ASSESSMENT OF FLOOD HAZARD USING GEOSPATIAL DATA AND FREQUENCY RATIO MODEL IN SUKHOTHAI PROVINCE, THAILAND

Kamonchat Seejata (1*), Aphittha Yodying (1), Sasithon Chatsudarat (2), Polpreecha Chidburee (2), Nattapon Mahavik (2), Charatdao Kongmuang (2) , Sarintip Tantanee (1)

[1] Department of Civil Engineering, Naresuan University, Phitsanulok, Thailand, 65000.
[2] Department of Natural Resources and Environment, Naresuan University, Phitsanulok, Thailand, 65000.
Email: kamonchatseejata@gmail.com

KEYWORDS: Geographic Information System, Frequency Ratio, Area Under the Curve, Sukhothai.

ABSTRACT: Flood is one of the most serious disasters. It frequently occurs in Thailand, Especially in, Sukhothai Province is one of the flood-prone areas in the Yom River basin. Due to the physical geography of the basin and lacking the regulated dam, the heavy rainfall over an upstream area causes a large amount of water flow downstream through the narrow river. Sukhothai regularly suffers from flood due to water overflowing and the high erosion along the riverbanks. Therefore, this research aims to develop a flood hazard map by using Remote Sensing (RS), the Geographic Information System (GIS), and Frequency Ratio (FR) method. In addition, the Area Under the Curve (AUC) and the Seed Cell Area Index (SCAI) methods were used to assess the accuracy and efficiency of this method. In this research, the flood extent map was generated from The Geo-Informatics and Space Technology Development Agency (GISTDA) that were obtained from flood inventory derived from 2005 to 2017 over the study area. The developing process of flood extent map was divided into two steps. Firstly, 70% of all flood inventory area has been used as the training dataset to generate flood hazard maps. Secondly, the remaining of 30% of all flood inventory has been applied as testing dataset to validate the map. To assess the flood hazards, eight parameters namely rainfall, elevation, slope, soil drainage, land use, drainage density, road density and the distances from the drainage were considered. Each parameter was classified using the Jenks natural breaks classification method. Then, the results indicated that 24.22%, 24.07%, 24.81%, 15.64% and 11.26% of the study area are very low, low, moderate, high, and very high flood hazard class, respectively. The validation result reveals that frequency ration method was found to be effective with 64.94% accuracy in the AUC success rate of assessing the flood hazard. The research output can then provide useful information for the local authorities, planners, government and related organizations to manage flood risk and improve flood response plans in the future.

1. INTRODUCTION

Flooding is one of the most serious and frequent disasters in Thailand. It occurs almost every year during the monsoon season and that causes a lot damage. Every part of Thailand struggles with flood-related damages annually (CFE-DM, 2018). The main causes of floods in Thailand are the influence of weather phenomena and heavy rain (Department of Disaster Prevention and Mitigation, 2015). The floods have been exacerbated and more severe and devastating as well as inflicting heavier tolls of losses, such as, floods of 2011 by heavy rain from monsoons in July. These impacted a total of 4,039,459 households and 13,425,869 people, as well as 2,329 houses, were destroyed, 96,833 houses were partially damaged; around 657 people died, and three were reported missing. According to the Word Bank damage worth of around THB 1,440 billion were estimated (Thai Water, 2012). Sukhothai province is located in one of the flood-prone areas. Due to geographical characteristics of the upstream basin in North of Sukhothai Yom, river flows through a narrowed path receives heavy rainfall leading to lower valleys. The amount of water flowing into the province of Sukhothai has a large volume that causes an overflow of the riverbanks, erosion, damage, high and severe flooding (Royal Irrigation Department, 2009). Flood hazard assessment is essential tool understand the effect of floods. The flood situation can be analyzed by combining flood hazard assessment with Geographic Information System to present mapped information defining flood-prone areas (Organization of American States, 1991). Therefore, this research aims to develop a flood hazard map with remote sensing and the geographic information system that are handy tools for disaster management, and flood hazard assessment (Cao et al., 2016; Tehrany et al., 2018). For flood hazard analysis, frequency ratio method is used to estimate flood probability and generate flood hazard map with geographic information system technique (Anucharn and Iamchuen, 2017; Samanta et al., 2018; Youssef et al., 2015). Furthermore, the Area Under the Curve (AUC) and the Seed Cell Area Index (SCAI) were used to assess the accuracy and efficiency of this method

2. STUDY AREA

Sukhothai province is located in a Yom river basin of Thailand, covering an area of is 6,663 km^2. It contains nine districts (or Amphoe) including Mueang Sukhothai, Ban Dan Lan Hoi, Khiri Mat, Kong Krailat, Si Satchanalai, Si Samrong, Sawankhalok, Si Nakhon and Thung Saliam (Figure 1). The topography has mostly featured the basin plains in the north with highlands and mountains stretching in the west, central plains and southern high lands. The Yom River flows from the north to the south through Sri Satchanalai, Sawankhalok, Sri Samrong, Mueang Sukhothai and Kong Krailat. The highest Luang hill is located about 1,185 meters above mean sea level (m. MSL). The Yom River is one of the important river for agriculture subsistence in Thailand. The feature of Yom River has a high slope especially, upstream of the river. Consequently, it can cause problems during the rainy season which the exceed water flow southward rapidly, causing flooding in the lower basin plain. On the other hand, there is less water during summer season, causing lack of water for agriculture (Sukhothai Provincial Office, 2018).

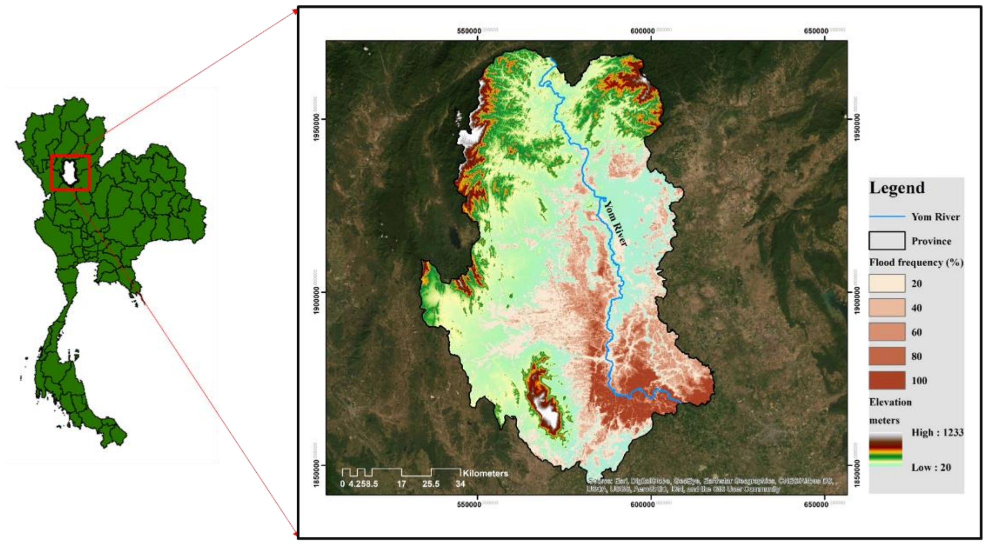


Figure 1 Study area location and flood area occurred.

The climate in Thailand is generally influenced by the southwest monsoons in the rainy season and the northeast monsoons in winter season whereas summer season is period of the change in monsoons from northeast to southwest. The average temperature is 27.6 degree Celsius. The average highest and lowest temperature is 37.7 and 18.5 degree Celsius, respectively. April is the hottest month while January is known to be coldest. The average annual rainfall is 1,144.95 milliliters. The highest rainfall occur during September at around 267.9 milliliters, and the least rainfall is in November at around 5.4 milliliters (Land Development Department, 2015).

3. GEOSPATIAL DATA AND METHOD

To develop a flood hazard map, this research used the Frequency Ratio method. The methodological framework is shown in Figure 2.

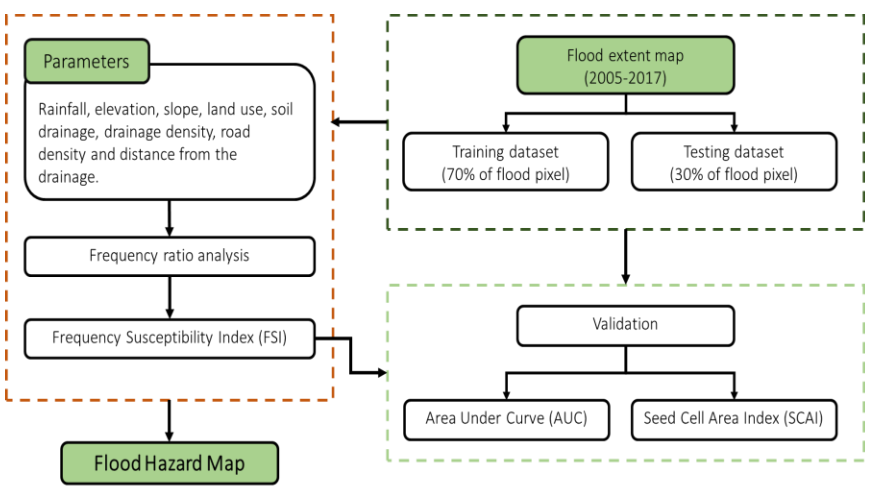
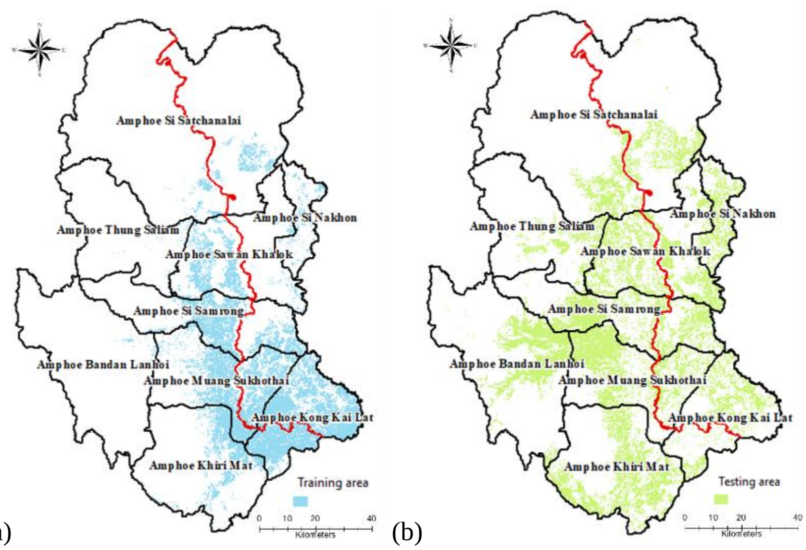


Figure 2 Methodological framework for flood hazard mapping of Sukhothai province, Thailand.

3.1 Flood extent map
The flood extent map was generated from The Geo-Informatics and Space Technology Development Agency (GISTDA) that were obtained from flood inventory derived from 2005 to 2017 over the study area. It was divided into two parts that 70% of all flood inventory area will be used as the training dataset to generate flood hazard map and the remaining of 30% will be used as testing dataset for validation (Figure 3). Subset feature in Geostatistical analyst tool is used to create a random dataset over the area.



(a) (b)

Figure 3 Flood extent location divided into parts (a) Training area and (b) Testing area.

3.2 Flood conditioning parameters
In order to assess flood hazard, eight conditioning parameters were selected and prepared based on literature reviews and data availability. These parameters were considered by rainfall, elevation, slope, land use, soil drainage, drainage density, road density, and the distance from the drainage. Each conditional parameter was analyzed in the GIS environment. The cell size of the raster for these parameters is 30 meters which were divided into classes using the natural breaks (Jenks) classification method (Figure 4a-h).

Rainfall data (1988-2017) is a vital importance parameter in the occurrence of flood. A flood occurs during the monsoon and season including the amount of flood derived from upstream. In this research, obtained data from Thai Meteorological Department (TMD) contained 19 meteorological stations (12 stations in Sukhothai province and 7 stations in the nearby province). The Inverse Distance Weighted (IDW) interpolation method was also used in calculation process.

Low elevation and less slope area will be more floods occurrence. Hence, elevation, and slope were obtained from the Digital Elevation Model (DEM) of Shuttle Radar Topography Mission (SRTM) with 30 x 30-meter spatial resolution. DEM tiles were downloaded from the official United States of Geological Survey (USGS) website (https://earthexplorer.usgs.gov/). The slope was created from the DEM using ArcMap 10.2 software.

Land-use changes causes a risk of flooding, for instance, the well-forested area is less likely to experience a flood due to water infiltration, evapotranspiration, and runoff (Cao et al., 2016; Duangpiboon et al., 2018). Therefore, it was classified into five classes according to Land Developed Department; Agriculture land (A), Forest land (F), Urban land (U), Water land (W), Miscellaneous land (M).

Soils with good drainage can help reduce flooding. It was classified into seven classes as adopted from Land Developed Department; No survey, poorly to somewhat poorly drained, well to moderately well-drained, very well-drained, urban area, miscellaneous area and water area.

Drainage and road density was calculated by the ratio of the length within the cell to the cell size. The drainage density area that has more stream drainage might be more flood occurs. The road density, acts as obstruction along the drainage system and leads to flooding.

The buffer distance from the drainage is also considered, the area that located nearby waterway might be affected by overflowing riverbank during heavy rainfall. Hence, five concentric buffers, each of 1,000-meter width, were demarcated around each stream to generate the map.

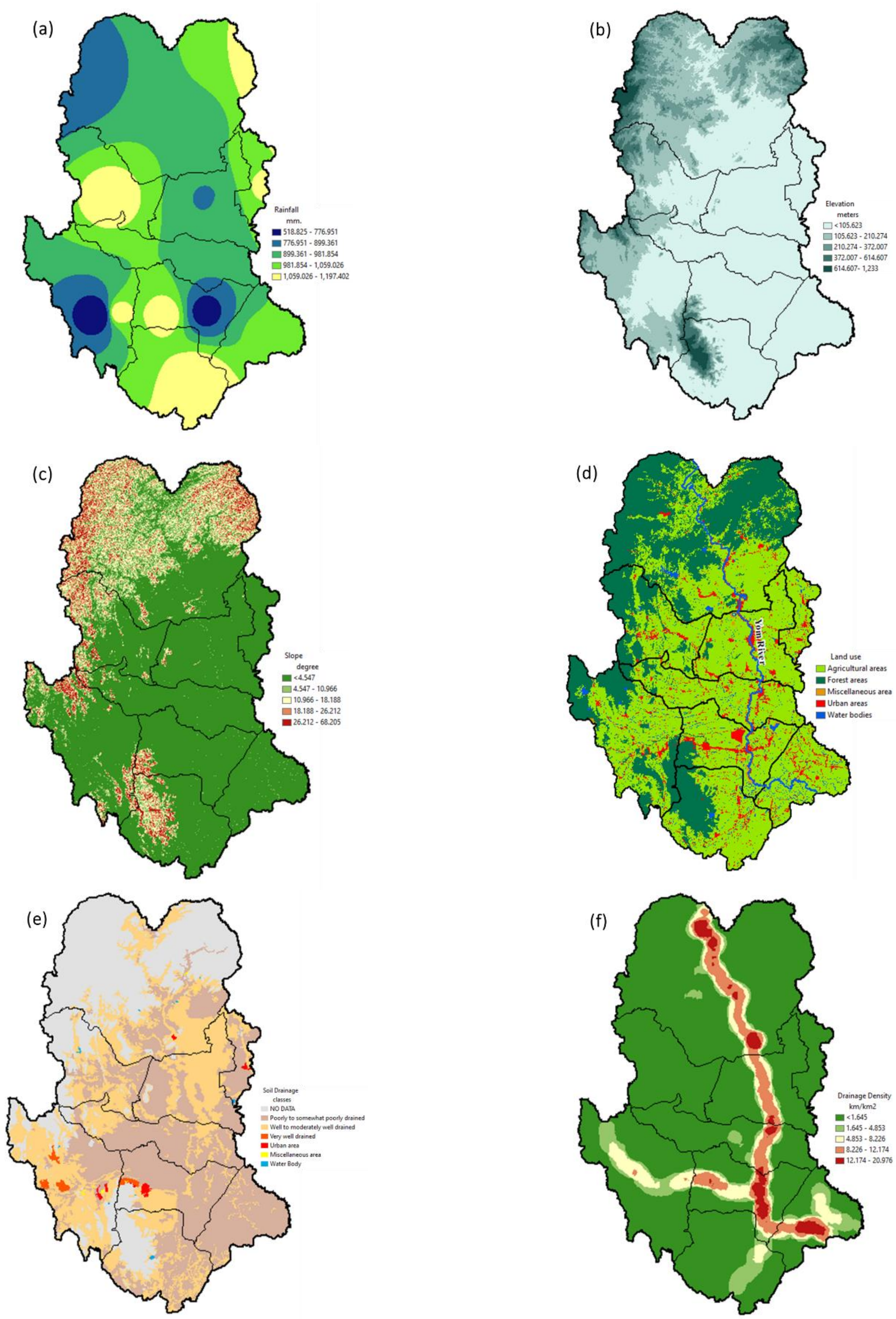


Figure 4 Flood conditioning parameters namely (a) average rainfall, (b) elevation, (c) slope, (d) land use, (e) soil drainage, (f) drainage density, (g) road density, and (h) the distances from the drainage.

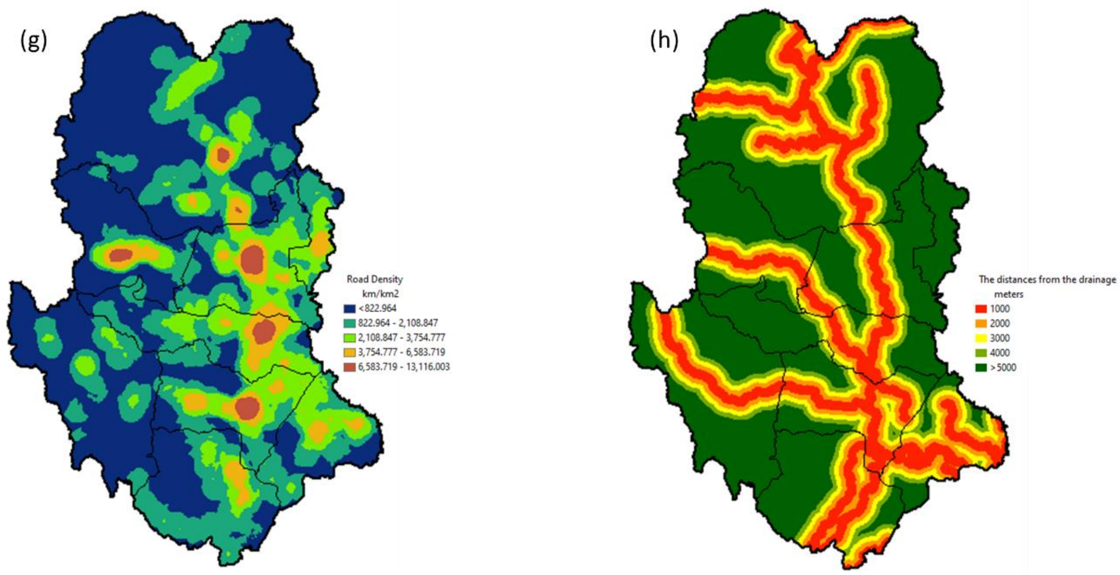


Figure 4 (Continue).

3.3 Frequency Ratio method

The Frequency Ratio (FR) is a statistical analysis method, based on the correlation between the spatial distribution of parameters and flood inventory (AltThuwaynee, 2018). This method is very popular for assessing landslide and flood hazard assessment. Generally, the ratio is defined by the area where flood occurrences are found in the total study area. In order to developed flood hazard map, the FR values were calculated using the following equation (1).

$$FR = \frac{PH}{PS} \tag{1}$$

Where PH is the percentage of flood hazard in each class and PS is the percentage of the study area in each class. In this analysis, If FR value lower than 1 indicates weak correlation, on the other hand, the value of FR more than 1 indicates strong correlation. The frequency susceptibility index was calculated using the following equation (2).

$$FSI = \sum_{i=1}^{n} FR_i \tag{2}$$

Where FR_i is the value of FR in each parameter and n is the number of parameters. Finally, the FSI value that was applied for flood hazard mapping and classified by Natural Breaks (Jenks) into five classes, which are very high, high, moderate, low, and very low.

3.4 Model Validation

The Area Under the Curve (AUC) method was used to validate the results obtained from the flood hazard map regarding the success rate and prediction accuracy of the model. The success rate was calculated using the training flood area and the prediction accuracy was calculated using a testing flood area. It started with arranging the flood susceptibility index. Then, the arranged flood susceptibility index was classified into 100 categories on the x-axis, with a cumulative of flood occurrence on the y-axis (Tehrany et al., 2019). The model is successful if the range of AUC close to 1, the model is considered to be more reliable, while a value below 0.5 indicates that the model cannot efficiently predict flood. In the Seed Cell Area Index (SCAI) method, flood hazard zones area was divided by the percentage of the flood of the training and testing dataset as follow equation (Suzan and Doyaran, 2004).

$$SCAI = \frac{Area\ extent\ of\ flood\ hazard\ classes\ (\%)}{Flood\ hazard\ zone\ of\ the\ training\ and\ testing\ set\ in\ each\ hazard\ class\ (\%)} \tag{3}$$

If the SCAI values were lower for the indices of the model, then high flood hazard could be expected. On the other hand, if SCAI values were higher, progressively lower flood hazard would be denoted (Shahana and petal, 2019; Liuzzo et al., 2019).

4. RESULT AND DISCUSSION

4.1 Flood Hazard using frequency ratio

The results revealed that the average rainfall, class of average rainfall of 518.825- 776.951 mm. has the highest FR value. Consequently, higher rainfall than this class does not affect the flood hazard. Due to the highest rainfall often occurs at the highest elevation where areas are not flooding occur, agreeing with the previous work in the literature review (Cao et al., 2016; Liuzzo et al., 2019). The most floods occurred at elevation lower than 105.623 m. MSL, of which FR has the highest values, whereas, for elevation higher than 105.623 m. MSL, the FR is equal to 0. For the slope parameter, the highest FR values occur at lower than 4.547 degrees, while all the other classes provide very low FR values. Flooding was not possible on a slope higher than 10 degrees. Agriculture and miscellaneous area are most flooding occurrence (FR = 1.54 and 1.56, respectively), while FR value in forest area is equal to 0 that is least effect floods occur. Soil drainage types of poorly to somewhat poorly drain have the highest FR value of 2.52 causing floods to occur. As for the drainage density parameter, a class of 1.645-4.853 km/km^2 has the highest FR value (2.27), whereas the class of less than 1.645 has a lower FR value. For the road density, a class of 2,108.847-3,754.777 km/km^2 has the highest FR value (1.83), whereas the class of less than 822.964 km/km^2 has the lower FR value. In terms of the distances from the drainages of less than 1,000 and 2,000 meters might get the effect of floods, whereas distance more than 5,000 meters will be least affected (Table 1).

Based on equation (2), The FSI values are calculated for each cell of the grid. The obtained, FSI ranges between 41.729 and 170.329. Hence, FSI values are related to a flood hazard, which lower values are attributed to the location where floods are less to occur. On the other hand, for the higher values, floods are more likely to occur. The FSI values were classified into five classes namely very low (41.729- 67.449), low (67.449 - 86.109), moderate (86.109 - 104.769), high (104.769 - 128.471), and very high (128.471 - 170.329) as shown in Figure 5. The flood hazard map, then, can be obtained from the frequency ratio method.

Table 1 Frequency ratio analysis of flood conditioning parameters

Parameter	Classes	Number of total pixels in this study		Number of flood occurrence pixels		FR value
		Pixels number	Percentage (%)	Pixels number	Percentage (%)	
Rainfall (mm.)	518.825 - 776.951	182530	2.46	33362	3.03	1.23
	776.951 - 899.361	951679	12.85	117948	10.70	0.83
	899.361 - 981.854	2960604	39.98	469780	42.61	1.07
	981.854 - 1,059.026	2217474	29.95	358740	32.54	1.09
	1,059.026 - 1,197.402	1092704	14.76	122640	11.12	0.75
Elevation (m.)	<105.623	4481626	60.53	1102369	100.00	1.65
	105.623 - 210.274	1657147	22.38	7	0.00	0.00
	210.274 - 372.007	791362	10.69	10	0.00	0.00
	372.007 - 614.607	350531	4.73	0	0.00	0.00
	614.607 - 1,233	123005	1.66	0	0.00	0.00
Slope (degree)	<4.547	5104010	68.94	1096624	99.48	1.44
	4.547 - 10.966	798416	10.78	5632	0.51	0.05
	10.966 - 18.188	757663	10.23	87	0.01	0.00
	18.188 - 26.212	528531	7.14	33	0.00	0.00
	26.212 - 68.205	215051	2.90	10	0.00	0.00
Land Use	Urban	382153	5.16	33155	3.01	0.58
	Agriculture	4402429	59.45	1011659	91.76	1.54
	Forest	2359723	31.87	447	0.04	0.00
	Water	175384	2.37	37433	3.40	1.43
	Miscellaneous	85274	1.15	19776	1.79	1.56

Table 1 (Continue).

Parameter	Classes	Number of total pixels in this study		Number of flood occurrence pixels		FR value
		Pixels number	Percentage (%)	Pixels number	Percentage (%)	
Soil Drainage	No Data	2280462	30.80	1254	0.11	0.00
	Well to moderately well drained	2320392	31.34	74100	6.72	0.21
	Poorly to somewhat poorly drained	2737477	36.97	1027097	93.16	2.52
	Very well drained	35464	0.48	0	0.00	0.00
	Urban area	20902	0.28	19	0.00	0.01
	Miscellaneous area	4202	0.06	0	0.00	0.00
	Water	6092	0.08	0	0.00	0.00
Drainage Density (km/km^2)	<1.645	5902900	79.72	671828	60.94	0.76
	1.645 - 4.853	502064	6.78	170003	15.42	2.27
	4.853 - 8.226	444658	6.00	132501	12.02	2.00
	8.226 - 12.174	404917	5.47	81434	7.39	1.35
	12.174 - 20.976	150452	2.03	46704	4.24	2.09
Road Density (km/km^2)	<822.964	3538697	47.79	278094	25.22	0.528
	822.964 - 2,108.847	2237577	30.22	423421	38.41	1.271
	2,108.847 - 3,754.777	1129826	15.26	307405	27.88	1.827
	3,754.777 - 6,583.719	392447	5.30	83099	7.54	1.422
	6,583.719 - 13,116.003	106444	1.44	10451	0.95	0.659
The distances from the drainage (m.)	1000	1162542	15.70	296183	26.87	1.71
	2000	923228	12.47	234680	21.29	1.71
	3000	807338	10.90	190442	17.27	1.58
	4000	703222	9.50	121188	10.99	1.16
	>5000	3808661	51.43	259977	23.58	0.46

According to the results (Figure 5 and Table 2), areas of high and very high hazard classes are about 15.64% and 11.26%, of the whole province, respectively. These areas are in Muang Sukhothai, Kong Kai Lat, and Khiri Mat, in particular, where located in the south of the study area. In addition, mostly lower elevation and flat slope areas are more risks of flooding, including agriculture land where is found in the vulnerable area. People who live in high and very high hazard should be aware of flood hazards. Moderate hazard class covers an area of 24.84%. Areas of very low and low hazard classes about 24.22% and 24.07% where are located in high elevation and high slope degree area of Si Satchanalai, Thung Saliam, Bandan Lanhoi, and Khiri Mat, in particular. Most of the area of very low and low hazard class are forest and somewhat agricultural area.

Table 2 flood hazard classification in Sukhothai Province.

Level	Area (km^2)	Percentage
Very low	1614	24.22
Low	1604	24.07
Moderate	1653	24.81
High	1042	15.64
Very high	750	11.26
Total	6663	100

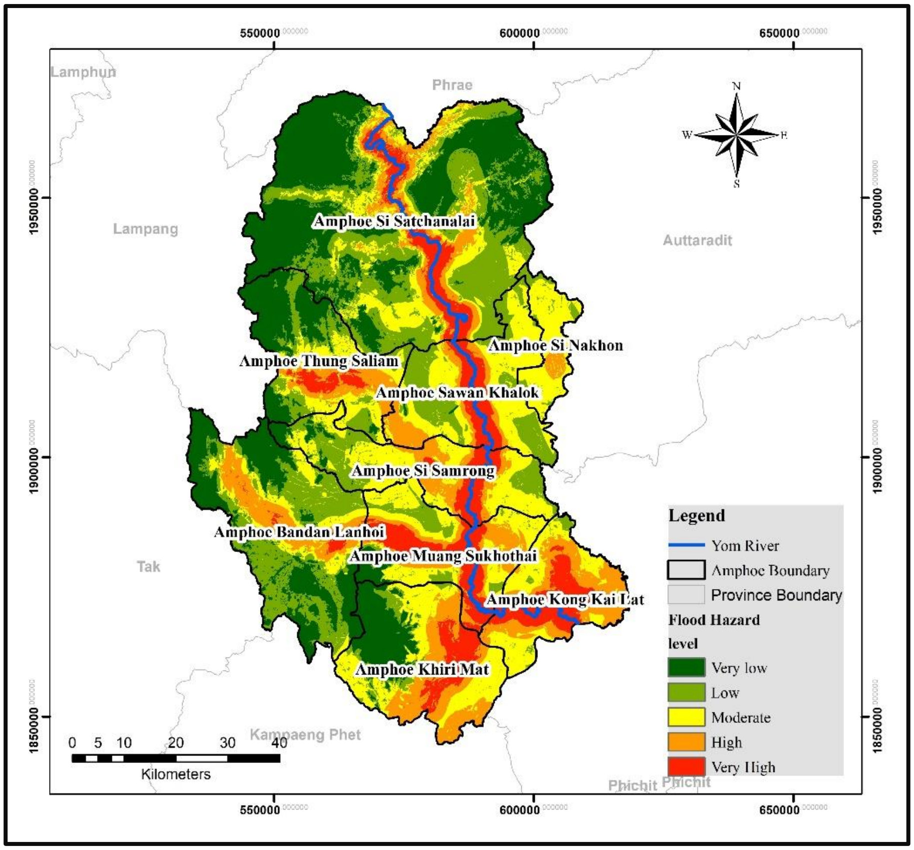


Figure 5 Flood hazard map using Frequency ratio method in Sukhothai Province.

4.2 Model validation

Flood hazard map generated through the frequency ratio model was validated by the area under the curve to examine the reliability and efficiency of the flood map. For the rank of flood susceptibility index was classified into 100 categories on the x-axis, with a cumulative of flood occurrence on the y-axis (Figure 6). The AUC of the training dataset was 64.94% for the frequency ratio model, whereas the testing was 60.87%. However, when comparing the efficiency of the model with previous studies. In this case, both the success and prediction rate curves showed the ability to predict flood hazard. Figure 8 shows both the success rate and prediction rate for the frequency model.

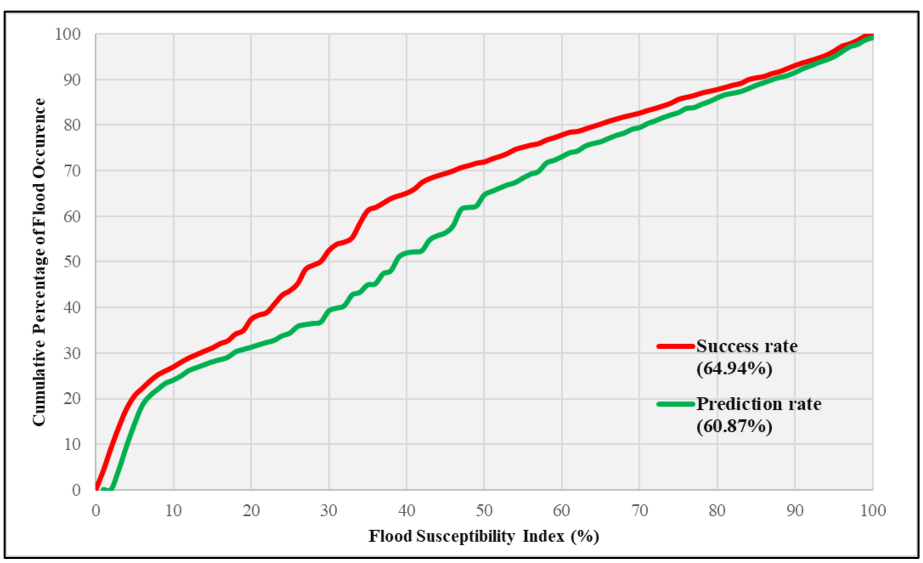


Figure 6 Success rate and prediction rate curve for the model derived from the frequency ratio.

For the seed cell area index method, flood hazard map generated through the FR model shows sufficient accuracy since the values of SCAI are high and very high for low and very low flood hazard classes, whereas the SCAI values are low and very low for high and very high hazard classes as shown in table 3.

Table 3 Seed cell area index values of the flood hazard zones.

Flood Hazard Zones	Percentage of flood hazard area	Percentage of training and testing data set	SCAI
Very low	24.22	0.33	74.26
Low	24.07	18.04	1.33
Moderate	24.81	36.24	0.68
High	15.64	24.86	0.63
Very high	11.26	20.54	0.55

5. CONCLUSION

A flood hazard map is an essential tool for identifying areas most at risk of flooding. It helps not only mitigation and flood risk management but also assisting the local authorities to manage the flood. This study aimed at developing a flood hazard map through utilizing remote sensing, the geographic information system, and frequency ratio (FR) method. In addition, the area under the curve (AUC) and the seed cell area index (SCAI) methods were used in validation process. In analysis using Frequency Ratio model, the flood extent map divided into two parts that are 70% of all flood inventory area was used as the training dataset. The remaining of 30% was used as testing dataset which applied to validate with combining layers in a GIS environment. The main eight parameters (e.g. rainfall, elevation, slope, soil drainage, land use, drainage density, road density and the distances from the drainage) were considered to assess the flood hazards. Consequently, the values of flood hazard map were classified into five classes using the Jenks natural breaks classification method. These zone areas were distributed over an area of 1,614 km^2, 1,604 km^2, 1,653 km^2, 1,042 km^2, and 750 km^2 for very low, low, moderate, high, and very high zone, respectively. The area under the curve (AUC) was used as validation method. AUC result of FR model was 64.94% for success rate, whereas the prediction rate was 60.87%. Therefore, the applied model can effectively help to develop flood hazard maps. Moreover, the results of this research can provide useful information for helping the local authorities, and related organizations to manage flood risk and improve flood response plans in the future to reduce potential damage and losses due to floods.

6. ACKNOWLEDGEMENTS

This research was supported by “Advancing Co-design of Integrated Strategies with Adaptation to Climate Change in Thailand (ADAP-T)” (Grant Number: JPMJSA1502) supported by the Science and Technology Research Partnership for Sustainable Development (SATREPS), JST-JICA. Including thanks to the Thai Meteorological Department, Land Develop Department, Geo-Informatics and Space Technology Development Agency (GISTDA) for providing various datasets for this research.

7. REFERENCES

Aditian, A., Kubota, T., & Shinohara, Y. (2018). Comparision of GIS-based landslide susceptibility models using frequency ratio, logistic, and artificial neural network in a tertiary region of Ambon, Indonesia. Geomorphology, 318, pp. 101-111.

AltThuwaynee, O. (2018). How to Produce Prediction Map in GIS with ArcGIS and Excel, Retrieved from https://www.udemy.com/user/omar-f-althuwaynee/

Anucharn, T., & Iamchuen, N. (2017). Flood Susceptibility Map Based on Frequency Ratio Method at Songkhla Lake Basin in the Southern of Thailand. BURAPHA SCIENCE JOURNAL, 22(3), pp. 106-122.

Cao, C., Xu, P., Wang, Y., Chen, J., Zheng, L., & Niu, C. (2016). Flash Flood Hazard Susceptibility Mapping Using Frequency Ratio and Statistical Index Methods in Coalmine Subsidence Areas. Sustainability, 8(9), 948.

CFE-DM. (2018). Thailand Disaster Management Reference Handbook. Hickam, Hawaii.

Department of Disaster Prevention and Mitigation. (2015). National disaster Risk Management Plan. Retrieved from http://www.disaster.go.th/upload/download/file_attach/584115d64fcee.pdf.

Duangpiboon, S., Suteerasak, T., Rattanakom, R., & Towanlong, W. (2018). Flood Susceptibility Mapping Using Geographic Information System and Frequency Ratio Analysis in the Lang Suan Watershed, Southern Thailand. The Journal of King Mongkut's University of Technology North Bangkok, 28(2), pp. 259-272.

Land Development Department, (2015). Soil resource survey report, Sukhothai province.

Liuzzo, L., Sammartano, V., & Freni, G. (2019). Comparison between Different Distributed Methods for Flood Susceptibility Mapping. Water Resources Management, 33, pp.3155-3173.

Organization of American States, (1991). Primer on Natural Hazard Management in Integrated Regional Development Planning. Washington, D.C.

Royal Irrigation Departmen, (2009). Chapter 3 Water management for flood mitigation in each province. Retrieved from http://irrigation.rid.go.th/rid4/jone_water/data/2552/pland/25-06-52/ES03.doc.

Sahana, M., & Sajjad, H. (2017a). Evaluating effectiveness of frequency ratio, fuzzy logic and logistic regression models in assessing landslide susceptibility: a case from Rudraprayag district, India. Journal of Mountain Science, 14(11), pp. 2150-2167

Sahana, M., & Petal, P, P., (2019b). A comparison of frequency ratio and fuzzy logic models for flood susceptibility assessment of the lower Kosi River Basin in India. Environmental Earth Sciences, 78(289).

Samanta, R. K., Bhunia, G. S., Shit, P. K., & Pourghasemi, H. R. (2018). Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India. Modeling Earth Systems and Environment, 4(1), pp. 395-408.

Sukhothai Provincial Office, (2018). General information of sukhothai province. Retrieved from http://www.sukhothai.go.th/3584/data/data_01.htm

Tehrany, M, S., Kumar, L., Jebur, M, N., & Shabani, F. (2019). Evaluating the application of statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. GEOMATIC, NATURAL, HAZARD AND RISK, 10(1), pp. 79-101.

ThaiWater. (2012). 2011 Thailand Flood Executive Summary. Retrieved from http://www.thaiwater.net/web/index.php/ourworks2554/379-2011flood-%20summary.html

Youssef, A. M., Pradhan, B., & Sefry, S. A. (2015). Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models. Environmental Earth Sciences, 75(12).