on the areas of Xiu Yan county, the town of An Shan, and the Liao Ning province in China (Fig. 1) due to the frequent occurrence of debris flow in this region, which is approximately located between latitudes 40°00′N to 40°39′N and longitudes 122°52′E to 123°46′E. The Xiu Yan county covers an area of 4502 km^2, and the temperature of the study area ranges from −30.9° to 37.3°. The terrain consists of hills, low mountains, and middle-sized mountains.

The average annual precipitation is 896 mm, with a minimum precipitation of 172.2 mm and a maximum of 1451.3 mm. The rainy season occurs primarily from June to September, with precipitation in this period accounting for 73.5–80.2% of the yearly amount. The precipitation distribution decreases from the southeast to the northwest.

The debris flows in the Xiu Yan country represent natural degradation processes that are largely triggered by heavy rain due to either a single heavy thunderstorm/rain or successive days of moderate rain (especially during the rainy season), which causes flash flooding that leads to failure of the rock surfaces along fractures, joints, and cleavage planes (Jin, 2011).

2.2. Data collection

The formation of debris flow is a complicated process that requires favorable terrain conditions, source conditions and hydrodynamic conditions. Precipitation is the most important triggering factor for inducement of a debris flow disaster. The debris flow is the joint result of antecedent rainfall and short-term intense rainfall (Tang et al., 1994; Chen, 1985). Several factors were selected as the environmental factors for debris flow: the drainage area (DA), the basin relative height difference (BRHD), supply length ratio (SLR), brook longitudinal slope (BLS), vegetation coverage (VC), ditch shore hillside slope (DSHS), and loose materials reserves along the ditch (LMRD). The 5-day accumulative rainfall (FAR), the maximum hours of rainfall intensity (MHRI), and the daily rainfall (DR) were selected as components of the rainfall influence index of the debris flow warning model. The rainfall intensity is an important initiation factor for debris flow, and the daily rainfall data can be provided by the meteorological observatory 24-h rainfall forecast; the greater the rainfall, the greater the likelihood of debris flow occurrence. The long-term continuous rainfall significantly increases the void water pressure in deep-slope soil, and the soil produces sliding fracture behavior, which is an important contribution to the occurrence of debris flow. Therefore, the accumulative rainfall amount responds to the saturated liquefaction degree of the soil, to a certain extent, according to the degree of influence of the antecedent precipitation. When debris flow occurs in the Xiu Yan area, the 5-day accumulative rainfall acts as the effect index of the antecedent precipitation.

The largest quantity of debris flow that mobilizes at one time in a single gully (BQDR) directly determines the destructiveness of the debris flow, and thus this quantity was viewed as the dominant assessment index of the debris flow risk. The abbreviations and units of the predictors of debris flow are shown in Table 1.

2.3. Methodology

The most important tasks in this method are the establishment of a data retrieval model and selection of the input variables for

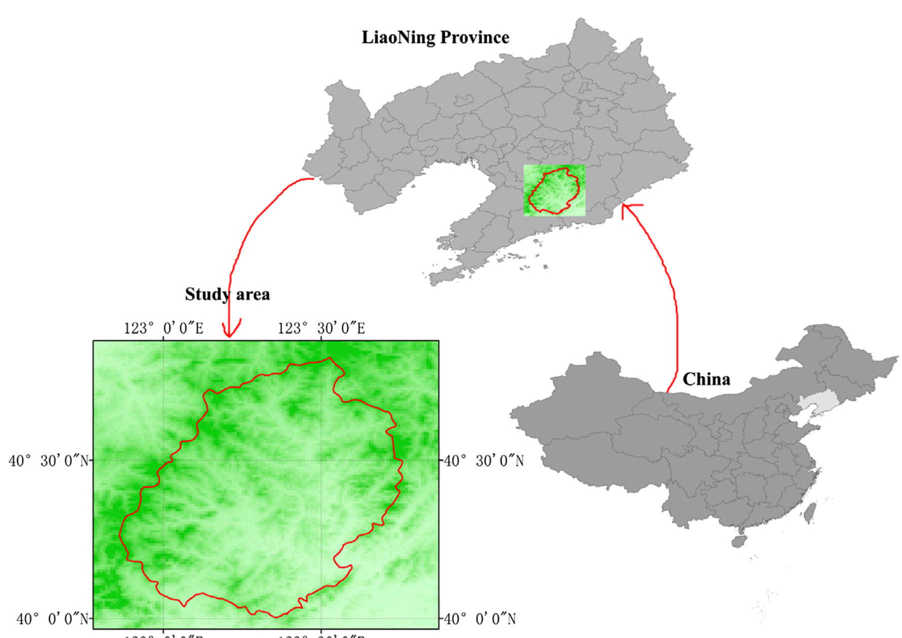

Fig. 1. Location map in Xiu Yan county, An Shan town, Liao Ning province, China.