PAPER • OPEN ACCESS

Frequency ratio application for mapping flood susceptibility in Welang Watershed, East Java

To cite this article: R Aprilia *et al* 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **930** 012095

ICWRDEP 2021 IOP Publishing
IOP Conf. Series: Earth and Environmental Science **930** (2021) 012095 doi:10.1088/1755-1315/930/1/012095

Frequency ratio application for mapping flood susceptibility in Welang Watershed, East Java

R Aprilia[1], E Hidayah[1,*], and D Junita K[1]

[1]University of Jember, Kalimantan St. Number 37, Jember, 68121, Indonesia

*Corresponding author: entin.teknik@unej.ac.id

Abstract. Flood is one of the disaster threats downstream of Welang river, Pasuruan. A flood susceptibility map is needed to anticipate floods disasters. This research aimed to map flood Susceptibility in the Welang watershed using a Geographical Information System. In determining flood hazard, the Frequency Ratio (FR) approach was used. Flood locations were identified from the interpretation of field survey data as training data and model validation. The data were represented in a Digital Elevation Model (DEM) map, geological data, land use, river data, and Landsat Satellite Imagery and processed into a spatial database on the GIS platform. The factors that caused flooding consisted of Flood inventory, slope, Elevation, *Topographic Wetness Index* (TWI), Standardized Precipitation Index (SPI), Flow Accumulation, Distance to the river, River Density, Rainfall, Vegetation Index (NDVI), and Landuse. The map results with acceptable accuracy showed that the FR model gained an Area Under Curve (AUC) value of 90%, and the incidence for the Area Under Curve (AUC) was 93%. It is known that 1% of the flood-prone area is very high. The local Government can use the research to minimize the risk of flooding in the Welang watershed.

keywords: frequency ratio, flood susceptibility, geomorphology, mapping

1. Introduction

Pasuruan is one of the regions of East Java province which recently is always hit by floods every year in the rainy season and also causes flash floods. Floods or flash floods lead to relevant damages to property, infrastructure, public facilities, economic activities, and services, especially when hitting urban areas with historic and productive sites, and thus affect the entire community [1]. Floods in Pasuruan are caused by the flow of Welang River that crosses Kraton District, Pohjentrek District, and Gading Rejo District.

Floods in Welang River generally occur from January to April, where floods often appear in the Welang River area. On February 22, 2018, and April 3, 2018, Welang River experienced a very high water discharge which made Welang River unable to accommodate the volume of water. It was worsened by the sea tide, resulting in an overflow that flooded houses as high as approximately 50 cm to 2 meters along the river and transport stream of Pantura road (Northern Coast of Java Island). It caused the flow of transportation was paralyzed for several days. Floods due to the overflow of Welang River are presented in Figure 1.

Figure 1. Occurrence of flooding due to overflowing of Welang River

This is very detrimental to many parties, the residents of Pasuruan and those who pass along the Pantura road. The Government's efforts to carry out rehabilitation and reconstruction after disasters are often in a rather hasty manner to meet the urgent demands of disaster victims [2]. Thus, it is necessary to have the policy take mitigation strategy steps in dealing with floods during the rain influenced by the tides of seawater to reduce the risk caused by flooding in the Welang River area. In developing a mitigation strategy, the availability of a flood hazard map is an absolute requirement that must be provided. Based on the map-making process, an assessment can be carried out to determine the factors that cause flooding.

Flooding can be caused by changes in land use due to urbanization (especially the construction of villas/hotels in the forest). Many trees are cleared, and more forests are converted to residential areas. It leads to an imbalance in the catchment hydrology [3]. In addition, several other factors are predicted as factors that trigger flood events. They are slope, elevation, aspect, profile curvature, distance to the river, *Normalized Difference Vegetation Index* (NDVI), *Topographic Wetness Index* (TWI), geology, and precipitation [4], [5].

One approach that has been widely used to model the level of flood hazard is frequency ratio (FR). Several successes of FR statistics to map flood hazards have been carried out by [4], [5]. Based on the performance measurement of *Area Under Curve* (AUC) in mapping flood hazard, the FR approach is better than the Weight of evidence [6], and FR-natural breaks remain better than the statistical index (SI)-natural breaks [7].

With the flood occurrence, an assessment is needed in identifying the risk of flood disaster. The flood disaster reduction strategy as a mitigation step is necessary to reduce the risks that arise so that the level of losses and flood victims can be minimized. This paper discusses the mapping of flood Susceptibility through weighting the factors that affect flooding using the FR method assisted by GIS.

2. Material and Methods

In this study, the Susceptibility of flooding used the frequency ratio method assisted with spatial analysis using GIS application. The following are the stages of the research methodology in figure 2; 1). Inventory data on flood events and data on each factor; 2). Each flood conditioning factor is reclassified into seven classes using the quantile schema; 3) Data on flood occurrence are separated into 70% training and 30% validation; 4). The process of calculating the distribution of occurrence using the FR method is carried out to determine the weight value of each factor according to the occurrence. A higher Frequency Ratio (FR) indicates a strong correlation between conditioning factors and flood occurrence [8].

IOP Conf. Series: Earth and Environmental Science **930** (2021) 012095 doi:10.1088/1755-1315/930/1/012095

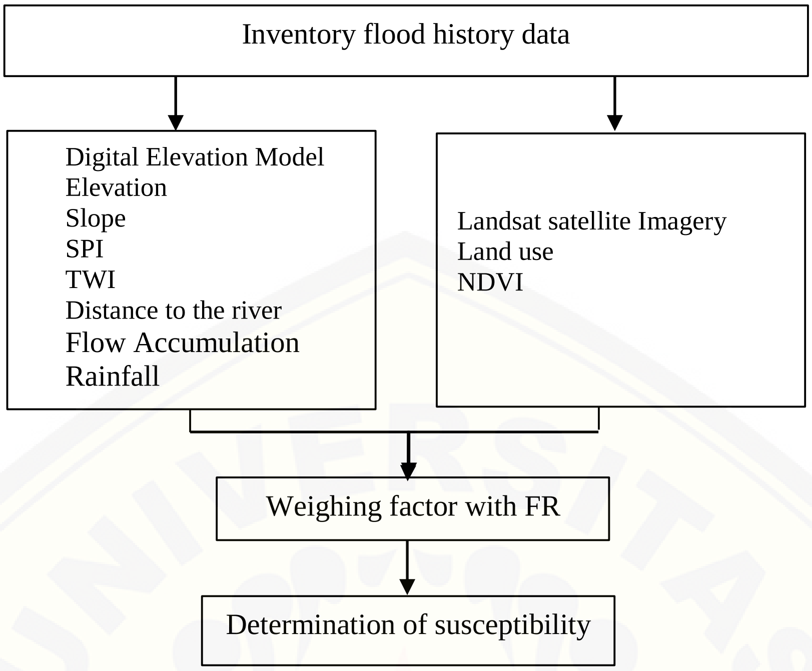


Figure 2. Research Step

2.1. *Study area and Data Requirement*

In general, the research location was in Welang watershed with 532,120 km^2, with coordinates from 112^{0}36 E to 113^{0}53 E and 7^{0}53 S to 7^{0}35 S with an altitude value of 0 - 3,233 m (figure 3).

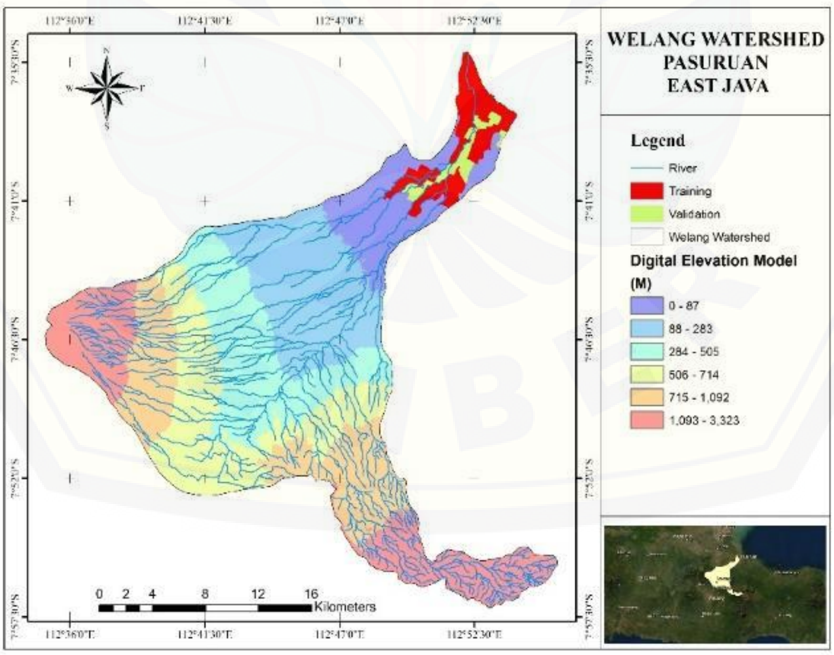


Figure 3. Map of the study area

Data for the environment spatial modeling to create a potential flood susceptibility map used SRTM DEM 30x30 m resolution. The derivative of the DEM produces a geomorphological layer in the form of the slope, TWI, SPI, and Flow accumulation. The river vector layer was obtained from the RBI map

IOP Conf. Series: Earth and Environmental Science **930** (2021) 012095 doi:10.1088/1755-1315/930/1/012095

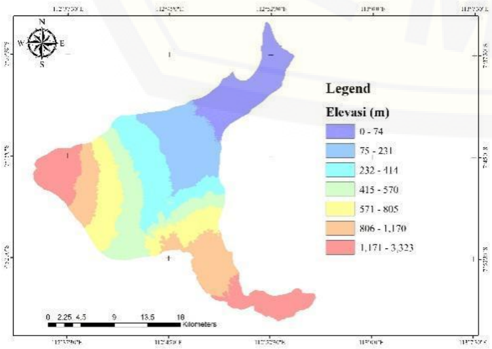
derived into River Density, Distance to the river, and land use factors. The NDVI factor was obtained from the Infrared Band derivative and the Landsat 8 image red band. NDVI was used to identify the reflectance value of the object against the red wave. Value in NDVI has a range from -1 to 1. A value approaching -1 means that low absorption is identified as non-plant, and a value close to 1 has high absorption, which can be identified as plantations. The *Rainfall* Layer was obtained from the Irrigation Agency of Pasuruan. In detail, the source data layer can be seen in table 1.

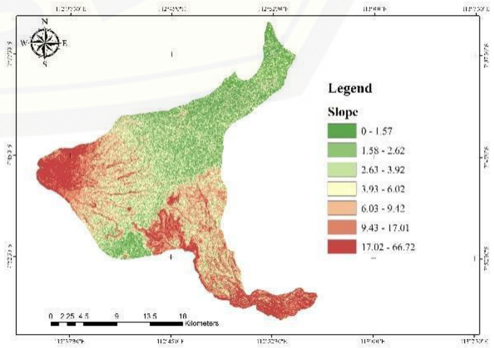
Table 1. Source of data

No.	Classification	Sub-classification	Data source	Resolution	Source
1	Flood Inventory map	Flood inventory	Polygon Coverage		Survey the flood inventory map
2	Topographic map	Slope	SRTM	30 x 30 m	Derivated from SRTM
3		Elevation	SRTM	30 x 30 m	www2.jpl.nasa.gov
4		SPI	SRTM	30 x 30 m	Derivated from SRTM
5		TWI	SRTM	30 x 30 m	Derivated from SRTM
6		Flow Accumulation	SRTM	30 x 30 m	Derivated from SRTM
7	Map layer	Distance to river	River Layer	30 x 30 m	Derivated from Vector
8		River Density	River Layer	30 x 30 m	Derivated from Vector
9	Hydrology	Rainfall	Raint station point	30 x 30 m	Derivated from Vector
10	NDVI	NDVI	Landsat 8	30 x 30 m	Derivated from Landsat
11	Map layer	Landuse	Rupa Bumi Indonesia	1: 25,000	Derivated from Vector

2.2. *Flood conditioning factor*

Each factor is divided into seven classes using the classification method based on the quantile, except land use and rain, which are divided based on natural breaks. Figures 4. a to j are maps of each classified factor.





IOP Conf. Series: Earth and Environmental Science **930** (2021) 012095 doi:10.1088/1755-1315/930/1/012095

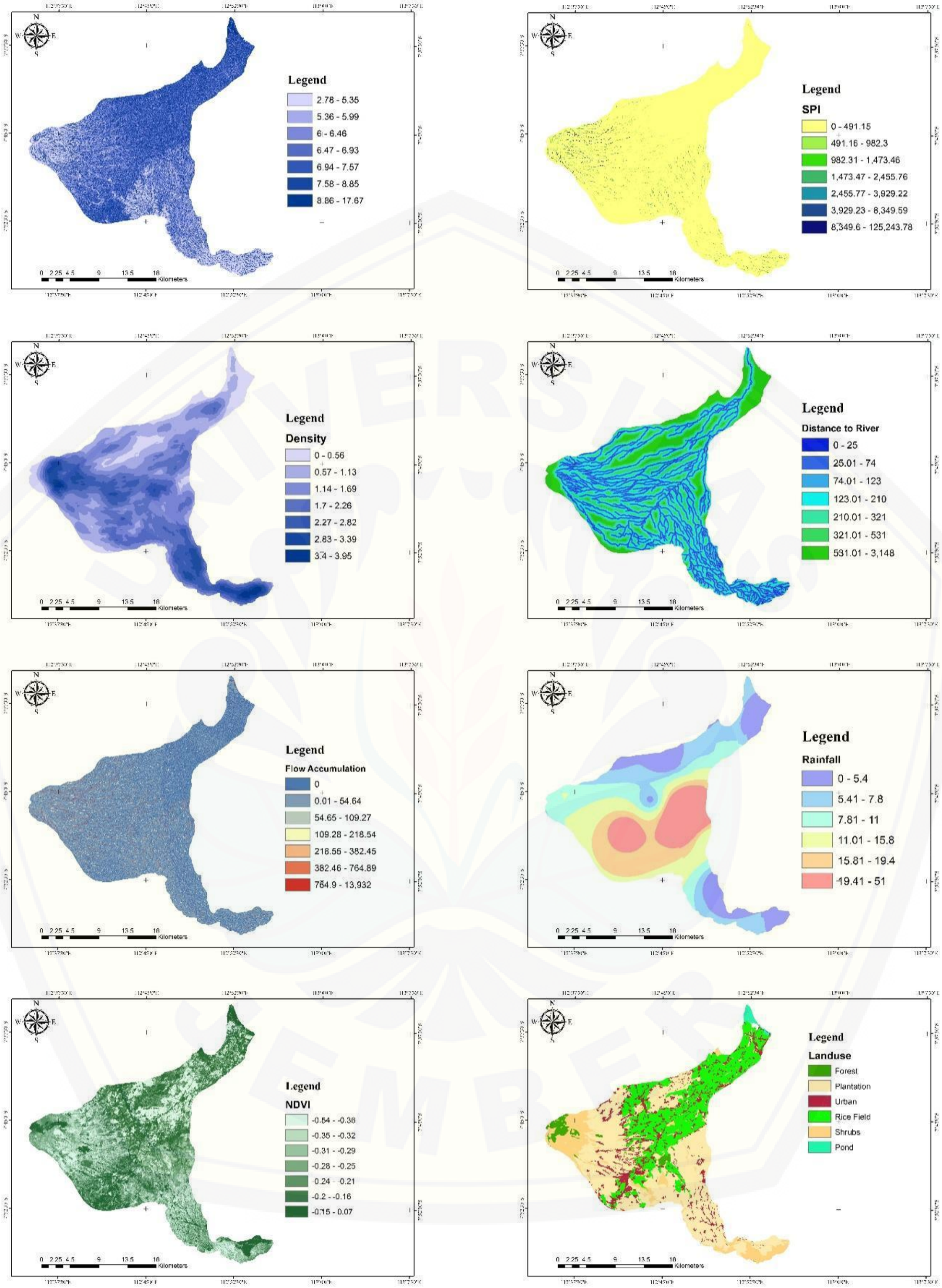


Figure 4. Maps of Each Classified Factors a) Elevation; b) Slope; c) TWI; d) SPI; e) River Density; f) Flow Accumulation; g) Distance to River; h) Rainfall; i) NDVI; j) Landuse

Stream Power Index (SPI) and Topographic Wetness Index (TWI) are two important factors in hydrology. They have the function to evaluate the spatial variation of flood-prone areas, basins erosive strength and relative discharge to the watershed area confirmed by SPI that shows the abrasive strength of flooding. High values mean high flood power, and lower values indicate areas with the potential for flow accumulation in the watershed [3].

$$SPI = As \tan \beta \tag{3}$$

where *As* is the specific basin, and β is the graph of the gradient of local slope (in degrees).

The Topographic Wetness Index (TWI) indicates water accumulation in a river basin, which causes the possibility of a watershed approaching saturated conditions [3].

$$TWI = Ln\left(\frac{\alpha}{\tan \beta}\right) \tag{4}$$

where $\boldsymbol{\alpha}$ is the cumulative drainage slope area through a point (union of contour lengths) and β is the angle of inclination at that point.

Data on land use is obtained from the RBI map at a scale of 1:25,000 with an updated 2020 base map projection. The land use map is of the vector type, converted into a raster grid with a resolution of 30 x 30 m.

Precipitation data are obtained from data recorded by rainfall stations when a flood occurred, converted into GIS, then interpolated using IDW.

2.3. *Frequency Ratio*

The frequency ratio method is used to estimate the probabilistic relationship between dependent and independent variables and to determine the weight coefficient value for each flood class of the related variable [7], [9] applying the following formula:

$$\text{FR} = \frac{FCi}{FS} / \frac{ACi}{AS} \tag{1}$$

Information ;

a. *Fci* is the number of pixels with a flash flood for each class of each variable *i*;
b. *Fs* is the total number of pixels with flash floods in the study area;
c. *Aci* is the number of pixels for each class of each variable *i*;
d. *As* is the total number of pixels in the study area.

Determination of the flood Susceptibility index used the total calculation of the Weight of each parameter. The following is the formula for determining the flood index [9]

$$FSM_{FR} = FR_{elevasi} + FR_{slope} + FR_{SPI} + FR_{twi} + FR_{density} + FR_{landuse} + FR_{distance} + FR_{flow\ accumulation} + FR_{rainfall} + FR_{NDVI} \tag{2}$$

3. Result and Discussion

From the collected data, the calculation results are presented in Table 2.

Table 2. Calculation of Frequency Ratio (FR)

No.	Factor	Class Number	Class	Calculation	%	Occurrence	%	FR
1	Elevation (m)	1	0 - 74	76586.00	13.88	29,528	100.00	7
		2	75 - 231	84721.00	15.36	0	0.00	0
		3	232 - 414	80011.00	14.50	0	0.00	0
		4	415 - 570	77775.00	14.10	0	0.00	0
		5	571 - 805	78342.00	14.20	0	0.00	0
		6	806 - 1,170	78336.00	14.20	0	0.00	0
		7	1,171 - 3,323	75975.00	13.77		0.00	0
				551746.00	86.23	29,528	100.00	7

IOP Conf. Series: Earth and Environmental Science **930** (2021) 012095 doi:10.1088/1755-1315/930/1/012095

No.	Factor	Class Number	Class	Calculation	%	Occurrence	%	FR
2	Slope (degree)	1	0 - 1.57	65960.00	11.95	12,005	40.66	3
		2	1.63 - 2.62	77540.00	14.05	9,042	30.62	2
		3	2.62 - 3.92	83652.00	15.16	5,849	19.81	1
		4	3.93 - 6.02	89091.00	16.15	2,340	7.92	0
		5	6.03 - 9.42	80493.00	14.59	283	0.96	0
		6	9.43 - 17.01	78580.00	14.24	9	0.03	0
		7	17.02 - 66.72	76430.00	13.85		0.00	0
				551746.00	100.00	29,528	100.00	7
3	SPI	1	0 - 491	216161.00	39.18	16,348	55.36	1
		2	491 - 982	318915.00	57.80	13,155	44.55	1
		3	983 - 1473	7077.00	1.28	21	0.07	0
		4	1474 - 2456	4425.00	0.80	4	0.01	0
		5	2457 - 3929	2212.00	0.40	0	0.00	0
		6	3930 - 8350	1610.00	0.29	0	0.00	0
		7	8351 - 125244	1346.00	0.24		0.00	0
				551746.00	99.76	29,528	100.00	2
4	TWI	1	2.78 - 5.35	73309	13.29	19	0.06	0
		2	5.36 - 5.99	84646	15.34	756	2.56	0
		3	6 - 6.46	83221	15.08	3,327	11.27	1
		4	6.47 - 6.93	85416	15.48	6,049	20.49	1
		5	6.94 - 7.57	78058	14.15	6,582	22.29	2
		6	7.58 - 8.85	74789	13.55	6,870	23.27	2
		7	8.86 - 17.67	72307	13.11	5,925	20.07	2
				551746.00	100.00	29,528	100.00	7
5	River network density	1	0 - 0.68	81590.00	14.02	19,314	65.35	5
		2	0.69 - 1.07	83504.00	14.35	2,632	8.91	1
		3	1.08 - 1.38	82473.00	14.17	1,544	5,22	0
		4	1.39 - 1.66	82950.00	14.25	2,909	9.84	1
		5	1.67 - 1.94	84336.00	14.49	1,997	6.76	0
		6	1.95 - 2.4	85164.00	14.63	1,160	3.92	0
		7	2.41 - 3.95	82068.00	14.10	0	0.00	0
				582085.00	100.00	29,556	100.00	7
6	Land use	1	Forest	12547.00	2.16	0	0.00	0
		2	Garden	269651.00	46.43	776	2.72	0
		3	Settlement	66343.00	11.42	3,714	13.04	1
		4	Rice fields	162139.00	27.92	18,954	66.55	2
		5	Bush	65063.00	11.20	53	0.19	0
		6	Pond	5056.00	0.87	4,984	17.50	20
				580799.00	100.00	28,481	100.00	24
7	River distance (m)	1	0-25	36031.00	6.19	1,028	3.48	1
		2	26-74	118137.00	20.30	3,043	10.30	1
		3	75-123	75925.00	13.04	2,113	7.15	1
		4	124-210	102482.00	17.61	3,128	10.58	1
		5	211-321	91167.00	15.66	3,557	12.03	1

IOP Conf. Series: Earth and Environmental Science **930** (2021) 012095 doi:10.1088/1755-1315/930/1/012095

No.	Factor	Class Number	Class	Calculation	%	Occurrence	%	FR
		6	322-531	80702.00	13.86	4,460	15.09	1
		7	532-3148	77641.00	13.34	12,227	41.37	3
				582085.00	100.00	29,556	100.00	7
8	Accumulated flow	1	0-55	518816.00	94.03	29,234	99.00	1
		2	56-109	11446.00	2.07	188	0.64	0
		3	110-219	8386.00	1.52	45	0.15	0
		4	220-328	3898.00	0.71	13	0.04	0
		5	329-546	3682.00	0.67	13	0.04	0
		6	547-929	2820.00	0.51	31	0.10	0
		7	930-13932	2698.00	0.49	4	0.01	0
				551746.00	100.00	29,528	100.00	2
9	Rainfall (mm)	1	0-5.2	80474	13.83	10,592	35.84	3
		2	5.21-7	87122	14.97	12,179	41.21	3
		3	7.01 - 9.4	87557	15.04	6,778	22.93	2
		4	9.41 - 13.2	83163	14.29	7	0.02	0
		5	13.2 - 16.6	79677	13.69	0	0.00	0
		6	16.61-20.2	83203	14.29	0	0.00	0
		7	20.21-51	80889	13.90	0	0.00	0
				582085	100.00	29,556	100.00	8
10	NDVI (ratio)	1	.-0.54 - -0.38	78604	13.50	3,096	10.47	1
		2	.-0.37 - -0.32	80854	13.89	2,330	7.88	1
		3	.-0.31- -0.27	84934	14.59	2,301	7.78	1
		4	.-0.26 - -0.22	86740	14.90	2,772	9.38	1
		5	.-0.21- -0.17	85075	14.61	3,679	12.44	1
		6	.-0.016 - -0.12	81163	13.94	5,269	17.82	1
		7	.-0.11 - 0.07	84749	14.56	10,116	34.22	2
				582119	100.00	29,563	100.00	7

Based on table 2 above, it can be seen that each class of each flood conditioning factor influences flood hazard, which depends on the value of the resulting *Frequency Ratio* (FR). The value above 1 shows the locations that affect the occurrence of flooding. Sequentially, the factors that have a strong influence on flash floods Susceptibility are land use, rain, and then six other factors that have the same effect value are slope, elevation, TWI, river network density, distance to the river, and NDVI. Furthermore, from figures 5-13, the graph of the relationship between the ratio frequency to each factor will explain the details of the factors and classes that trigger flooding.

a. Elevation

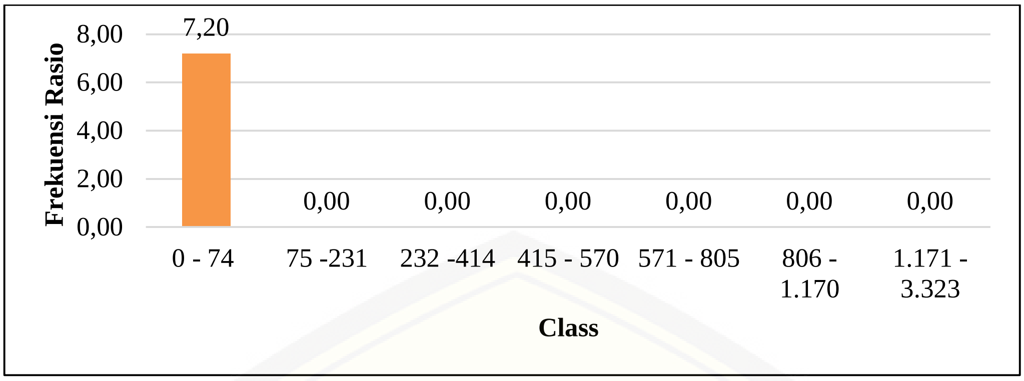


Figure 5. Elevation Cumulative Graph

Based on figure 5, the class of 0–74 with a value of > 1 is 7.20. This class is the only class that has the potential to trigger flooding. In this case, a low elevation is an elevation that triggers a flood-prone.

b. Slope

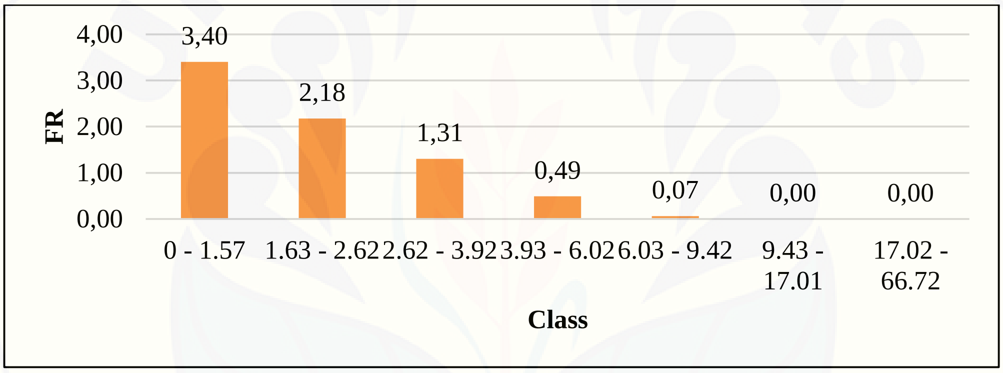


Figure 6. Slope Cumulative Graph

Based on figure 6, the lower the slope, the higher the level of flood Susceptibility potential. The slope class that is prone to flooding in the first class is between 0-3.93.

c. Stream Power Index (SPI)

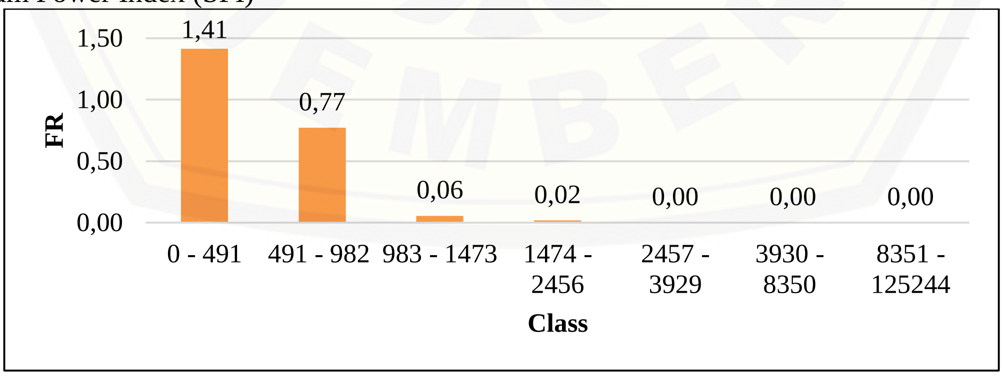


Figure 7. SPI Cumulative Graph

Based on figure 7, the lower the SPI class, the more Susceptibility to flooding. Flood-prone locations are in the SPI class 0 - 491.

d. Topographic Wetness Index (TWI)

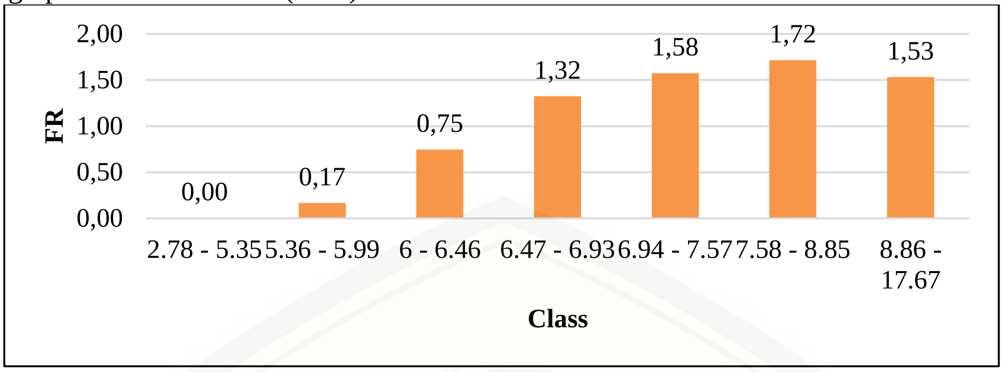


Figure 8. TWI Cumulative Graph

Based on figure 8, the Topographic Wetness Index (TWI) factor is inversely proportional to the two previous factors. The higher the TWI value, the more prone to flooding. Frequency Ratio value greater than the class 6.47-6.93 is a class that triggers floods.

e. The River Network Density

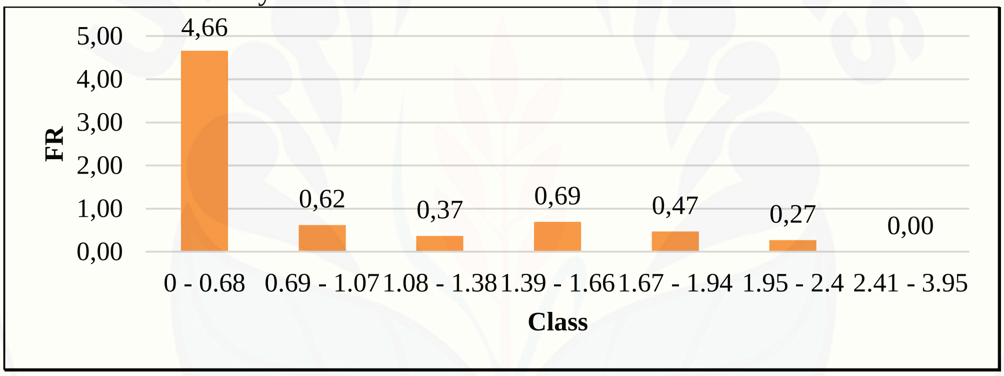


Figure 9: Cumulative River Density Graph

Based on figure 9, the network density factor with a Frequency Ratio value > 1 by 4.66 is a class that triggers Susceptibility to flooding. It can be interpreted that in the lowest class, river network density is the highest, which is potentially prone to flooding.

f. Landuse

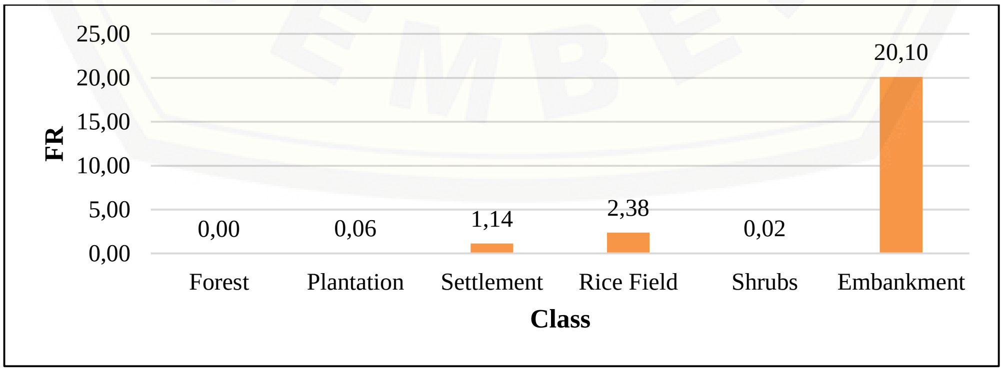


Figure 10. Cumulative Land Use Graph

Based on figure 10, the weight value of FR > 1 is available in the residential class, rice fields, and ponds. The three classes are prone to flooding because settlements, rice fields, and ponds are areas with small infiltrations, so they are easily triggered by flooding.

g. Distance to River

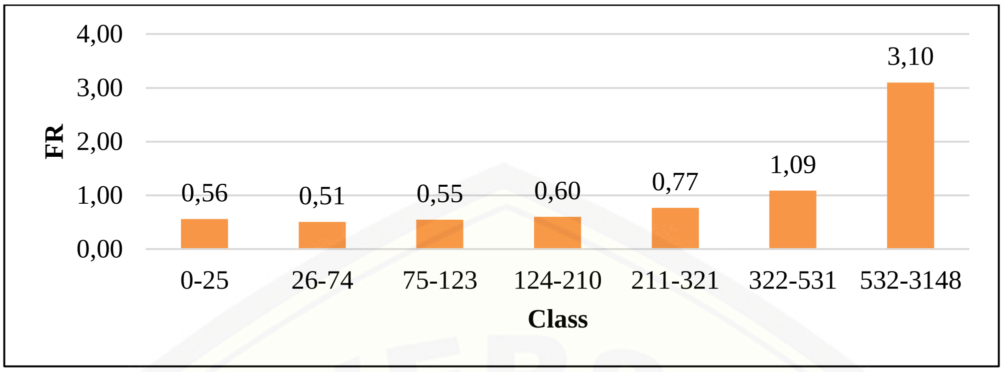


Figure 11. Cumulative Distance to River Graph

Based on figure 11, the Frequency Ratio value > 1 is in class 322 - 3148, which is a potentially flood-prone location. The closer to the river, the more potentially affected by river overflow.

h. Accumulation

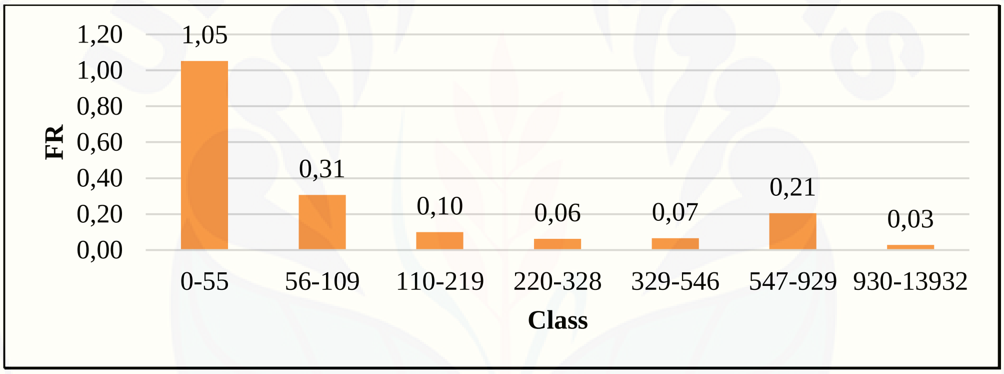


Figure 12. Flow Accumulation graph

Based on figure 12, Frequency Ratio value > 1 is in the class 0 - 55 value by 1.05. It shows that the direction of the flood flow is more to that location. In this Welang watershed, the flow accumulation value is small, not as great as other factors triggering the flooding.

i. Rain

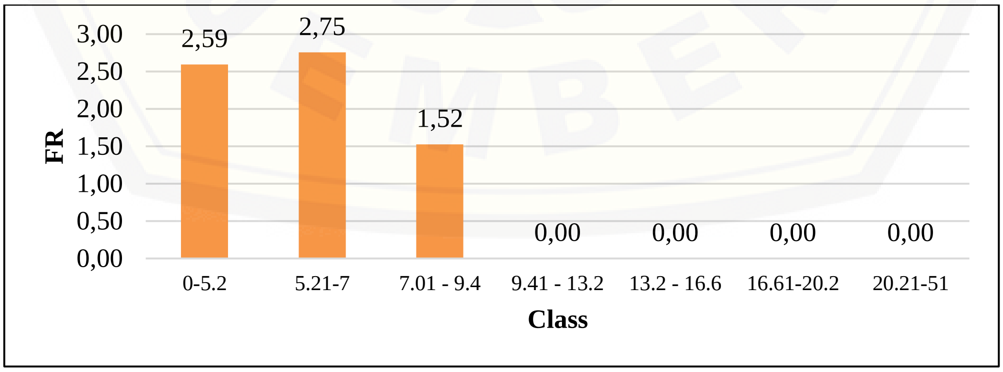


Figure 13. Cumulative rainfall graph

Based on figure 13, the FR value for rain factor > 1 lies in the first, second, and third classes. The low rain value is a location with the potential for flash floods, which means it opposes the actual rain occurrence. It shows that the location of the flood is the highland area with low rainfall, considering that the area is near the coast. Raindrops from upstream cause this flood.

j. NDVI

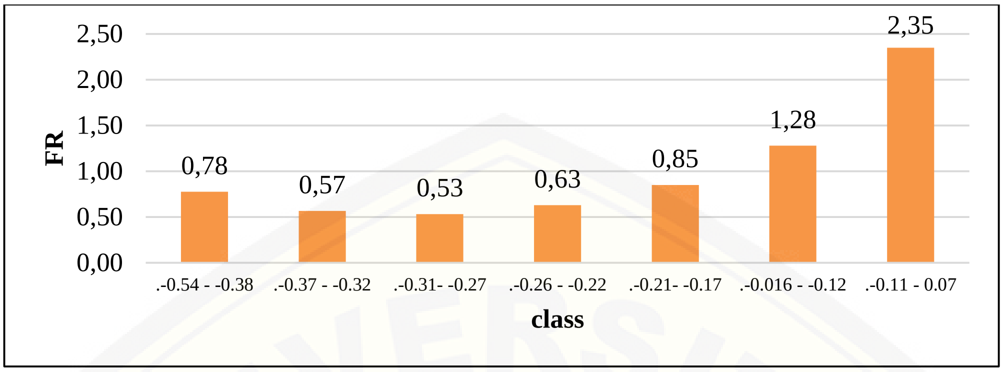


Figure 14. NDVI Graph

Based on figure 14, the FR value of more than 1 is in the sixth and seventh grades. The results show that the greater the NDVI value, the greater the potential for flooding to occur.

The susceptibility index class in Figures 15 and 16 is divided into six classes: Very Low, Low, Moderate, Moderate to High, High, and Very High. The percentage values obtained for Welang watershed are: Very Low 32%, Low 37%, Moderate 15%, Moderate to High 9%, High 6%, and Very High 1% as shown in Figure 16 below.

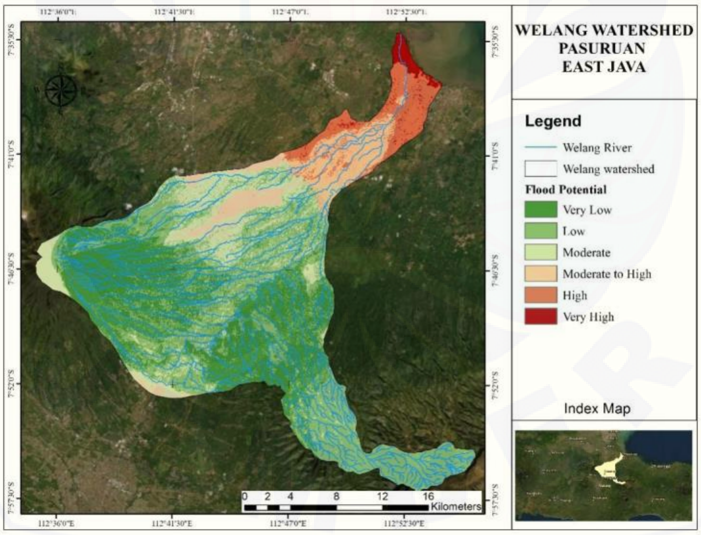


Figure 15. Map of Flood Susceptibility in Welang Watershed

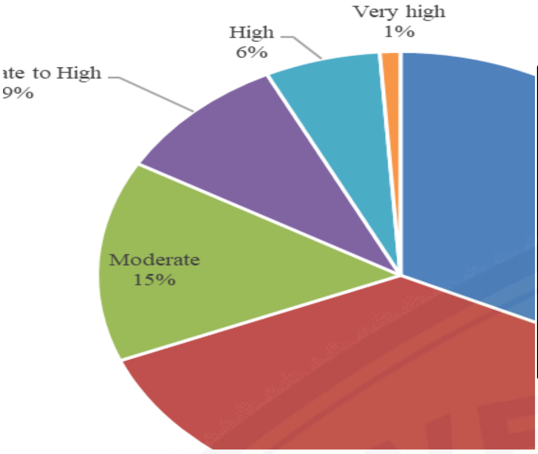


Figure 16. The percentage value of potential flooding

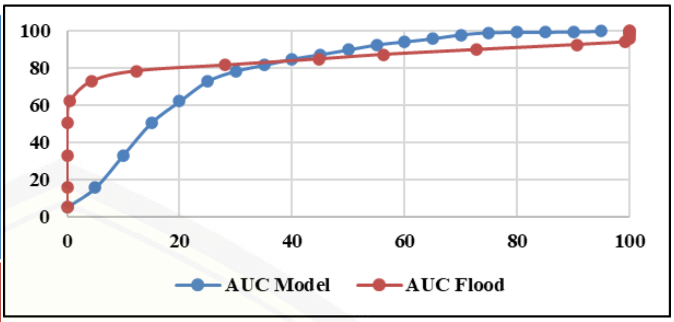


Figure 17. AUC graph

In Figure 17, the results of the calculation of model validation show that the AUC model is 90% and the AUC model is 93%. These results indicate that the flood Susceptibility map has a significantly good level of accuracy.

4. Conclusion

Flood hazard modeling using FR can represent locations that are potentially prone to flooding. The accuracy of FR calculated based on AUC can be applied to the flood hazard map of the Welang watershed since it has an AUC accuracy value of 93%. The percentage of the susceptible area from moderate to very high is 22%.

In this study, factors that contribute to modeling flood events are land use, elevation, river network density, distance to the river, and slope.

The mapping of potential flood Susceptibility using frequency ratio is a method that utilizes the spatial similarity of each model. It allows it to be applied in areas that have the same characteristics in potential flooding.

References

[1] C. Arrighi *et al.* 2018 Quantification of flood risk mitigation benefits: A building-scale damage assessment through the RASOR platform *J. Environ. Manage.* vol **207** pp 92–104 doi: 10.1016/j.jenvman.2017.11.017.

[2] J. Vogt and P. Stephan 2012 Using microencapsulated fluorescent dyes for simultaneous measurement of temperature and velocity fields *Meas. Sci. Technol.* vol 23 no 10 pp 48–56 doi: 10.1016/j.ijdrr.2017.09.019.

[3] K. Khosravi *et al.* 2018 A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran *Sci. Total Environ.* vol 627 pp 744–755 doi: 10.1016/j.scitotenv.2018.01.266

[4] R. Mind *et al.* 2019 x *Int. J. Disaster Risk Reduct.* vol 38 no April 2018 p 101211 doi: 10.1016/j.ijdrr.2019.101211

[5] S. Mirzaei, M. Vafakhah, B. Pradhan, and S. J. Alavi 2021 Flood susceptibility assessment using extreme gradient boosting (EGB), Iran *Earth Sci. Informatics* vol 14 no 1 pp 51–67 doi: 10.1007/s12145-020-00530-0

[6] O. Rahmati, H. Zeinivand, and M. Besharat 2016 Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis *Geomatics, Nat. Hazards Risk* vol 7 no 3 pp 1000–1017 doi: 10.1080/19475705.2015.1045043

[7] C. Cao, P. Xu, Y. Wang, J. Chen, L. Zheng, and C. Niu 2016 Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas

Sustain. vol 8 no 9 doi: 10.3390/su8090948

[8] K. Khosravi, E. Nohani, E. Maroufinia, and H. R. Pourghasemi 2016 A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique', *Nat. Hazards*, vol. 83, no. 2, pp. 947–987, 2016, doi: 10.1007/s11069-016-2357-2.

[9] S. Lee and B. Pradhan, 'Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models', *Landslides*, vol. 4, no. 1, pp. 33–41, 2007, doi: 10.1007/s10346-006-0047-y.