GIS-Based Flood Susceptibility Mapping Using Statistical Index and Weighting Factor Models

Worawit Suppawimut*

Department of Geography, Faculty of Humanities and Social Sciences, Chiang Mai Rajabhat University, Thailand

ARTICLE INFO

Received: 12 May 2021
Received in revised: 8 Jul 2021
Accepted: 12 Jul 2021
Published online: 25 Aug 2021
DOI: 10.32526/ennrj/19/2021003

Keywords:
Flood susceptibility/ Hazard mapping/ Statistical index/ San Pa Tong District/ Bivariate statistics

*** Corresponding author:**
E-mail: worawit_sup@cmru.ac.th

ABSTRACT

Floods are one of the most devastating natural hazards, causing deaths, economic losses, and destruction of property. Flood susceptibility maps are an essential tool for flood mitigation and preparedness planning. This study mapped flood susceptibility using statistical index (SI) and weighting factor (WF) models in San Pa Tong District, Chiang Mai Province, Thailand. The conditioning factors used to perform flood susceptibility mapping were elevation, slope, aspect, curvature, topographic wetness index, stream power index, rainfall, distance from rivers, stream density, soil drainage, land use, and road density. The flood data were randomly classified as training data for mapping (70% of data) and testing data for model validation (30% of data). The results revealed that the SI and WF models classified 49.49% and 51.74% of the study area, respectively, as very highly susceptible to flooding. In the WF model, the factors with the greatest influence were land use, soil drainage, and elevation. The validation of the models using the area under the curve revealed that the success rates of the SI and WF models were 91.80% and 93.06%, while the prediction rates were 92.05% and 93.52%, respectively. The results from this study can be useful for local authorities in San Pa Tong District for flood preparedness and mitigation.

1. INTRODUCTION

Flooding is the most devastating category of natural disaster, affecting people and properties around the world (Paul et al., 2019; Samanta et al., 2018), with an average of 71.9 million people reported as affected by flooding annually (CRED, 2020). In the monsoon-dominated tropical and subtropical regions of the world, flooding occurs frequently and across wide areas (Khaing et al., 2021). In Asia, flooding is one of the most destructive natural disasters, with the highest proportion (79.9%) of the total population affected by floods globally (CRED, 2020). In Thailand, floods are one of the most destructive natural disasters. In 2011, the country suffered the worst floods in more than half a century; these floods inundated more than six million hectares of land in 66 provinces and affected more than 13 million people. The estimated damage and losses totaled approximately USD 46.5 billion (World Bank, 2012).

Flood occurrence is affected by various factors. Heavy rainfall is one of the main factors, leading to the rapid accumulation and release of runoff waters from upstream to downstream areas (Kongmuang et al., 2020). Climate change also causes flood occurrences with rising frequency and magnitude (Khaing et al., 2021; Tehrany et al., 2019), while human activities such as urbanization and deforestation can also increase flooding incidence rates (Cabrera et al., 2019; Paul et al., 2019). In flood-prone areas, flood risk can be assessed to prevent damage to residential areas, agriculture, public properties, etc. (Paul et al., 2019; Samanta et al., 2018). Flood susceptibility mapping is an essential tool for flood preparedness and mitigation, in particular, planning using reliable information can help support communities and government authorities to precisely implement flood protection strategies.

Various approaches have been applied for flood susceptibility mapping. Hydrological models have been developed by various researchers such as SWAT (Igarashi et al., 2019) and HEC-RAS (Khaing et al., 2021; Rahmati et al., 2016). Although hydrological model can predict and simulate flood hazard, there are

Citation: Suppawimut W. GIS-based flood susceptibility mapping using statistical index and weighting factor models. Environ. Nat. Resour. J. 2021;19(6):481-493. (https://doi.org/10.32526/ennrj/19/2021003)

some limitations such as the requirement of vast data budget, unavailability of large-scale data and time consuming for preparation and calibration of parameters (Cabrera et al., 2019; Hoang et al., 2020). In the few past decades, geographic information system (GIS) and remote sensing (RS) have made remarkable contributions in flood hazard mapping (Rahmati et al., 2016; Samanta et al., 2018). Several techniques have been applied with GIS and RS, including analytical hierarchy process (AHP) (Hoang et al., 2020; Khaing et al., 2021; Rahmati et al., 2016), frequency ratio (FR) (Anucharn, 2019; Cao et al., 2016; Samanta et al., 2018; Tehrany et al., 2019), logistic regression (LR) (Tehrany et al., 2019), weight of evidence (WoE) (Tehrany et al., 2017), statistical index (SI) (Cao et al., 2016; Khosravi et al., 2016; Tehrany et al., 2019), and artificial neural network (ANN) (Anucharn, 2019; Kia et al., 2012). The results of the research mentioned above, demonstrate slightly variance from place to place and each type of model is still necessary to be examined. Thus, testing and valuation of these models can provide optimal and more reliable results.

Statistical index (SI) modeling has been applied to various hazard mapping efforts and has performed efficiently with acceptable results (Khosravi et al., 2016). It has been widely used in mapping landslide susceptibility (Budha et al., 2016; Pourghasemi et al., 2013) and has also been applied to flood susceptibility mapping (Khosravi et al., 2016; Tehrany et al., 2019). However, the main limitation of SI is the lack of consideration of the relationship between the causative factors themselves which needs further research. The weighting factor (WF) method has been applied widely in the field of landslide studies (Yalcin, 2008), however, its use is still lacking in the field of flood susceptibility mapping (Khosravi et al., 2016). Additionally, some causative factors have not been applied to the WF model to measure their impact on flood occurrence such as aspect and road density. Given this context, a comparative study of the SI and WF models may contribute to the assessment of flood susceptibility.

This research, therefore, aimed to perform flood susceptibility mapping of the San Pa Tong District, Chiang Mai, which suffered from flooding in 2005, 2009, 2010, and 2011, by applying SI and WF models and to examine the performance of these two models. This research also aimed to find the most influential factors for flood occurrence in the study area. The study results can be useful for local administrators to minimize the consequences of future floods, and, furthermore, these research methods can provide guidelines for further research.

2. METHODOLOGY

2.1 Study area

The study area is the San Pa Tong District, Chiang Mai Province, in northern Thailand. It is located between latitudes 18°30′ and 18°43′ N and longitudes 98°48′ and 98°57′E, covering an area of approximately 173.45 km^2 (Figure 1). The altitude ranges between 239 m and 640 m above sea level, with a mountainous area in the north and lowlands covering the central and the southern parts of the area. There are three main rivers in San Pa Tong, namely the Ping River, Khan River, and Mae Wang River. Due to the area's topographic and hydrological characteristics, San Pa Tong is flood-prone and has been frequently affected by floods. The main causes of flooding in the area are the high intensity of rainfall and runoff from the upper catchments flowing to the lower areas in the south. Regarding the flood data of 2005, 2009, 2010, and 2011, 59.49 km^2 of San Pa Tong has been recorded as a flooded area (Suppawimut, 2020). In this context, the San Pa Tong District was, therefore, selected as the study area.

2.2 Data collection

The historic flood data was collected from multi-source satellite imagery from 2005 to 2019 including RADARSAT, COSMO, and THAICHOTE, operated by the Geo-Informatics and Space Technology Development Agency (GISTDA). In this study, the flood inventory was prepared based on the floods that occurred in 2005, 2009, 2010, and 2011. Flood raster data were randomly classified as training data (70%) and testing data (30%) (Figures 2 and 3). The conditioning factor data were acquired from secondary data sources and government organizations. The digital elevation model (DEM) data were derived from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) with 30 m × 30 m resolution and was obtained from the Earthdata website (https://earthdata.nasa.gov). The DEM was used to extract the topographic factors, namely slope, aspect, curvature, topographic wetness index (TWI), and stream power index (SPI). The other sources of data were as follows: rainfall data from the Upper Northern Region Irrigation Hydrology Center; land use data of 2018 and soil drainage data from the Land Development Department; road data from Nostra Map; and river and stream data derived from 1:50,000 topographic maps.

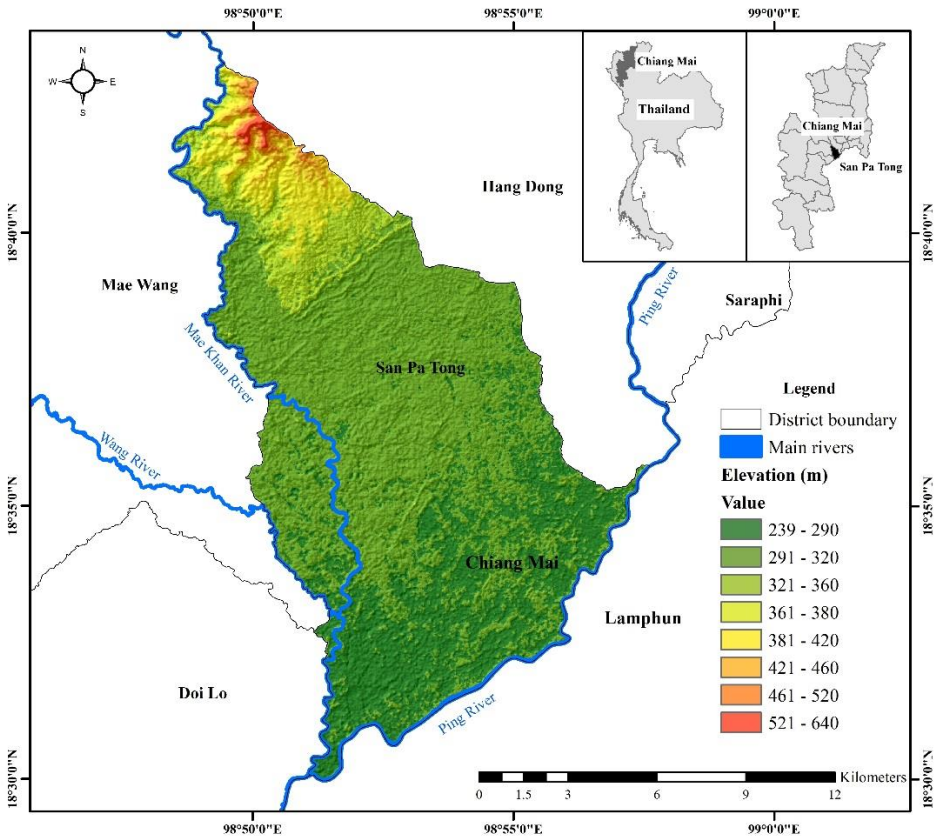

Figure 1. Map of the study area

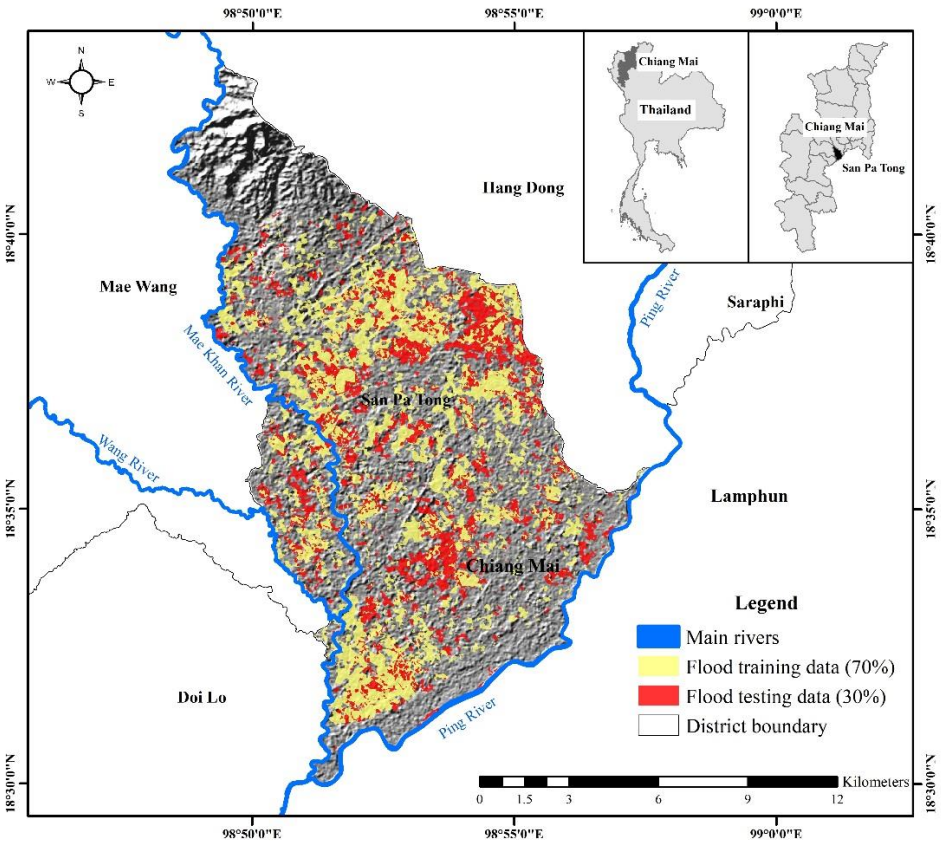

Figure 2. Map of the study area with training and testing data

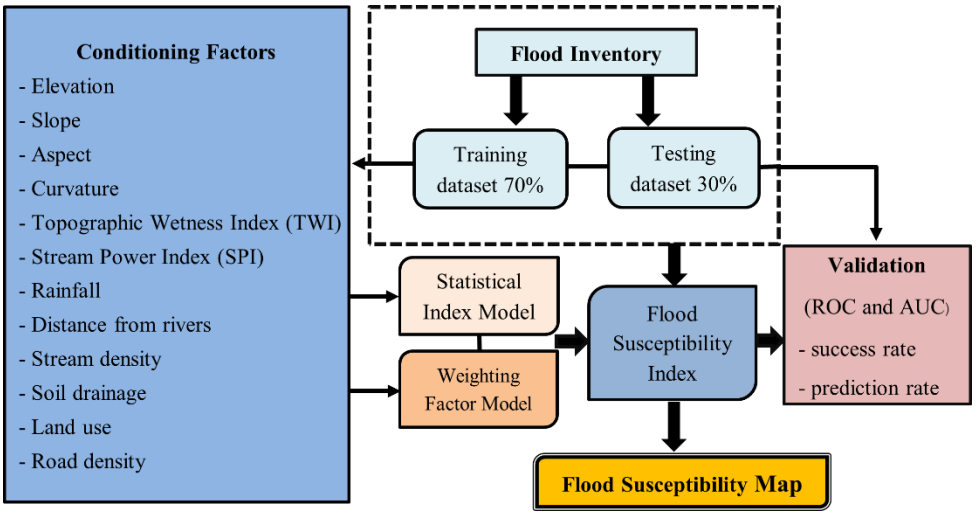

Figure 3. Flow chart of the research methodology

2.3 Flood conditioning factors

Flood conditioning factor data is essential for examining the relationships between the causative factors and flood occurrence (Khosravi et al., 2016). In this study, 12 conditioning factors were considered for flood susceptibility mapping based on the literature (Anucharn, 2019; Khosravi et al., 2016; Kongmuang et al., 2020; Paul et al., 2019; Tehrany et al., 2017; Tehrany et al., 2019), namely elevation, slope, aspect, curvature, TWI, SPI, rainfall, distance from rivers, stream density, soil drainage, land use, and road density (Figures 3 and 4). All conditioning factors were used to perform the flood susceptibility mapping. Each factor was prepared in raster format at a spatial resolution of 30 m × 30 m and was classified using the natural breaks method. Figure 3 shows a flowchart of the research methodology, and Table 1 shows the class values and the characteristics of the conditioning factors.

2.4 Flood susceptibility mapping

2.4.1 Statistical index model

The SI model is a bivariate statistical analysis (BSA) introduced by van Westen et al. (1997). It has been applied to various natural hazard studies, including landslides (Budha et al., 2016; Pourghasemi et al., 2013), floods (Khosravi et al., 2016; Tehrany et al., 2019), and flash floods (Cao et al., 2016). For SI, the weighted value of a conditioning class is calculated as the natural logarithm of flood existence in each class of a conditioning factor divided by the total flood density for the study as expressed in Equation 1 (Tehrany et al., 2019):

$$W_{ij} = \ln\left(\frac{D_{ij}}{D}\right) = \ln\left[\left(\frac{N_{ij}}{S_{ij}}/\frac{N}{S}\right)\right] \quad (1)$$

Where; W_{ij} is the weight given to class i of the factor j, D_{ij} is the flood density in class i of the factor j, D is the total flood density of the study area, N_{ij} is the number of flood pixels in class i of the factor j, S_{ij} is the total number of pixels in class i of the factor j, N is the total number of flood pixels, and S is the total number of pixels in the study area.

The conditioning factors were reclassified using the W_{ij} values. Then, the classified factors were combined using the raster calculator to calculate the flood susceptibility index (FSI). The FSI can be described by the following equation:

$$FSI_{SI} = \sum_{j=1}^{n} W_{ij} \quad (2)$$

Where; FSI_{SI} is the flood susceptibility index of the SI model, W_{ij} is the weight given to class i of the factor j, and n represents the number of conditioning factors.

2.4.2 Weighting factor model

The weighting factor model is a modified version of the SI model (Oztekin and Topal, 2005; Yalcin, 2008; Khosravi et al., 2016). Weights are derived for the conditioning factors to determine their influence on flood occurrence. TSI values are calculated by multiplying the SI values by the number of flood pixels in the same conditioning class, then, the values of all conditioning classes for a particular factor are summed (Oztekin and Topal, 2005). The

weighting factor values for each conditioning factor are calculated, ranging from 1 to 100, using the following equations (Yalcin, 2008; Khosravi et al., 2016):

$$TSI_{value} = \sum_{i=1}^{n} SI \times S.pixel \quad (3)$$

$$W_{wf} = \frac{(TSI_{value}) - (MinTSI_{value})}{(MaxTSI_{value}) - (MinTSI_{value})} \times 100 \quad (4)$$

Where; TSI is the total weighting index value of pixels in the conditioning class for each factor, $MinTSI_{value}$ and $MaxTSI_{value}$ are the minimum and maximum values of the total weighting index value among all conditioning factors, respectively, and W_{wf} is the weighting factor value for each conditioning factor.

To calculate the flood susceptibility index value using the WF model, the W_{ij} weighting value (i.e., W_{ij} of the SI method) of the conditioning class is multiplied by the weighting factor value. The FSI of the WF model is then calculated using the following equation:

$$FSI_{WF} = \sum_{i=1}^{n} SI \times WF \quad (5)$$

Where; FSI_{WF} is the flood susceptibility index of the WF model, SI is the weighting value of the conditioning class, and WF is the weighting factor value of each conditioning factor.

2.4.3 Validation of the model

The receiver operating characteristic (ROC) and the area under the curve (AUC) metrics were used to evaluate the performance of the results of the SI and WF methods. ROC and AUC methods are widely used in natural hazard research (Khosravi et al., 2016; Tehrany et al., 2019). Both susceptibility map results were compared with training data and testing data. The calculated AUC values represent the success rate and prediction rate performance for training data and testing data, respectively. The AUC has a value range from 0-1, where 1 indicates the highest accuracy; if the AUC is closer to 1, the map results are considered more precise and reliable (Tehrany et al., 2019). The AUC value can be classified as follows: weak (0.5-0.6), moderate (0.6-0.7), good (0.7-0.8), very good (0.8-0.9), or excellent (0.9-1.0) (Pourghasemi et al., 2013; Yesilnacar, 2005).

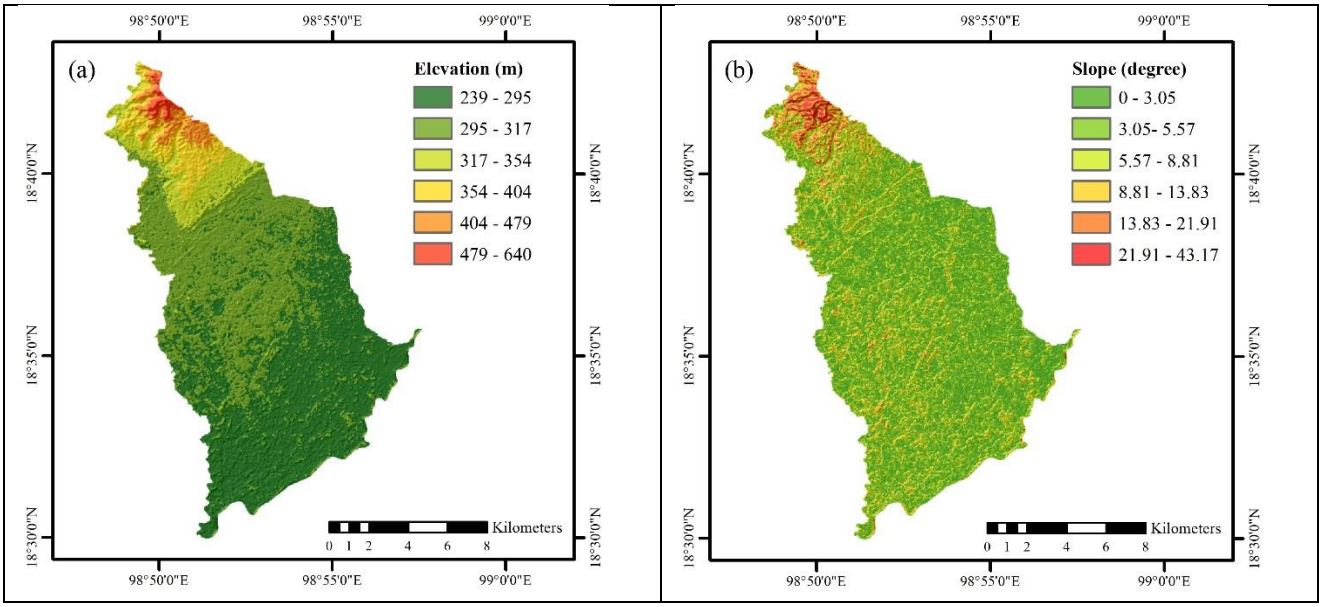

Figure 4. Conditioning factors: (a) elevation, (b) slope, (c) aspect, (d) curvature, (e) TWI, (f) SPI, (g) rainfall, (h) distance from rivers, (i) stream density, (j) soil drainage, (k) land use, and (l) road density

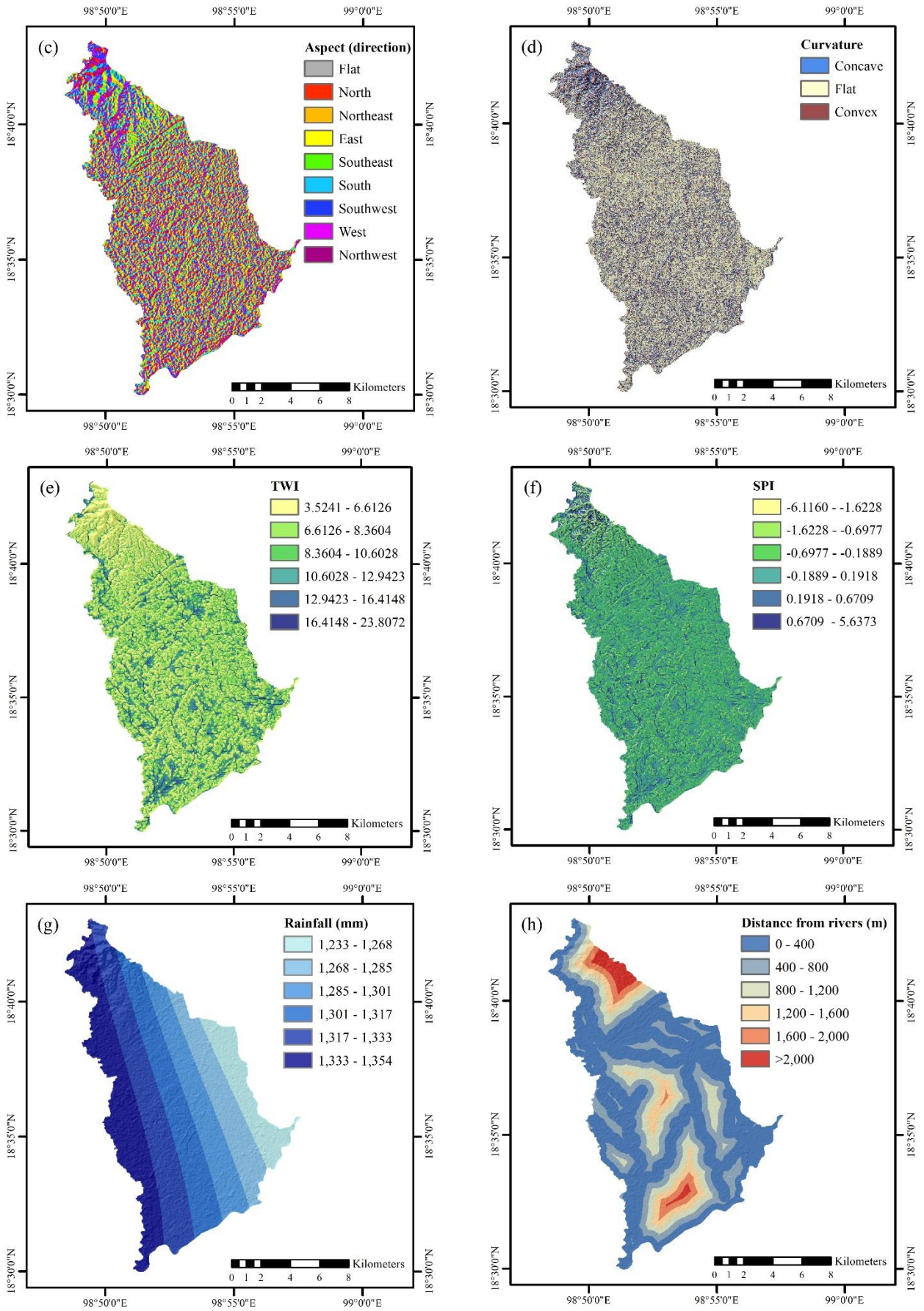

Figure 4. Conditioning factors: (a) elevation, (b) slope, (c) aspect, (d) curvature, (e) TWI, (f) SPI, (g) rainfall, (h) distance from rivers, (i) stream density, (j) soil drainage, (k) land use, and (l) road density (cont.)

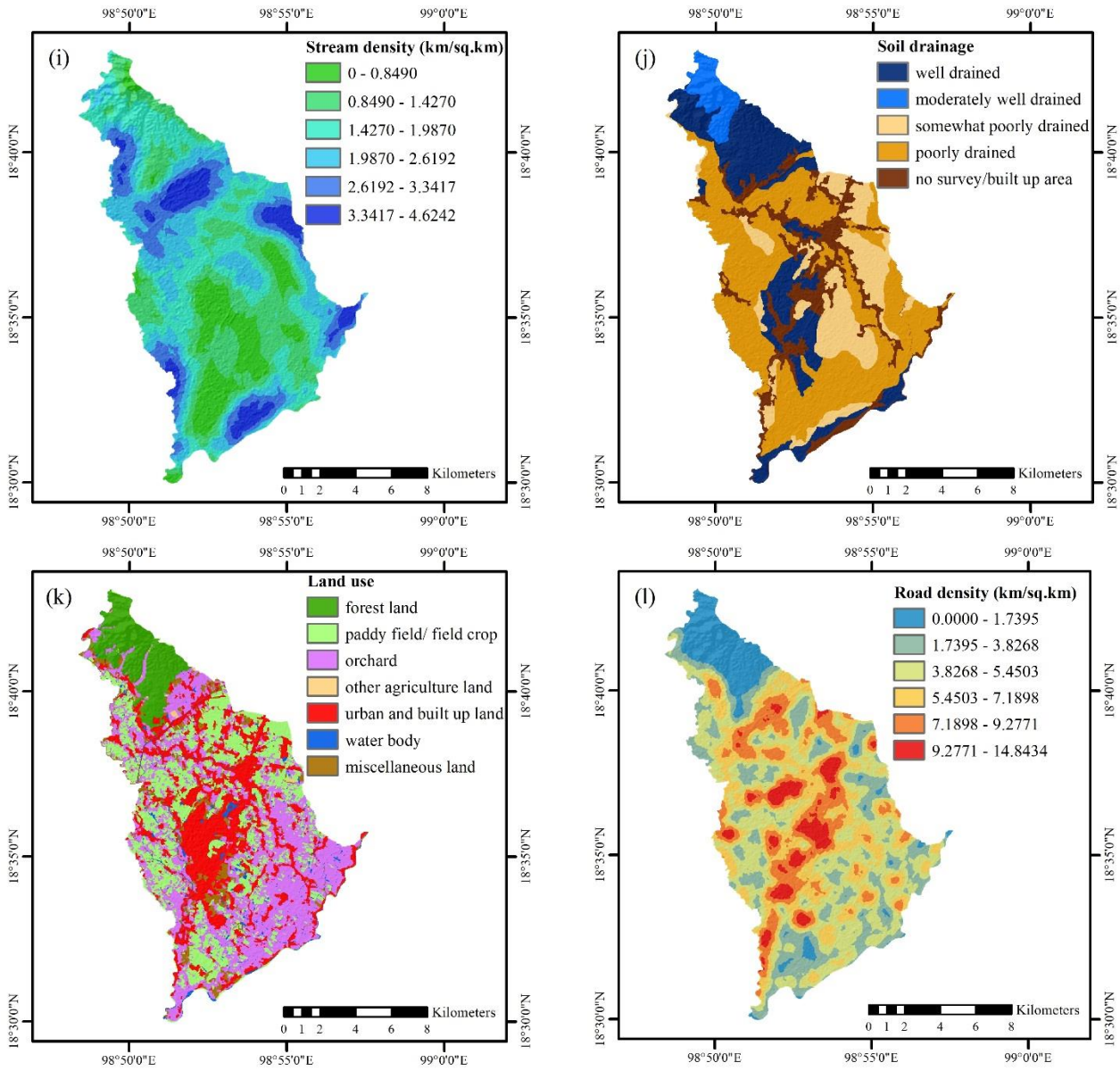

Figure 4. Conditioning factors: (a) elevation, (b) slope, (c) aspect, (d) curvature, (e) TWI, (f) SPI, (g) rainfall, (h) distance from rivers, (i) stream density, (j) soil drainage, (k) land use, and (l) road density (cont.)

3. RESULTS AND DISCUSSION

3.1 Flood susceptibility mapping using SI model

The results of using the SI model to calculate the weight of each flood conditioning factor class represent the correlation with flood occurrence and are presented in Table 1. A positive weight for the conditioning class indicates a high correlation with flood occurrence, whereas a negative weight means a low correlation. Each conditioning class of 12 factors was reclassified using its SI weight (W_{ij}) and was used to calculate the flood susceptibility index (FSI), as expressed in Equation (6). FSI values from the SI model were reclassified into five susceptibility classes (very low, low, moderate, high, and very high) using the geometrical interval classifier in ESRI ArcGIS 10.5 software.

$$FSI_{SI} = SI_{elevation} + SI_{slope} + SI_{aspect} + SI_{curvature} + SI_{TWI} + SI_{SPI} + SI_{rainfall} + SI_{distance\ to\ river} + SI_{stream\ density} + SI_{soil\ drainage} + SI_{land\ use} + SI_{road\ density} \quad (6)$$

The SI results presented in Figure 5 (a) show that about 49.49% of the study area is classified as very high susceptibility. The proportions of the study area classified as high, moderate, low, and very low susceptibility are 23.35%, 23.21%, 3.71%, and 0.24%, respectively (Table 1). The very highly susceptible areas are found in the west, the east, and the south,

while the very low susceptibility regions are mostly found in the northern area.

3.2 Flood susceptibility mapping using WF model

The WF model showed the weight of each conditioning factor as an influence on flooding, as shown in Table 1. The advantage of the WF approach is its consideration of the different weights among the factors. The results showed that land use, soil drainage, and elevation are the most influential factors, with WF weights of 100, 82.61, and 75.77, respectively (Table 1). This indicates the importance of these factors and their necessity for flood susceptibility mapping research. The remaining conditioning factors are, in order of influence: road density, slope, distance from rivers, TWI, SPI, stream density, rainfall, aspect, and curvature. In contrast, the study of Khosravi et al. (2016) in the Haraz watershed of Iran found that the most influential factors were distance from rivers, elevation, and TWI.

Table 1. Calculation of weight values of SI and WF models

Conditioning factors	Conditioning classes	No. pixels	Percentage of area	No. of flood pixels	Percentage of flood	SI(W_{ij})	TSI	WF
Elevation (m)	239-295	91,113	47.2764	21,527	51.3305	0.0823	1,771	75.77
	295-317	71,637	37.1708	18,885	45.0308	0.1918	3,623	
	317-354	13,411	6.9587	1,194	2.8471	-0.8937	-1067	
	354-404	9,366	4.8598	332	0.7916	-1.8146	-602	
	404-479	5,467	2.8367	0	0.0000	0.0000	0	
	479-640	1,730	0.8977	0	0.0000	0.0000	0	
Slope (degree)	0-3.0544	66,718	34.6184	16,253	38.7548	0.1129	1,834	26.41
	3.0544-5.5710	65,599	34.0378	15,426	36.7829	0.0776	1,196	
	5.5710-8.8112	37,982	19.7080	7,832	18.6752	-0.0538	-422	
	8.8112-13.8323	15,086	7.8278	2,181	5.2005	-0.4089	-892	
	13.8323-21.9126	5,310	2.7552	241	0.5747	-1.5675	-378	
	21.9126 43.1662	2,029	1.0528	5	0.0119	-4.4808	-22	
Aspect (direction)	Flat	363	0.1884	93	0.2218	0.1633	15	1
	North	23,113	11.9928	5,391	12.8547	0.0694	374	
	Northeast	21,506	11.1590	5,061	12.0678	0.0783	396	
	East	22,893	11.8786	5,237	12.4875	0.0500	262	
	Southeast	26,306	13.6496	5,771	13.7608	0.0081	47	
	South	27,207	14.1171	5,844	13.9349	-0.0130	-76	
	Southwest	24,598	12.7633	4,925	11.7435	-0.0833	-410	
	West	22,754	11.8065	4,600	10.9686	-0.0736	-339	
	Northwest	23,984	12.4447	5,016	11.9605	-0.0397	-199	
Curvature	-6.7778-(-)0.5312	41,911	21.7466	8,731	20.8188	-0.0436	-381	1
	-0.5312-0.4449	109,325	56.7262	24,574	58.5960	0.0324	797	
	0.4449-5.6667	41,488	21.5272	8,633	20.5851	-0.0447	-386	
Topographic Wetness Index (TWI)	3.5241-6.6126	66,549	34.5307	11,639	27.7529	-0.2185	-2,543	11.35
	6.6126-8.3604	54,317	28.1838	12,026	28.6757	0.0173	208	
	8.3604-10.6028	28,472	14.7735	6,490	15.4752	0.0464	301	
	10.6028-12.9423	29,072	15.0848	7,742	18.4606	0.2020	1,564	
	12.9423-16.4148	11,321	5.8742	3,244	7.7352	0.2752	893	
	16.4148-23.8072	2,993	1.5530	797	1.9004	0.2019	161	
Stream Power Index (SPI)	-6.1160-(-)1.6228	2,549	1.3226	103	0.2456	-1.6837	-173	8.85
	-1.6228-(-)0.6977	19,349	10.0397	3,441	8.2050	-0.2018	-694	
	-0.6977-(-)0.1889	4,7245	24.5143	10,627	25.3398	0.0331	352	
	-0.1889-0.1918	91,143	47.2920	21,136	50.3982	0.0636	1,345	
	0.1918-0.6709	28,111	14.5861	6,003	14.3140	-0.0188	-113	
	0.6709-5.6373	4,327	2.2452	628	1.4974	-0.4050	-254	

Table 1. Calculation of weight values of SI and WF models (cont.)

Conditioning factors	Conditioning classes	No. pixels	Percentage of area	No. of flood pixels	Percentage of flood	SI(W_{ij})	TSI	WF
Rainfall (mm)	1,233-1,267	18,887	9.8000	4,152	9.9003	0.0102	42	1.27
	1,267-1,285	25,591	13.2786	5,501	13.1170	-0.0122	-67	
	1,285-1,301	29,372	15.2404	6,946	16.5625	0.0832	578	
	1,301-1,317	34,939	18.1290	6,864	16.3670	-0.1022	-702	
	1,317-1,333	41,908	21.7451	9,747	23.2415	0.0665	649	
	1,333-1,353	42,027	21.8068	8,728	20.8117	-0.0467	-408	
Distance from rivers (m)	0-400	101,742	52.7916	23,400	55.7967	0.0554	1,295	13.82
	400-800	43,241	22.4367	9,238	22.0278	-0.0184	-170	
	800-1,200	23,561	12.2253	5,127	12.2252	0.0000	0	
	1,200-1,600	13,067	6.7802	2,895	6.9030	0.0180	52	
	1,600-2,000	6,466	3.3551	1,162	2.7708	-0.1913	-222	
	>2,000	4,647	2.4112	116	0.2766	-2.1653	-251	
Stream density (km/km^2)	0-0.8490	24,481	12.7026	6,208	14.8028	0.1530	950	5.66
	0.8490-1.4270	44,644	23.1647	8,726	20.8069	-0.1073	-937	
	1.4270-1.9870	48,947	25.3975	9,299	22.1732	-0.1358	-1,262	
	1.9870-2.6192	38,946	20.2082	9,571	22.8218	0.1216	1,164	
	2.6192-3.3417	22,452	11.6498	4,847	11.5575	-0.0080	-39	
	3.3417-4.6242	13,254	6.8772	3,287	7.8378	0.1307	430	
Soil drainage	Well drained	40,767	21.1530	4,632	11.0449	-0.6498	-3,010	82.61
	Moderately well drained	8,301	4.3072	4	0.0095	-6.1128	-24	
	Somewhat poorly drained	32,827	17.0332	9,324	22.2328	0.2664	2,484	
	Poorly drained	81,629	42.3554	22,746	54.2372	0.2473	5,624	
	No survey/built up area	29,200	15.1512	5,232	12.4756	-0.1943	-1,017	
Land use	Forest land	18,150	9.4176	592	1.4116	-1.8979	-1,124	100
	Paddy field/ field crop	48,414	25.1209	18,062	43.0683	0.5391	9,737	
	Orchard	64,831	33.6393	10,908	26.0098	-0.2572	-2,806	
	Other agricultural land	741	0.3845	120	0.2861	-0.2954	-35	
	Urban and built-up land	46,179	23.9612	9,502	22.6573	-0.0560	-532	
	Water body	4,191	2.1746	646	1.5404	-0.3448	-223	
	Miscellaneous land	10,218	5.3019	2,108	5.0265	-0.0533	-112	
Road density (km/km^2)	0-1.7395	18,578	9.6397	1,025	2.4441	-1.3722	-1,407	39.28
	1.7395-3.8268	35,663	18.5047	9,959	23.7470	0.2494	2,484	
	3.8268-5.4503	57,787	29.9843	13,429	32.0211	0.0657	883	
	5.4503-7.1898	43,113	22.3703	8,985	21.4245	-0.0432	-388	
	7.1898-9.2771	27,452	14.2442	6,344	15.1271	0.0601	382	
	9.2771-14.8434	10,131	5.2567	2,196	5.2363	-0.0039	-9	

The FSI from the WF model was calculated using both SI and WF weights, as shown in equation (7), and was categorized into five classes using the geometrical interval classifier method (Figure 5 (b)). Land classified as very highly susceptible occupies 51.74% of the study area, followed by 26.18%, 18.54%, 2.87%, and 0.66% of the area classified as high, moderate, low, and very low susceptibility, respectively. The WF model yields a greater area classified as very high or high susceptibility, accounting for 78.92% of the study area, compared to 72.84% according to the SI model (Table 1).

As Figures 5 (a) and (b) show, low and very low susceptibility areas are mainly in areas of high elevation. However, the very high and high susceptibility areas from the SI and WF models are

partly different. The high susceptibility areas of the WF map cover a greater extent than those in the SI map and mainly dominate the southeast of the study area located in low elevations with poor drainage conditions. The results indicate that the weighting values of the conditioning factors play an important role in obtaining the flood susceptibility mapping while the SI model relies on a calculation with the equal weight of the conditioning factors (Khosravi et al., 2016).

$$FSI_{WF} = (SI_{elevation} \times 75.77) + (SI_{slope} \times 26.41) + (SI_{aspect} \times 1) + (SI_{curvature} \times 1) + (SI_{TWI} \times 11.35) + (SI_{SPI} \times 8.85) + (SI_{rainfall} \times 1.27) + (SI_{dist.\ from\ river} \times 13.82) + (SI_{stream\ density} \times 5.66) + (SI_{soil\ drainage} \times 82.61) + (SI_{land\ use} \times 100) + (SI_{road\ density} \times 39.28) \quad (7)$$

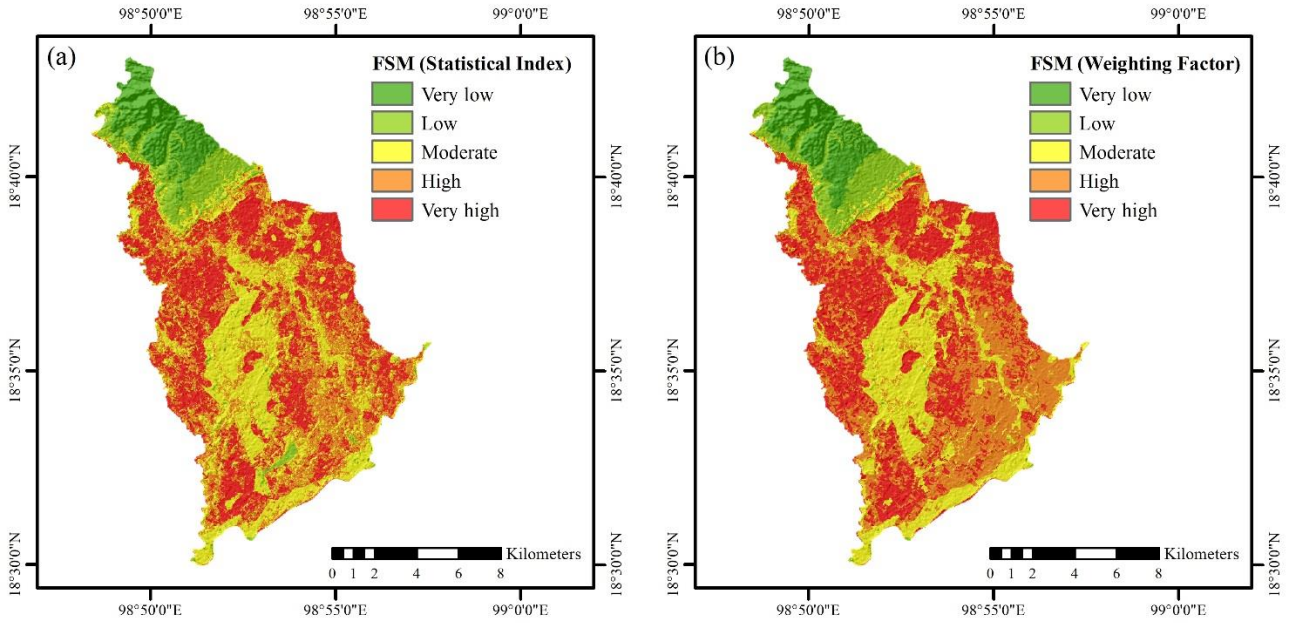

Figure 5. Flood susceptibility mapping: (a) SI model, (b) WF model

3.3 Validation of the models

The results from the SI and WF models were compared with the training data (70%) and testing data (30%) using the ROC curve and the AUC value methods. Figures 6 (a) and (b) show the ROC curves of the success and prediction rates of both results. The success rates of the SI and WF models were 91.80% and 93.06%, respectively. The prediction rate is widely used to clarify the predictability of future flood occurrences (Paul et al., 2019); the prediction rate values were 93.53% for the WF model and 92.05% for the SI model. Therefore, the results show that the WF model performed slightly better than the SI model for mapping flood-susceptible areas in San Pa Tong. However, both models provided excellent outcomes, with AUC values higher than 90% (Khosravi et al., 2016; Yesilnacar, 2005). A similar pattern of results was obtained by Khosravi et al. (2016), who found excellent results with SI and WF, and Cao et al. (2016), who also obtained a high prediction rate result using the SI model.

Additionally, the recurring flood data were utilized and compared with the results from SI and WF models. Figures 7 (a) and (b) show an excellent correlation between the recurring flood areas and the results from both SI and WF models. Comparatively, the WF model performed more accurately with 82.12% of the recurring flood areas falling into the very high susceptibility class, compared to 77.06% from the SI model (Table 2). The results also revealed that no recurring flood area was in the very low susceptibility class for either the SI or WI model. Based on the validation results, both models can be considered effective approaches for mapping flood susceptibility in other geographic areas.

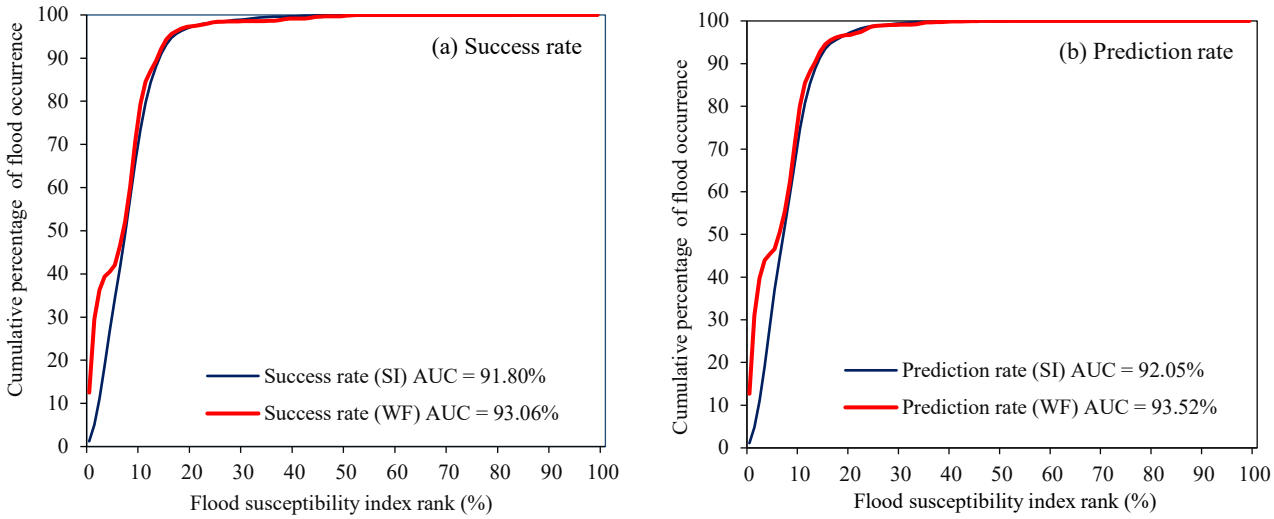

Figure 6. Validation using ROC and AUC: (a) success rate and (b) prediction rate

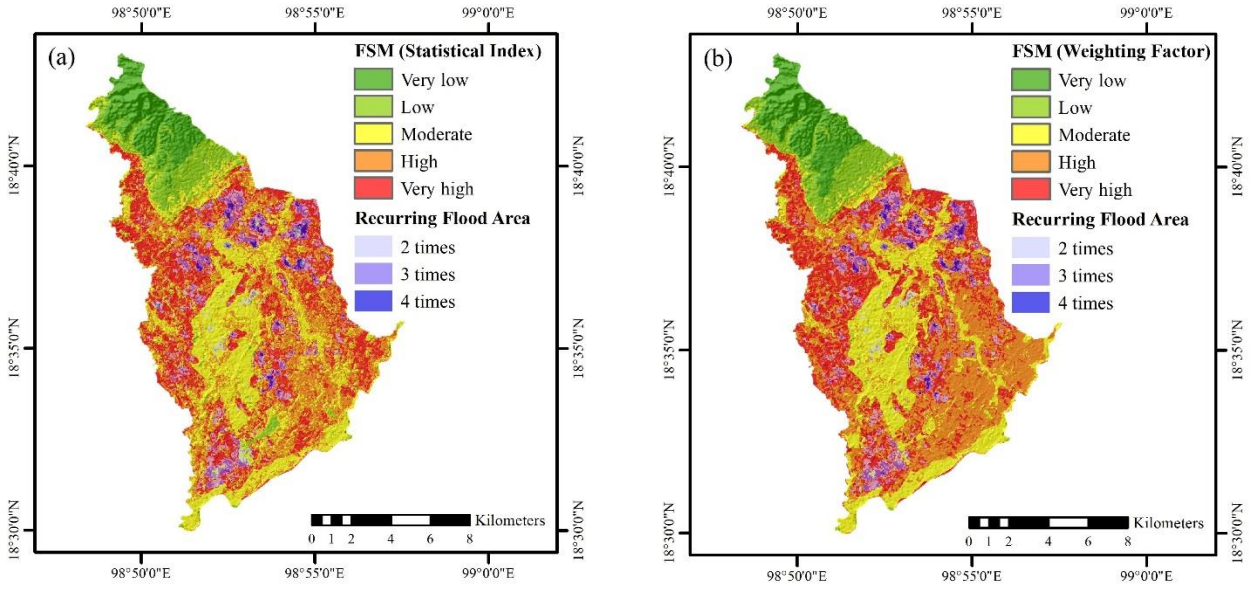

Figure 7. Flood susceptibility mapping using SI (a) and WF (b) models and recurring flood areas from 2005 to 2019

Table 2. Computation of the recurring flood area and flood susceptibility class

Model	Susceptibility class	Recurring flood area (km^2)				Percentage (%)
		2 times	3 times	4 times	Total	
SI	Very low	0.00	0.00	0.00	0.00	0.00
	Low	0.11	0.01	0.00	0.12	0.70
	Moderate	1.42	0.55	0.08	2.06	11.93
	High	1.40	0.32	0.06	1.78	10.30
	Very High	8.18	4.05	1.06	13.29	77.06
	Total	11.11	4.94	1.20	17.24	100.00
WF	Very low	0.00	0.00	0.00	0.00	0.00
	Low	0.06	0.01	0.00	0.07	0.39
	Moderate	0.96	0.17	0.01	1.14	6.62
	High	1.43	0.36	0.08	1.87	10.87
	Very High	8.65	4.40	1.10	14.16	82.12
	Total	11.11	4.94	1.20	17.24	100.00

3.4 Suggestion for further research

In terms of the results of this study, each conditioning factor has a different impact on flood occurrence, depending on the geographic context of the area. Thus, statistical models comparing the influencing factor are highly important (Kia et al., 2012). The DEM is also an essential component of these models and plays an important role in flood hazard research (Cabrera et al., 2019); the effects of flood susceptibility mapping using a different spatial resolution of DEM data might also therefore be investigated. Although both SI and WF results showed excellent and effective outcomes, there are some limitations in this study to be considered. Socio-economic factors could be integrated into the method as demonstrated by Hoang et al. (2020) and Khaing et al. (2021). Further conditioning factors could also be investigated, such as the Normalized difference vegetation index (NDVI), lithology, and land-use changes. For further research, FSM using statistical and machine learning models such as weight of evidence, artificial neural networks, logistic regression, and support vector machines could be considered (Anucharn, 2019; Paul et al., 2019; Tehrany et al., 2017; Tehrany et al., 2019). Moreover, integrating hydrological models and GIS-based technique is also a suggested point for further study (Khaing et al., 2021; Rahmati et al., 2016).

4. CONCLUSION

Flood susceptibility mapping is an essential tool for flood preparation. Identification of susceptible areas using reliable methods can help to reduce flood damage. This research applied SI and WF models for flood susceptibility mapping in the San Pa Tong District, Chiang Mai, Thailand, and compared their performance, in addition to investigating the most influential factors for flood occurrence in the study area. Flood data were randomly divided into training and testing data. Then, 12 conditioning factors, namely elevation, slope, aspect, curvature, TWI, SPI, rainfall, distance from rivers, stream density, soil drainage, land use, and road density were used to compare with training data and calculate the correlation with flood occurrence. The results from the SI and WF models revealed that very highly susceptible areas covered an estimated 49.49% and 51.74% of the study area, respectively. Regarding the WF results, the most influential factors were land use, soil drainage, and elevation, while the aspect and curvature were the least significant factors in determining flood susceptibility. ROC and AUC were then used to evaluate the success rate and the prediction rate of the SI and WF models. The results revealed that WF shows a better success rate than the SI model, with AUC values of 93.06% and 91.80%, respectively. WF also performed better, with a prediction rate of 93.52% compared to 92.05% for SI. In summary, both WF and SI results were shown as acceptable and reliable methods, with excellent performance rates for flood susceptibility mapping. The results of this research can be used to help planners implement flood preparedness and to minimize the impacts of future floods.

ACKNOWLEDGEMENTS

The author would like to express gratitude to GISTDA, the Upper Northern Region Irrigation Hydrology Center, and the Land Development Department for supporting the collection of the research data. The author is grateful to the Department of Geography, Faculty of Humanities and Social Sciences, Chiang Mai Rajabhat University for supporting the research facilities. Lastly, the author gratefully acknowledges the anonymous reviewers and the editor for their constructive comments and valuable suggestions which have helped to improve the quality of this manuscript.

REFERENCES

Anucharn T. A comparison the most appropriate method for flood susceptibility map in Khlong Nathawi Subwatershed, Songkhla Province. Journal of King Mongkut's University of Technology North Bangkok 2019;29(4):612-29 (in Thai).

Budha PB, Rai P, Katel P, Khadka A. Landslide hazard mapping in Panchase Mountain of Central Nepal. Environment and Natural Resources Journal 2020;18(4):387-99.

Cabrera JS, Lee HS. Flood risk assessment for Davao Oriental in the Philippines using geographic information system-based multi-criteria analysis and the maximum entropy model. Journal of Flood Risk Management 2020;13(2):e12607.

Cao C, Xu P, Wang Y, Chen J, Zheng L, Niu C. Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas. Sustainability 2016;8(9):948.

Center for Research on the Epidemiology of Disasters (CRED). Natural disaster 2019: Now is the time to not give up [Internet]. 2020 [cited 2020 Nov 10]. Available from: https://emdat.be/natural-disasters-2019-now-time-not-give/.

Hoang DV, Tran HT, Nguyen TT. A GIS-based spatial multi-criteria approach for flash flood risk assessment in the Ngan Sau-Ngan Pho mountainous river basin, North Central of Vietnam. Environment and Natural Resources Journal 2020;18(2):110-23.

Igarashi K, Koichiro K, Tanaka N, Aranyabhaga N. Prediction of the impact of climate change and land use change on flood

discharge in the Song Khwae District, Nan Province, Thailand. Journal of Climate Change 2019;5(1):1-8.

Khaing TW, Tantanee S, Pratoomchai W, Mahavik N. Coupling flood hazard with vulnerability map for flood risk assessment: A case study of Nyaung-U Township in Myanmar. Greater Mekong Subregion Academic and Research Network International Journal 2021;15:127-38.

Khosravi K, Pourghasemi HR, Chapi K, Bahri M. 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. Environmental Monitoring and Assessment 2016;188(12): 656.

Kia MB, Pirasteh S, Pradhan B, Mahmud AR, Sulaiman WNA, Moradi A. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Environmental Earth Sciences 2012;67(1):251-64.

Kongmuang C, Tantanee S, Seejata K. Urban flood hazard map using GIS of Muang Sukhothai District, Thailand. Geographia Technica 2020;15(1):143-52.

Oztekin B, Topal T. GIS-based detachment susceptibility analyses of a cut slope in limestone, Ankara-Turkey. Environmental Geology 2005;49(1):124-32.

Paul GC, Saha S, Hembram TK. Application of the GIS-based probabilistic models for mapping the flood susceptibility in Bansloi sub-basin of Ganga-Bhagirathi River and their comparison. Remote Sensing in Earth Systems Sciences 2019;2:120-46.

Pourghasemi HR, Moradi HR, Aghda SMF. Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. Natural Hazards 2013;69(1):749-79.

Rahmati O, Zeinivand H, Besharat M. Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis. Geomatics, Natural Hazards and Risk 2016;7(3): 1000-17.

Samanta S, Pal DK, Palsamanta B. Flood susceptibility analysis through remote sensing, GIS and frequency ratio model. Applied Water Science 2018;8(2):66.

Suppawimut W. Spatial analysis of recurrent flood in San Pa Tong District, Chiang Mai Province. Proceedings of the 6th Conference on Research and Creative Innovations; 2020 Sep 2-3; Rajamangala University of Technology Lanna, Chiang Mai: Thailand; 2020 (in Thai).

Tehrany MS, Shabani F, Jebur MN, Hong H, Chen W, Xie X. GIS-based spatial prediction of flood prone areas using standalone frequency ratio, logistic regression, weight of evidence and their ensemble techniques. Geomatics, Natural Hazards and Risk 2017;8(2):1538-61.

Tehrany MS, Kumar L, Jebur MN, Shabani F. Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. Geomatics, Natural Hazards and Risk 2019;10(1):79-101.

World Bank. Thai Flood 2011, Rapid Assessment for Resilient Recovery and Reconstruction Planning. World Bank; 2012.

van Westen CJ, Rengers N, Terlien MTJ, Soeters R. Prediction of the occurrence of slope instability phenomenal through GIS-based hazard zonation. Geologische Rundschau 1997; 86:404-14.

Yalcin A. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): Comparisons of results and confirmations. Catena 2008;72(1):1-12.

Yesilnacar EK. The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey [dissertation]. Melbourne, Australia: University of Melbourne; 2005.