Science of the Total Environment 625 (2018) 575–588

Contents lists available at ScienceDirect

Science of the Total Environment

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

Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China

Haoyuan Hong [a,b,c], Paraskevas Tsangaratos [d,*], Ioanna Ilia [d], Junzhi Liu [a,b,c,**], A-Xing Zhu [a,b,c,**], Wei Chen [e]

[a] Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing 210023, China
[b] Jiangsu Center for Collaborative Innovation in Geographic Information Resource Development and Application, Nanjing, Jiangsu 210023, China
[c] State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing 210023, China
[d] National Technical University of Athens, School of Mining and Metallurgical Engineering, Department of Geological Sciences, Laboratory of Engineering Geology and Hydrogeology, Zografou Campus: Heroon Polytechniou 9, 15780 Zografou, Greece
[e] College of Geology and Environment, Xi'an University of Science and Technology, Xi'an 710054, China

HIGHLIGHTS

- Identifying non-flood areas with fuzzy logic approach
- Combining fuzzy weight of evidence method and data mining techniques
- Fuzzy WofE-Support Vector Machine appears as the most accurate model.
- Assist researchers and local authorities in flood mitigation strategies

GRAPHICAL ABSTRACT

ARTICLE INFO

Article history:
Received 18 October 2017
Received in revised form 18 December 2017
Accepted 21 December 2017
Available online xxxx

Editor: Ouyang Wei

Keywords:
Flood susceptibility
Fuzzy WofE
Data mining methods
China

ABSTRACT

In China, floods are considered as the most frequent natural disaster responsible for severe economic losses and serious damages recorded in agriculture and urban infrastructure. Based on the international experience prevention of flood events may not be completely possible, however identifying susceptible and vulnerable areas through prediction models is considered as a more visible task with flood susceptibility mapping being an essential tool for flood mitigation strategies and disaster preparedness. In this context, the present study proposes a novel approach to construct a flood susceptibility map in the Poyang County, JiangXi Province, China by implementing fuzzy weight of evidence (fuzzy-WofE) and data mining methods. The novelty of the presented approach is the usage of fuzzy-WofE that had a twofold purpose. Firstly, to create an initial flood susceptibility map in order to identify non-flood areas and secondly to weight the importance of flood related variables which influence flooding. Logistic Regression (LR), Random Forest (RF) and Support Vector Machines (SVM) were implemented considering eleven flood related variables, namely: lithology, soil cover, elevation, slope angle, aspect, topographic wetness index, stream power index, sediment transport index, plan curvature, profile curvature and distance from river network. The efficiency of this new approach was evaluated using area under curve (AUC) which measured the prediction and success rates. According to the outcomes of the performed analysis, the fuzzy WofE-SVM model was the model with the highest predictive performance (AUC value,

* Corresponding author.
** Correspondence to: J. Liu and A.X. Zhu, Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing 210023, China.
E-mail addresses: ptsag@metal.ntua.gr (P. Tsangaratos), azhu@wisc.edu (A.-X. Zhu).

https://doi.org/10.1016/j.scitotenv.2017.12.256
0048-9697/© 2017 Elsevier B.V. All rights reserved.

0.9865) which also appeared to be statistical significant different from the other predictive models, fuzzy WofE-RF (AUC value, 0.9756) and fuzzy WofE-LR (AUC value, 0.9652). The proposed methodology and the produced flood susceptibility map could assist researchers and local governments in flood mitigation strategies.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Floods are among the natural disasters which are responsible for severe economic losses and serious damages recorded in agriculture and urban infrastructure (Zhang and Li, 2007; Garcia-Castellanos et al., 2009; Nie et al., 2012; Li et al., 2015). In China, floods are considered as the most frequent natural disaster, the occurence of which is attributed to the presence of the East Asian monsoon (Renyi and Nan, 2002; Yu et al., 2009). According to the National Climate Center of China, an average annual economic loss of about 17 billion USA dollars, since 1990 have been recorded, with the two-thirds of Chinese territory area and over half of the total population affected by floods almost every year (Huang et al., 2008; Nakayama and Watanabe, 2008; Zhang et al., 2011, 2015; Wang et al., 2012). Concerning the number of flood occurrence, damages and deaths associated with floods for the period 1990 till 2016 reveals a clear linear increase in the number of floods with a analogous increase in the total cost damages, while the number of deaths associated with floods after 2000 appear to have decreased (Fig. S.1a,b).

The magnitude of impact and the irreversible nature of damages caused by floods make the implementation of flood control and prevention measures a necessity (Billa et al., 2006; Shankman et al., 2006; Hu et al., 2014). The main efforts of the scientific community are focused in providing to disaster response teams and flood management officials, reference data and calibration information for flood susceptibility maps, dynamic flood models, and damage estimates (Kourgialas and Karatzas, 2011). Also a great part of flood management mitigation strategies and planning provide estimations about the associated flood hazard and risk index in terms of its location, magnitude and distribution. However, as many researchers have stated flooding is a complex phenomenon while the exact prediction of spatial and temporal flood occurrence is a difficult task (Pappenberger et al., 2006; Sarhadi et al., 2012; Chapi et al., 2017). The most common flood hazard and risk analysis is the implementation of hydrologic and hydraulic models which provide information about the flood inundation extent, the water depth and the velocity based on 1-D or 2-D hydraulic models (Mazzoleni et al., 2014; Gharbi et al., 2016). However, 1D or 2D models are mainly applied to small scale projects, since the implementation of such models require a huge amount of data with high precision. In addition, physically based models which integrates hydrologic and hydraulic models requires large processing power and computation time making them less attractive investigation tools for regional scale analysis (Li et al., 2012; Shrestha et al., 2013; Buahin and Horsburgh, 2015). Wing et al. (2017), also reported the need of complex datasets that are neccessary for implementing, calibrating and validating such models.

In the literature one can find alternative analyzing techniques that provide a qualitative or quantitative evaluation of the flood susceptibility, hazard and risk potential of an area (Chen et al., 2017c, 2017d). They differ from the conventional hydrological and hydraulic models in terms of the amount and characteristics of the data they need and also the way they handle data (Tien Bui et al., 2016). Qualitative techniques for flooding susceptibility, in which the process and the results depend on expert knowledge and quantitative techniques that depend on numerical expressions of the relation between independent parameters and flooding occurrence. Examples of qualitative techniques involve analytical hierarchy process (AHP) (Chen et al., 2011; Stefanidis and Stathis, 2013; Kazakis et al., 2015; Rahmati et al., 2016) and fuzzy logic (Pulvirenti et al., 2011; Zou et al., 2013), while some of the most popular quantitative techniques methods in flood susceptibility analysis are frequency ratio (FR) (Tehrany et al., 2014a), weights-of evidence (WOE) (Khosravi et al., 2016; Rahmati et al., 2016; Tehrany et al., 2014b), logistic regression (LR) (Fekete, 2009; Tehrany et al., 2014a), artificial neural network (ANN) (Kia et al., 2012; Nikoo et al., 2016; Kourgialas and Karatzas, 2017), support vector machine (SVM) (Tehrany et al., 2014b, 2015a, 2015b), naïve Bayes (Liu et al., 2015), random forest (Chapi et al., 2017), decision tree (Tehrany et al., 2013) and neural-fuzzy approach (Tien Bui et al., 2016).

The majority of the quantitative models that belong to the domain of data mining methods provide rapid and accurate prediction models. However, flood studies that are performed by data mining methods and especially, supervised methods, are influenced mainly by two particularities. The first involves mapping the exact extent of flood areas when an incidence occurs and the second involves the identification of non-flood areas (Diakakis et al., 2012; Tien Bui and Hoang, 2017). Non-flood areas are an essential component along with the actual flood areas in both the training and validation process. In most cases, floods are represented by points that are located in the centroid of the flood area that has been mapped by analyzing air photos, satellite images and also conducting field surveys. As for the identification of non-flood areas, the most common approach is based on random sampling techniques (Tehrany et al., 2013, 2014a; Nikoo et al., 2016). The same particularities appear in landslide susceptibility and hazard analysis. According to Melchiorre et al. (2008), it is rather easy to identify areas where landslides have already occurred, but it is difficult to identify statistically meaningful examples of stable areas.

Alternative sampling techniques could be found in the literature, in which probabilistic distance metric have been applied (Tsangaratos and Benardos, 2014; Kornejady et al., 2017). Based on the similarity between a set of conditions to a target set of conditions that describe unstable conditions, a first level susceptibility map is produced and then stable or positive areas in areas which exceed a certain threshold can be identified.

In this context, a similar approach was decided to be followed: the construction of a first level flood susceptibility map, the identification of non-flood areas and later the implementation of data mining methods for the final flood susceptibility map. The novelty of the followed approach was the introduction of a new concept of sampling technique based on fuzzy logic, which combines, a data-driven method characterized by its simplicity and straightforward interpretation of the weights, and a knowledge-driven method, which was based on a group of expert's judgment. As reported by several researchers, models that combine fuzzy logic and expert knowledge are characterized as simple and flexible models that require minimum data, capable of solving problems when the system is complex enough to solve by physically based or conceptual models (Zhu et al., 2001; Locatelli et al., 2011; Perera and Lahat, 2014; Hong et al., 2017a). In the present study, the fuzzy weight of evidence (FWofE) method developed by Cheng and Agterberg (Cheng and Agterberg, 1999) combined with was implemented in order to produce the first level flood susceptibility map, while data mining techniques, LR, RF and SVM following an optimized procedure were used for the construction of the final flood susceptibility map. The efficiency and usefulness of the developed approach was tested in Poyang County, located in the North of the Jiangxi Province, China.

2. Study area

The Poyang County is located in the North of the Jiangxi Province with Poyang Lake spreading in the west of it. The total population of Poyang

Download English Version:

https://daneshyari.com/en/article/8861149

Download Persian Version:

https://daneshyari.com/article/8861149

Daneshyari.com