IOP Conf. Series: Materials Science and Engineering **712** (2020) 012005 doi:10.1088/1757-899X/712/1/012005

Flash Flood Susceptibility Modelling: A Review

A Saleh[1], A Yuzir[1] and I Abustan[2]
[1]Disaster Preparedness & Prevention Centre (DPPC), Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM), 54100 Kuala Lumpur, Malaysia
[2]School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

Abstract. Flash flood is common problem in the developed or developing country especially in urban areas. This is due to the changes of land use, poor planning of drainage system, deforestation and unplanned land use and etc. Flash flood susceptibility modelling (FFSM) for urban area is important to avoid the loss and etc. FFSM can be vital component and essential tools for planning and management of natural disaster and environmental. The objective of this paper is to review selected paper on flash flood susceptibility modelling using Geographic Information System (GIS) coupling empirical models. A lot of methods have been coupling with GIS to developed flood susceptibility modelling, for example, Weight of Evidence, Random Forest, etc. However, there is still lack of susceptibility model for flash flood. The GIS coupling method can improve the accuracy of flood susceptibility modelling. Thus, this paper will review the ability of GIS coupling with empirical models for the flash flood susceptibility modelling.

1. Introduction

The occurrences of flash flood cannot be predicted and it is worst hydro-meteorological disaster [1]. Flash flood disasters occur frequently due to global climate change [2]. Flood events commonly cause destruction to agricultural crops and property, and may even result in the loss of human lives [3]. According to Elkhrachy [5], flash flood can be defined the flood which is begin in a short period of time and normally show high peak discharges. Elkhrachy [5] added usually the geomorphic low-lying zones will hit by flash flood when heavy rainfall.

By develop the flood susceptibility mapping and modelling, it can help local authority in flood management to identify the most sensitive zones for civil protective actions, assess damages, and make valid urban planning [5]. By taking the definition of landslide susceptibility mapping, Santangelo et al. [6] defined the susceptibility mapping as the probability that a risk occurrence happens in a particular area and in a not determined date. Susceptibility mapping was developed based on the relationship of the conditioning factors with the distribution of previous events [6]. That means susceptibility modelling is actually just an estimate of "where" disasters such as landslide or flood are likely to occur. Susceptibility map also known as natural hazard potential map for some groups.

According to Reichenbach et al. [7], there are confusion between "susceptibility" and "hazard" modelling. Even though these two terms usually used as same meaning but it these two words stating different models [7]. Reichenbach et al. [7] added that susceptibility model does not consider the size of natural hazard e.g., the length, width, depth, area or volumes. Meanwhile, hazard model would predicting "where" a flood or landslide will take place, "when" or "how frequently" it will happen, and "how large" it will be [8]. Flood susceptibility map will manage any future flood problems [9] and important step to predict and manage the future flood event [10].

2. Modelling Methods

There are many methods of flash flood susceptibility mapping; traditional bases hydrological methods, statistical, and machine learning based methods. Rule-based and automated modelling methods have outperform outdated traditional flood models due to their more suitable for hazard analyses [11]. In recent times, researchers have coupling many methods with GIS to increase the accuracy prediction of flooded areas (flood susceptibility). These methods are including qualitative method such as analytical hierarchy process (AHP), quantitative techniques such as weight of evidence (WoE) and frequency ratio (FR), and machine learning method such as artificial neural networks (ANN). However, there are still lack of studies have been carried out for flash flood susceptibility modelling. This paper aims to review selected study of flash flood susceptibility modelling.

3. Flash Flood Susceptibility Modelling Study

Bui et al. [12] is using feature selection method (FSM) and tree based ensemble methods. The study area located at the Bao Yen district and the Bac Ha district of Lao Cai Province, Vietnam. Bui et al. [12] were used 654 floods and 12 conditioning factors (stream density, toposhade, slope, curvature, stream power index (SPI), elevation, topographic wetness index (TWI), rainfall, aspect, normalize difference vegetation index (NDVI), soil type cover and lithology). This method used a fuzzy rule based algorithm (FURIA) which is as attribute evaluator, while Genetic Algorithm (GA) as the search method. This is to get best set of conditioning factors used in modelling assessments. Then, the FURIA-GA method was combined with LogitBoost, Bagging and AdaBoost ensemble algorithms for prediction model. Based on the result, FURIA-GA-Bagging (93.37%) outperformed the other ensemble algorithm, FURIA-GA-LogitBoost (92.35%) and FURIA-GA-AdaBoost (89.03%). The flash flood susceptibility mapping developed by using FURIA-GA-Bagging method is shown in Figure 1.

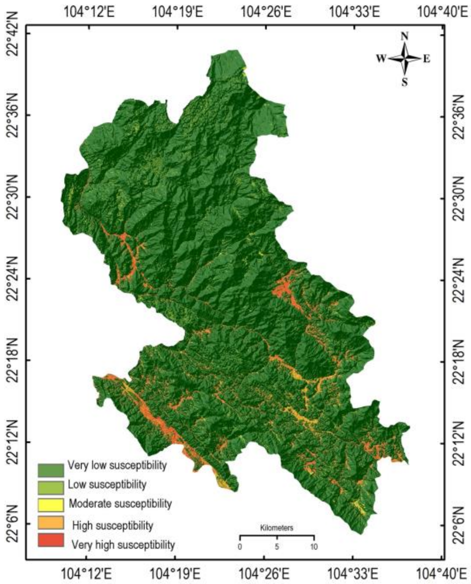

Figure 1. Flash flood susceptibility mapping developed by using FURIA-GA-Bagging method [12]

Youssef and Hegab [13] were applied Analytic hierarchy process (AHP) which is can categorize in qualitative method. The main objective of this study is to examine the effectiveness and reliability of AHP. The location of study area is Ras Gharib, Egypt. This study also using high-resolution images obtained after previous flood events in 2016. These high-resolution images are used to validate the susceptibility model by using a slicing technique and historical flood data. Youssef and Hegab [13] only used 7 flood factors - distance from streams, slope, curvature, lithological units, angle, elevation, and topographic wetness index (TWI). However, results of this study shown that AHP give good result for flash flood susceptibility which is 83.3%. The flash flood susceptibility mapping developed by using AHP method is shown in the Figure 2.

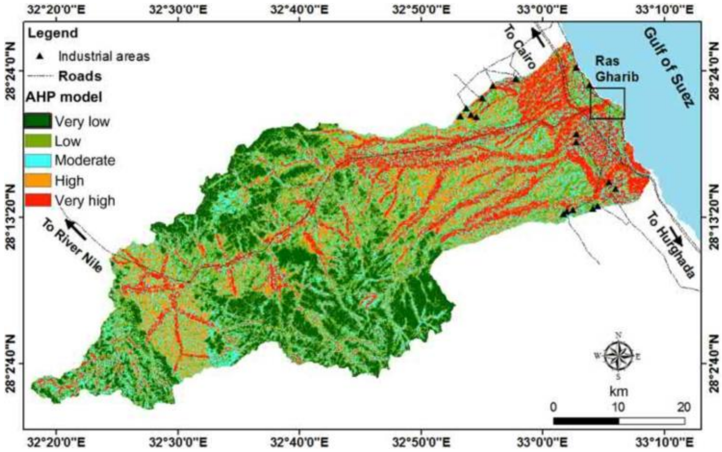

Figure 2. Flash flood susceptibility mapping developed by using AHP method [13]

Ngo et al. [14] analysed flash flood in two districts Bac Ha and Bao Yen, Vietnam by using machine learning techniques which are Firefly algorithm (FA), Levenberg–Marquardt (LM) Backpropagation, and an artificial neural network (FA-LM-ANN), LM-ANN, FA-ANN, support vector machine (SVM) and classification tree (CT). For this study, Ngo et al. [14] used 12 flood factors which are aspect, elevation, stream power index (SPI), slope, topographic wetness index (TWI), stream density, rainfall, curvature, normalized difference vegetation index (NDVI), lithology, toposhade, and soil type and 654 flash flood locations were identified to evaluate the flash flood model. According to the results, the integrated FA-LM-ANN gives the good results which is 97.0% following by LM-ANN - 92.6%, FA-ANN – 91.9%, SVM – 92.9% and CT – 90.8%. Figure 3 shows the flash flood susceptibility mapping developed by using FA-LM-ANN method.

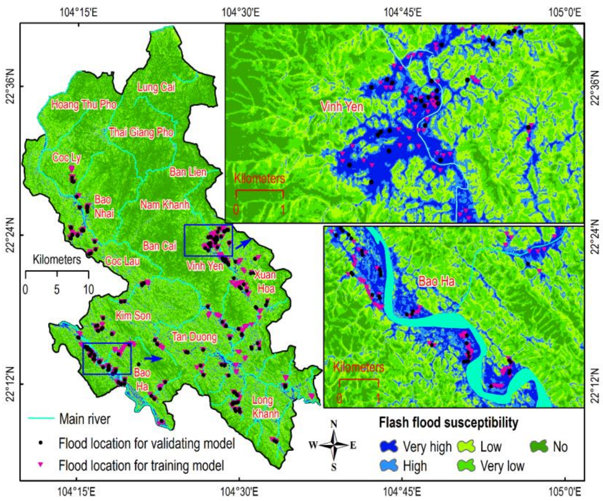

Figure 3. Flash flood susceptibility mapping developed by using FA-LM-ANN method [14]

Khosravi et al. [15] has done the comparisons between Logistic Model Trees (LMT), Reduced Error Pruning Trees (REPT), Naïve Bayes Trees (NBT), and Alternating Decision Trees (ADT) which is these four methods is decision trees algorithms. The study area is located in Haraz watershed, northern Iran. The 201 flood points were identified and 11 flood factors were selected which are altitude, ground slope, curvature, river density, normalized difference vegetation index (NDVI), topographic wetness index (TWI), distance from river, land use, stream power index (SPI), rainfall, and lithology. The results show that the ADT is the highest AUC value – 97.6 %, followed by NBT – 97.4 %, LMT –

97.1% and REPT - 81.1%, respectively. Before that, Khosravi et al. [16] compare the results between statistical index, Shannon's entropy, and weighting factor models for flash flood susceptibility model. For this study, Khosravi et al. [16] was used 211 flood locations were identified and 10 flash flood conditioning factors was used in this study area which are plan curvature, stream power index (SPI), land use, geology, distance from river, topographic wetness index (TWI), rainfall, slope angle, altitude, and normalized difference vegetation index (NDVI). The result show that the statistical index model with the prediction of 98.72% is the highest rate of prediction followed by weighting factor and Shannon's entropy models with the 97.6% and 92.42% for the prediction rate. The flash flood susceptibility mapping produced by using Alternating Decision Trees (ADT) and statistical index (SI) are presented in the Figure 4.

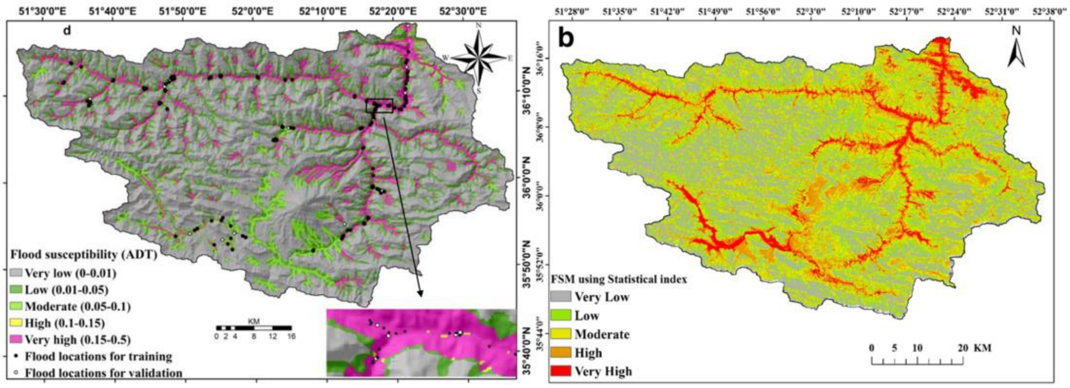

Figure 4. Flash flood susceptibility mapping produced by using Alternating Decision Trees (ADT) [15] and statistical index (SI) [16]

Youssef et al. [17] were used bivariate (frequency ratio) and multivariate statistical models (ensemble frequency ratio and logistic regression) to develop flash flood susceptibility modelling in Jeddah, Saudi Arabia. In the study, 127 flood locations were identified and 7 flood factors were used to develop flash flood model which are slope, elevation, curvature, geological units, land use, soil drain, and distance from streams. The results showed that the prediction accuracy achieved using ensemble technique between Frequency Ratio (FR) and Logistic Regression (LR) has the best (91.3%) followed by the traditional Frequency Ratio method - 89.6 %. This results show that bivariate statistical analysis (BSA) method can improve by integrated with Multivariate statistical analysis (MSA) method. Table 1 summarizes the methods used in flash flood susceptibility model. The flash flood susceptibility mapping developed by using multivariate statistical models (ensemble frequency ratio and logistic regression) is shown in Figure 5.

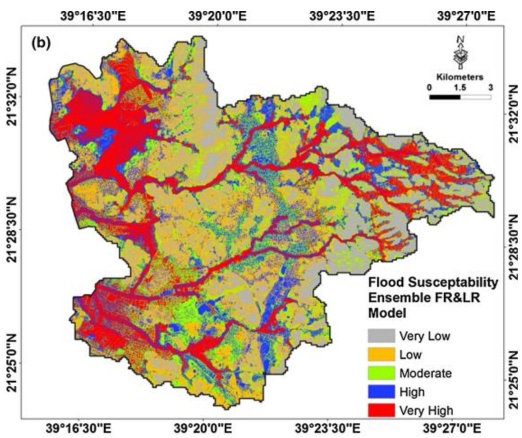

Figure 5. Flash flood susceptibility mapping developed by using multivariate statistical models (ensemble frequency ratio and logistic regression) [17]

Table 1. The summarizes of previous flash flood susceptibility model studies.

Authors	Method	Data/Conditioning Factors	Finding/Result
Bui et al. [12]	The feature selection method (FSM), used a fuzzy rule based algorithm Fuzzy Unordered Rules Induction Algorithm (FURIA), as attribute evaluator, whereas Genetic Algorithms (GA) were used as the search method, in order to obtain optimal set of variables used in flood susceptibility modeling assessments. The novel FURIA-GA was combined with LogitBoost, Bagging and AdaBoost ensemble algorithms.	654 flood locations were identified 12 flood factors - Elevation, slope, aspect, curvature, TWI, SPI, toposhade, stream density, rainfall, NDVI, soil type cover and lithology cover	FURIA-GA-Bagging - 93.37% FURIA-GA-LogitBoost - 92.35% FURIA-GA-AdaBoost -89.03%
Youssef and Hegab [13]	Analytical hierarchy process (AHP)	232 flood locations were identified 7 flood factors - Slope, angle, elevation, distance from streams, lithological units, TWI, and curvature.	AHP - 83.3%
Ngo et al. [14]	Firefly algorithm (FA), Levenberg–Marquardt (LM) Backpropagation, and an artificial neural network (FA-LM-ANN), LM-ANN, FA-ANN, support vector machine (SVM) and classification tree (CT)	654 flash flood locations were identified 12 flood factors - Elevation, slope, aspect, curvature, TWI, SPI, toposhade, stream density, rainfall, NDVI, soil type, and lithology	FA-LM-ANN -97.0% LM-ANN - 92.6% FA-ANN – 91.9% SVM – 92.9% CT – 90.8%
Khosravi et al. [15]	Decision trees algorithms Logistic Model Trees (LMT), Reduced Error Pruning Trees (REPT), Naïve Bayes Trees (NBT), and Alternating Decision Trees (ADT)	201 present and past flood locations were identified 11 flood factors - Ground slope, altitude, curvature, SPI, TWI, land use, rainfall, river density, distance from river, lithology, and NDVI.	ADT – 97.6 % NBT – 97.4 % LMT – 97.1% REPT - 81.1%
Khosravi et al. [16]	Shannon's entropy, statistical index, and weighting factor models	211 flood locations were identified 10 flood factors - Slope angle, plan curvature, altitude, TWI, SPI, distance from river, rainfall, geology, land use, and NDVI	SI - 98.72 % WF - 98.1% SE - 92.53%
Youssef et al. [17]	bivariate and multivariate statistical models (frequency ratio and ensemble frequency ratio and logistic regression)	127 flood locations were identified 7 flood factors - Slope, elevation, curvature, geological units, landuse, soil drain, and distance from streams	FR - 89.6 % FR+LR - 91.3%

4. Conclusion

In summary, the purpose of this paper was to review selected study on flash flood susceptibility modelling (FFSM). Based on the selected studies, the ability of integration between GIS and analytical model shows the good result in developing flash flood susceptibility model. The ensemble and hybrid model of machine learning can increase the accuracy of result. The methods of FFSM are also play important role. The studies carried out by using machine learning method shows that improvement of result when they applied the machine learning technique at same study area. It also found that the result of machine learning method has a higher accuracy than qualitative and quantitative method (WoE, FR, and AHP) even the number of flooded points was decrease compare with previous study. It is proved that the ability of integration between GIS and analytical model is able to generate the flash flood susceptibility modelling which are very useful tools to plan and manage the flood.

5. References

[1] Mahmood S and Rahman A 2019 Flash flood susceptibility modeling using geo-morphometric and hydrological approaches in Panjkora Basin, Eastern Hindu Kush, Pakistan *Environ. Earth Sci.* **78** 0

[2] Li H, Zhang S, Li Q, Zhang X and Guo L 2019 Research on flash flood disaster warning index: Case study of Luoning County *IOP Conf. Ser. Earth Environ. Sci.* **218** 012075

[3] Dahri N and Abida H 2017 Monte Carlo simulation-aided analytical hierarchy process (AHP) for flood susceptibility mapping in Gabes Basin (southeastern Tunisia) *Environ. Earth Sci.* **76** 1–14

IOP Conf. Series: Materials Science and Engineering **712** (2020) 012005 doi:10.1088/1757-899X/712/1/012005

[4] Elkhrachy I 2015 Flash Flood Hazard Mapping Using Satellite Images and GIS Tools: A case study of Najran City, Kingdom of Saudi Arabia (KSA) *Egypt. J. Remote Sens. Sp. Sci.* **18** 261–78
[5] Sadek M and Li X 2019 Low-Cost Solution for Assessment of Urban Flash Flood Impacts Using Sentinel-2 Satellite Images and Fuzzy Analytic Hierarchy Process : A Case Study of Ras Ghareb City , Egypt **2019**
[6] Santangelo N, Santo A, Di Crescenzo G, Foscari G, Liuzza V, Sciarrotta S and Scorpio V 2011 Flood susceptibility assessment in a highly urbanized alluvial fan: The case study of Sala Consilina (southern Italy) *Nat. Hazards Earth Syst. Sci.* **11** 2765–80
[7] Reichenbach P, Rossi M, Malamud B D, Mihir M and Guzzetti F 2018 A review of statistically-based landslide susceptibility models *Earth-Science Rev.* **180** 60–91
[8] Basith A 2011 Landslide Susceptibility Modelling Under Environmental Changes: A Case Study of Cameron Highlands, Malaysia 1–290
[9] Kourgialas N N, Karatzas G P, Kourgialas N N and Karatzas G P 2011 Flood management and a GIS modelling method to assess flood-hazard areas — a case study Flood management and a GIS modelling method to assess flood-hazard areas — a case study **6667**
[10] Kourgialas N N and Karatzas G P 2017 A national scale flood hazard mapping methodology: The case of Greece – Protection and adaptation policy approaches *Sci. Total Environ.* **601–602** 441–52
[11] Tehrany M S, Kumar L, Neamah Jebur M and Shabani F 2019 Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods *Geomatics, Nat. Hazards Risk* **10** 79–101
[12] Bui D T, Tsangaratos P, Ngo P T, Pham T D and Pham B T 2019 Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods *Sci. Total Environ.* **668** 1038–54
[13] Youssef A M and Hegab M A 2019 Flood-Hazard Assessment Modeling Using Multicriteria Analysis and GIS: A Case Study—Ras Gharib Area, Egypt *Spatial Modeling in GIS and R for Earth and Environmental Sciences* pp 229–57
[14] Ngo P T, Hoang N and Pradhan B 2018 A Novel Hybrid Swarm Optimized Multilayer Neural Tropical Areas Using Sentinel-1 SAR Imagery and
[15] Khosravi K, Pham B T, Chapi K, Shirzadi A, Shahabi H, Revhaug I, Prakash I, Tien Bui D, Shahabi H, Chapi K, Shirzadi A, Pham B T, Khosravi K and Revhaug I 2018 A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran *Sci. Total Environ.* **627** 744–55
[16] Khosravi K, Pourghasemi H R, Chapi K and Bahri M 2016 Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models *Environ. Monit. Assess.* **188**
[17] Youssef A M, Pradhan B and Sefry S A 2016 Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models *Environ. Earth Sci.* **75** 1–16

Acknowledgments
The first author would like to acknowledge Malaysia-Japan International Institute of Technology (MJIIT), UTM for providing Incentive Scheme to undertake his PhD program.