VIVA 2022

PREDICTION OF FLOOD IN TROPICAL REGION USING GEOGRAPHIC INFORMATION SYSTEM AND MULTI-LAYER PERCEPTRON NEURAL NETWORK

PRESENTED BY:

TAN KAI YING
17009034
BACHELOR OF ENGINEERING WITH HONS., CIVIL

SUPERVISED BY:

DR MUHAMMAD RAZA UL MUSTAFA

TABLE OF CONTENT

INTRODUCTION

BACKGROUND

Floods are considered as one of the most dangerous natural hazards that have tremendous impact on both property and communities all around the world every year

	Flood Incidents
Globally	1.47 billion (19% of global population) people exposed to flood in 2022
Asia	40% of total flood disaster, highest among all continents (past 10 years)
Malaysia	6.6 out of ten risk index, 23.8% happened as of April 2022
Klang Valley (Klang River Basin)	1228 flood incidents happened from 2014 to 2018 34 flood hotspots detected in Q1 2022

BACKGROUND

Flood damage

- resulted in economic loss of 1 trillion US dollars (RM 4.4 trillion) yearly by 2050

Advanced models developed using hydrological and machine learning algorithms (GIS + ML)

- an effective mitigation where helps in developing effective flood prevention and management plans

PROBLEM STATEMENT

Flood Causative Factors:

- Different studies used different factors and different amount, affect the accuracy of result
- Appropriate and suitable number of factors are considered
- Refer to previous studies: Avand et al. (2021), Islam et al. (2022), and Luu et al. (2021)

ANN is mostly adopted and proven by other researchers to be an effective technique

- stock prediction
- groundwater potentiality

GIS:

- the most suitable technique in dealing with database that contain geographic data

No study on Klang River Basin using ANN-MLP with GIS

OBJECTIVE

The main objective of this research study is to predict flood in a tropical region using Geographic Information System (GIS) in conjunction with Multi-layer Perceptron (MLP) Neural Network.

- To investigate the flood causative factors in the tropical region.
- To predict the land use using Multi-layer Perceptron (MLP) neural network.
- To develop flood mapping using Geographic Information System (GIS).

SCOPE OF STUDY

FLOOD CAUSATIVE FACTOR

- Elevation, slope, land use, rainfall, drainage density, geology, curvature, Topographical Wetness Index (TWI), Normalized Difference Vegetation Index (NDVI), and flow accumulation

GEOGRAPHIC INFORMATION SYSTEM (GIS)

- Spatial analysis
- Flood Mapping

ARTIFICIAL NEURAL NETWORK (ANN)

- Multi-layer Perceptron (MLP) neural network

LITERATURE REVIEW

FLOOD CAUSATIVE FACTORS

16% to 98% of the world population are exposing to at least medium-level flood risk

Rentschler et al. (2021)

Di Salvo et al. (2018);
Liu et al. (2015);
Osei et al. (2021)

The common flood causative factors:

- low capacity of the watercourse
- high river discharge
- the incapacity of the natural courses to transport excess water
- improper land use planning
- urban development or river morphology changes that causes the alteration of the natural water cycle

GEOGRAPHIC INFORMATION SYSTEM

Komolafe et al. (2018)

GIS is an effective system due to its ability to analyse spatial and distributed data which allow many variables to be included in the calculation and evaluation to predict complicated flood damages.

Di Salvo et al. (2018)

In GIS environment, all the data such as location, shape, and related attributes are able to be managed or to be assigned weights which result in analysation can be done effortless

ARTIFICIAL NEURAL NETWORK - MULTI LAYER PERCEPTRON

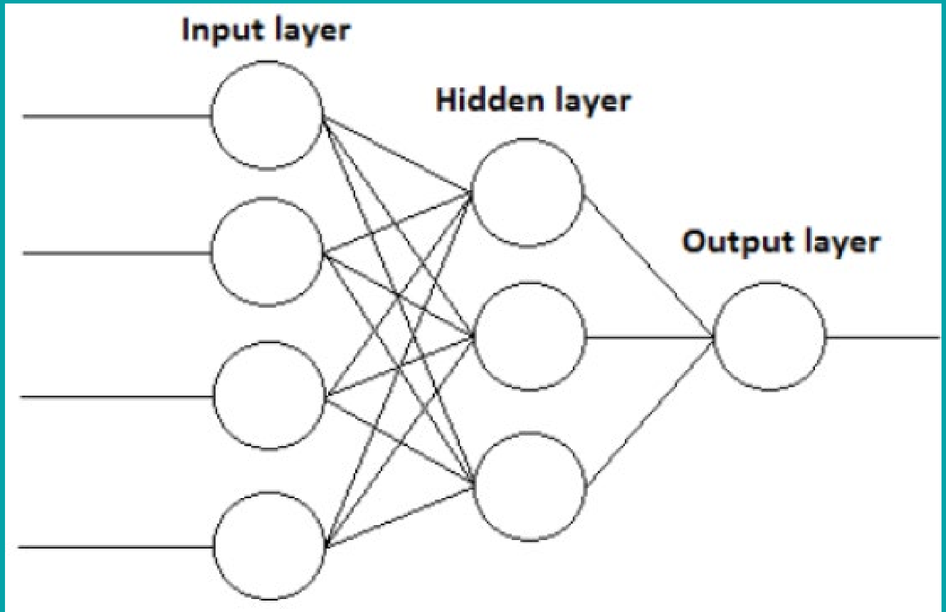
Das & Saha (2022)

ML techniques are adopted because it is more reliable and precise in dealing with nonlinear data compared to probabilistic methods

Liu et al. (2015)

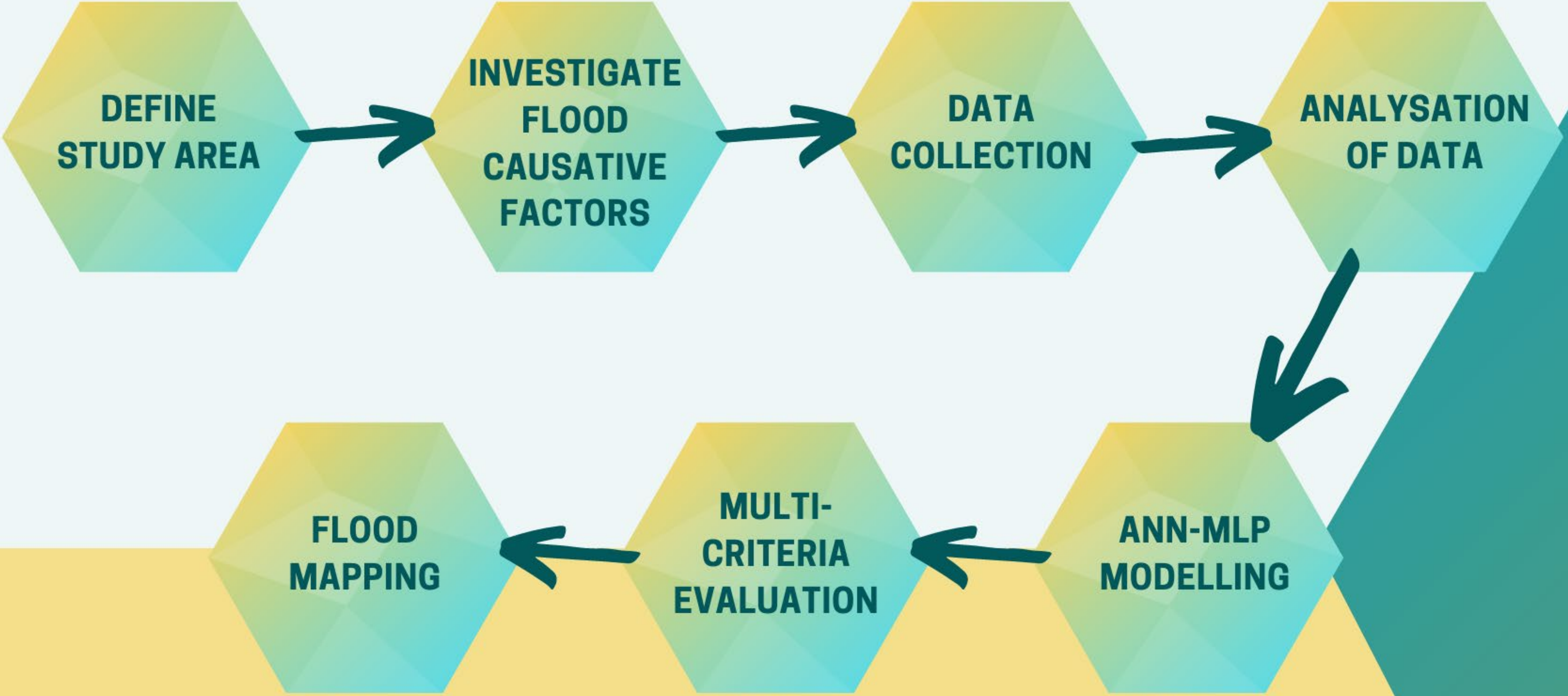
MLP has become an important tool for prediction purposes due to its abilities such as nonlinear activation and generalization skills

Carlsson (2015)



METHODOLOGY

FLOW CHART


DEFINE STUDY AREA
INVESTIGATE FLOOD CAUSATIVE FACTORS
DATA COLLECTION
ANALYSATION OF DATA
ANN-MLP MODELLING
MULTI-CRITERIA EVALUATION
FLOOD MAPPING

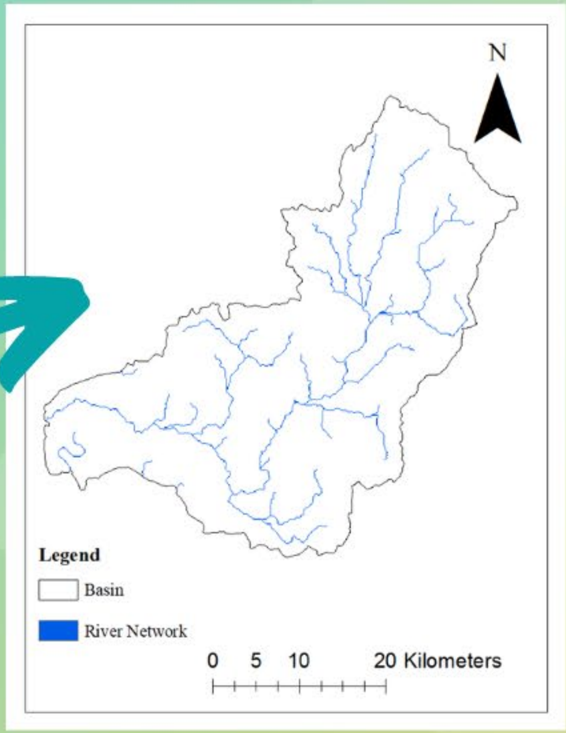
STUDY AREA

Located at Southwest of Peninsular Malaysia
Catchment area of Klang River Basin = 1288 km2

More than 50% - infrastructure, residential, and industrial areas
15% - permanent forest reserves and agriculture services
other land uses - community facilities, open space, and commercial areas

Klang River Basin was chosen:

- Frequent flood
- A tropical region



KLANG RIVER BASIN

According to Summary Environmental Impact Assessments, flooding has always been affecting the quality of life in Klang Valley, approximately 14% of Klang Valley is flood prone area.

RESEARCH APPROACH

INVESTIGATION OF FLOOD CAUSATIVE FACTORS

Based on the research done, 10 factors were investigated

- elevation, slope, land use, rainfall, drainage density, geology, curvature, Topographical Wetness Index (TWI), Normalized Difference Vegetation Index (NDVI), and flow accumulation.

Previous study done by Avand et al. (2021); Darabi et al. (2021); Islam et al. (2022); Luu et al. (2021); Panahi et al. (2021); Priscillia et al. (2021) were used as references

RESEARCH APPROACH

DATA COLLECTION

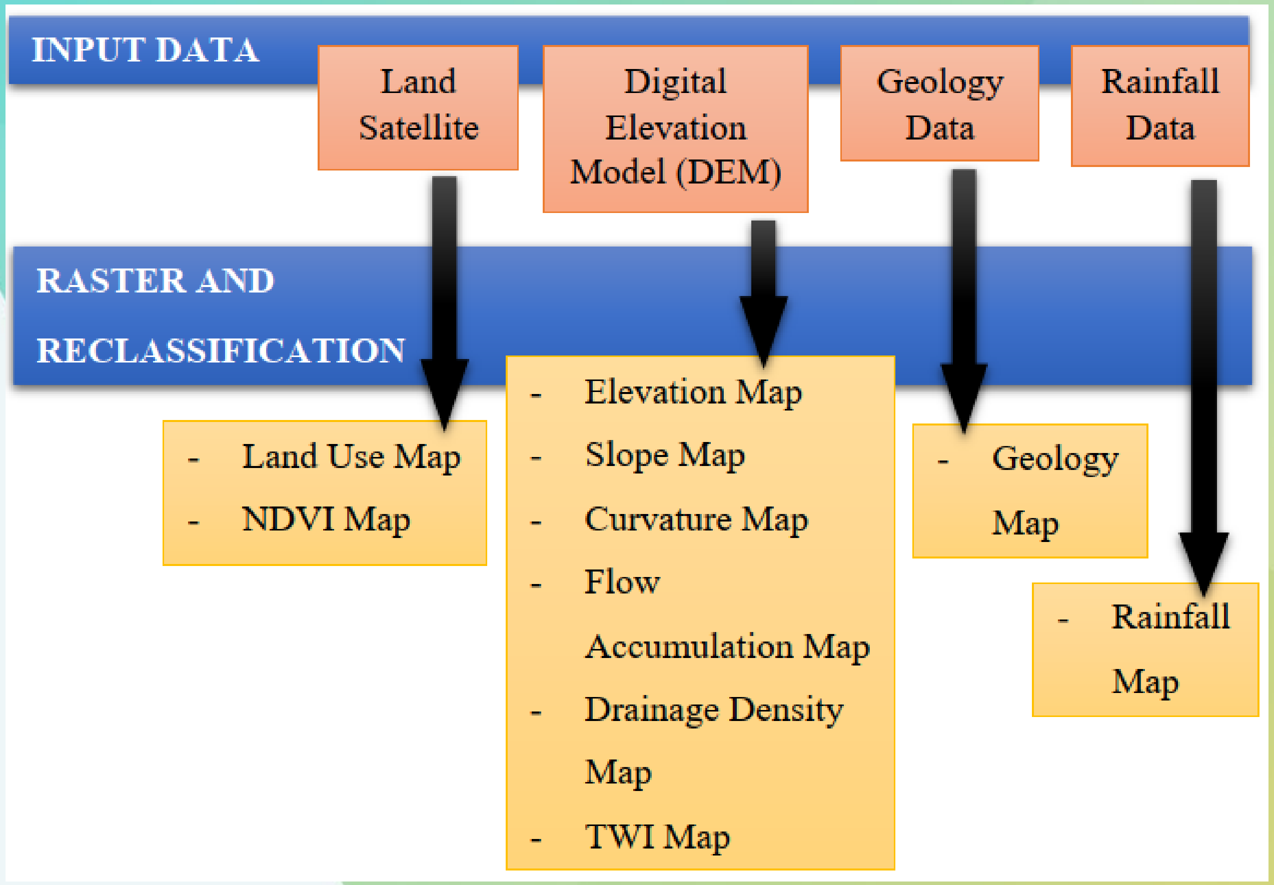
Year Range: 2011 - 2021

DATA	SOURCES (URL)	TYPE	DATE / PERIOD
Land Satellite	USGS https://earthexplorer.usgs.gov/	Landsat 4-5 TM C2 L1 Landsat 8-9 OLI/TIRS Collection 2 Level 2	2011/07/06 (path/row – 127/058) 2021/02/07 (path/row – 127/058)
Digital Elevation Model (DEM)	USGS https://earthexplorer.usgs.gov/	Digital Elevation: STRM 1 Arc-Second Global	2014/09/23 (Coordinates: 2, 101 and 3, 101)
Geology Data	USGS Geology Map https://certmapper.cr.usgs.gov/data/apps/world-maps/	Generalized Geology of Southeast Asia	-
Rainfall Data	CHRS Data Portal https://chrsdata.eng.uci.edu/	PERSIANN Yearly Dataset	2011 - 2021

RESEARCH APPROACH

ANALYSATION OF DATA USING GEOGRAPHIC INFORMATION SYSTEM (GIS)

- To obtain the spatial map of the flood causative factors
- From the data collected, the 10 flood causative factors have been analysed, calculated, and mapped by using GIS software, ArcGIS.



RESEARCH APPROACH

ARTIFICIAL NEURAL NETWORK (ANN) - MULTI LAYER PERCEPTRON (MLP)

Land Use Prediction

To train the model, the land use of the year 2011 was used as the initial input, land use of the year 2021 was used as the final input, and 3 factors (elevation, curvature, and slope) were used as the spatial variables

- only these factors will change and affect the land use condition

10 years interval was used - reference map was changed

QGIS with Plugin Toolbox, MOLUSCE was used

Data training, data testing, and data evaluation were done where the system will help to calculate the accuracy of result automatically

Trials and errors were done until the highest accuracy of result was obtained

RESEARCH APPROACH

MULTI-CRITERIA EVALUATION - ANALYTIC HIERARCHY PROCESS (AHP) METHOD

Multi-criteria evaluation using Analytic Hierarchy Process (AHP) was utilized
- to organize the decision
- to determine the rank and the weightage of each flood causative factor

Matrix consistency is checked to ensure the matrix is consistent by calculating the consistency ratio (CR) of the pairwise comparison matrix of the factors

CR value is less than or equal to 0.1 (10%) - the inconsistency of the matrix is acceptable

CR value is more than 0.1 (10%) - some revisions will need to be done

$$CR = \frac{CI}{RI}$$

$$CI = \frac{\lambda max - n}{n - 1}$$

where:

CI = Consistency index,

lambda max = principal eigenvalue

n = number of parameters used

RI= Random index

RESEARCH APPROACH

MULTI-CRITERIA EVALUATION - ANALYTIC HIERARCHY PROCESS (AHP) METHOD

No.	PARAMETERS	WEIGHTS
1	Rainfall	29.1%
2	Drainage Density	20.9%
3	Flow Accumulation	15.3%
4	Land Use	10.1%
5	Elevation	6.5%
6	Slope	6.7%
7	TWI	4.5%
8	Curvature	3.1%
9	Geology	2.2%
10	NDVI	1.7%

Principal eigenvector value = 10.45
Consistency Index (CI) value = 0.05
Random Index (RI)= 1.49
Consistency Ratio (CR) = 0.034 (3.4%)

CR value determined is less than 0.1 (10%)

.: this signified that the weightage and the rank of the parameters are acceptable

Result
Eigenvalue — Lambda: 10.450 — MRE: 31.3%
Consistency Ratio — 0.37 — GCI: 0.12 — Psi: ##### — CR: 3.4%

N	1	2	3	4	5	6	7	8	9	10
Random Index (RI)	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

RESEARCH APPROACH

DEVELOP FLOOD MAPPING USING GIS

The flood risk maps were developed

- using a geoprocessing tool (Weighted Overlay) in ArcGIS

All the 10 parameters were fused into one layer and produced a flood risk map

Flood Risk Map of Year 2031, 2041, 2051

- 10 spatial data of flood causative factors were used
- Replace year 2021 land use map with the new predicted land use obtained (year 2031, 2041, 2051)

SOFTWARE

ArcGIS

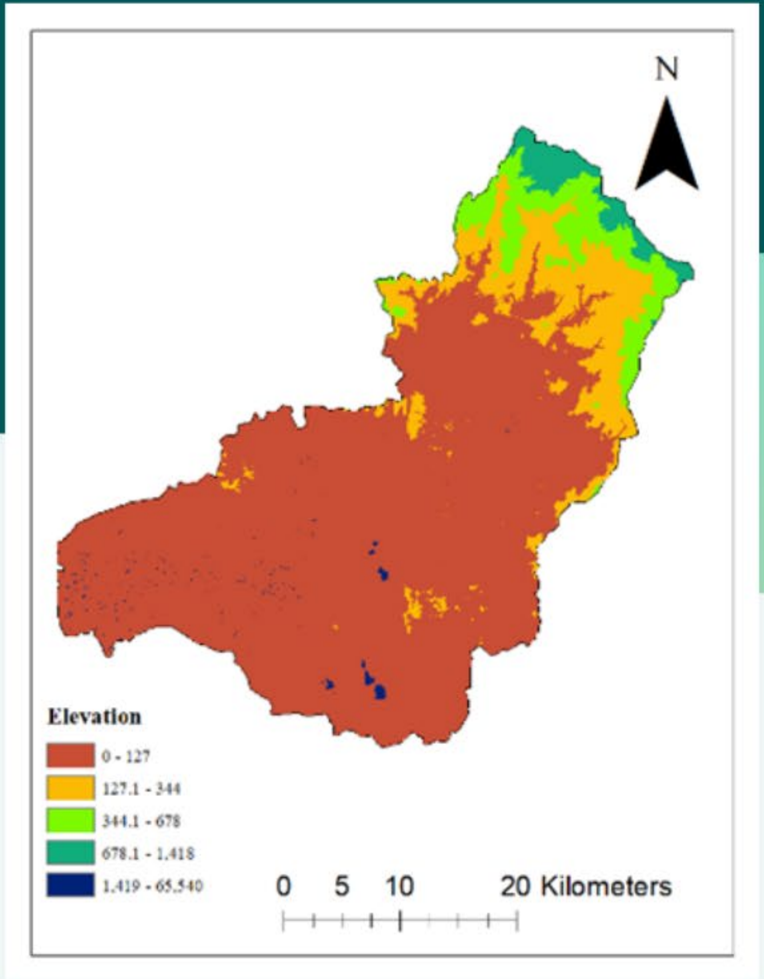
- A geospatial software that - manage, edit, and analyse geographic data

QGIS

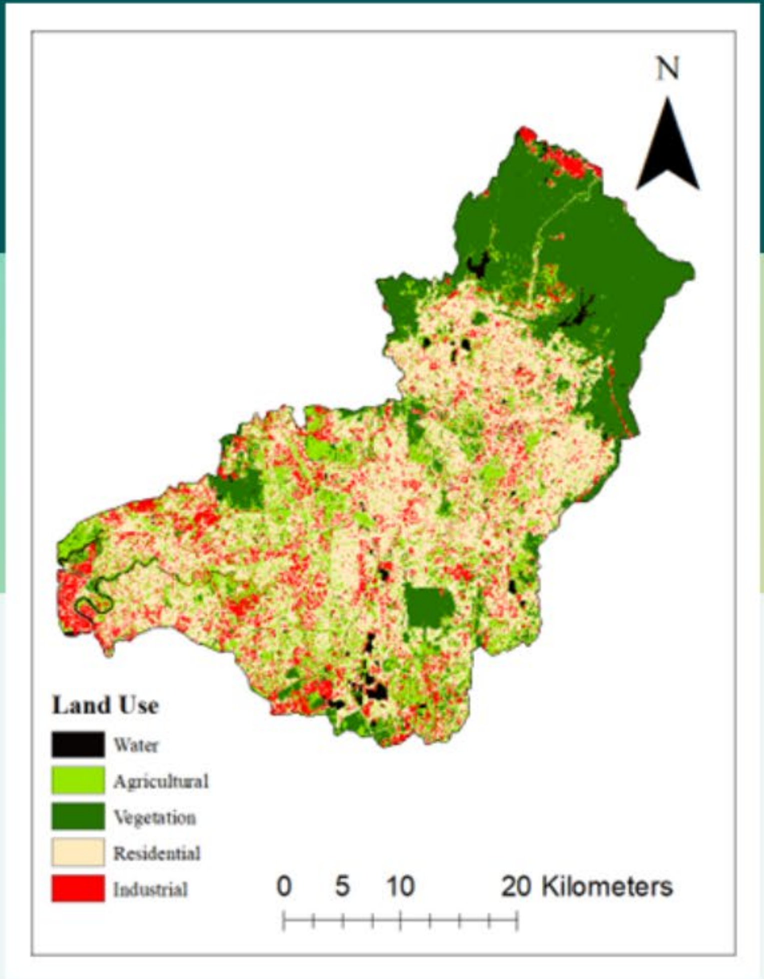
- Quantum Geographic Information System
- MOLUSCE plugin

RESULT AND DISCUSSION

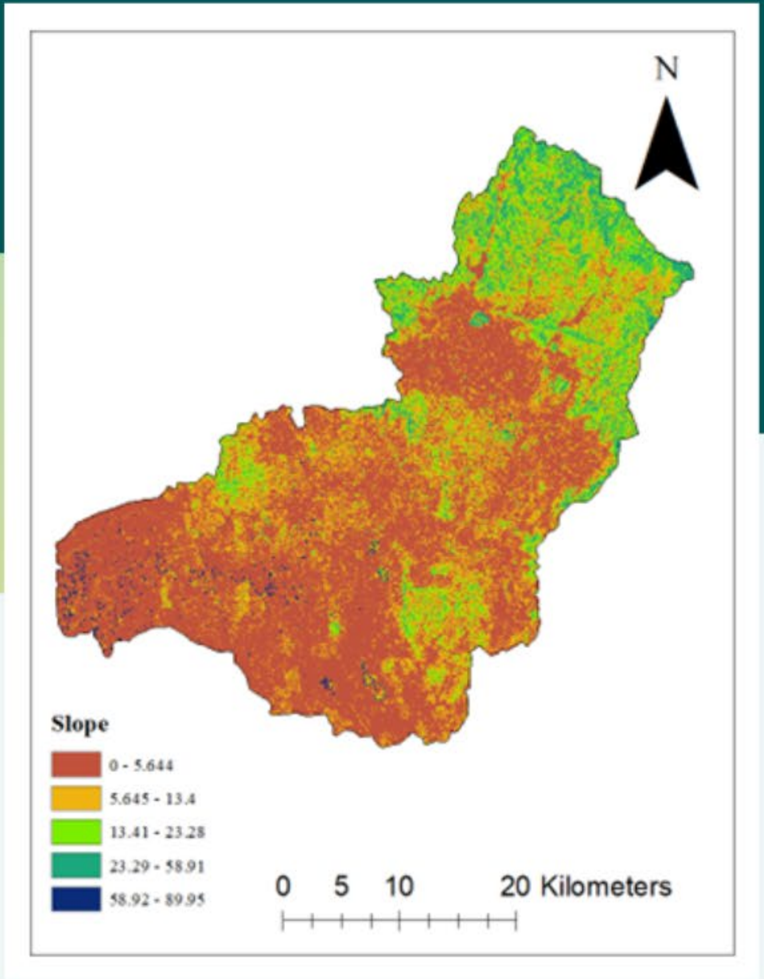
FLOOD CAUSATIVE FACTORS



ELEVATION

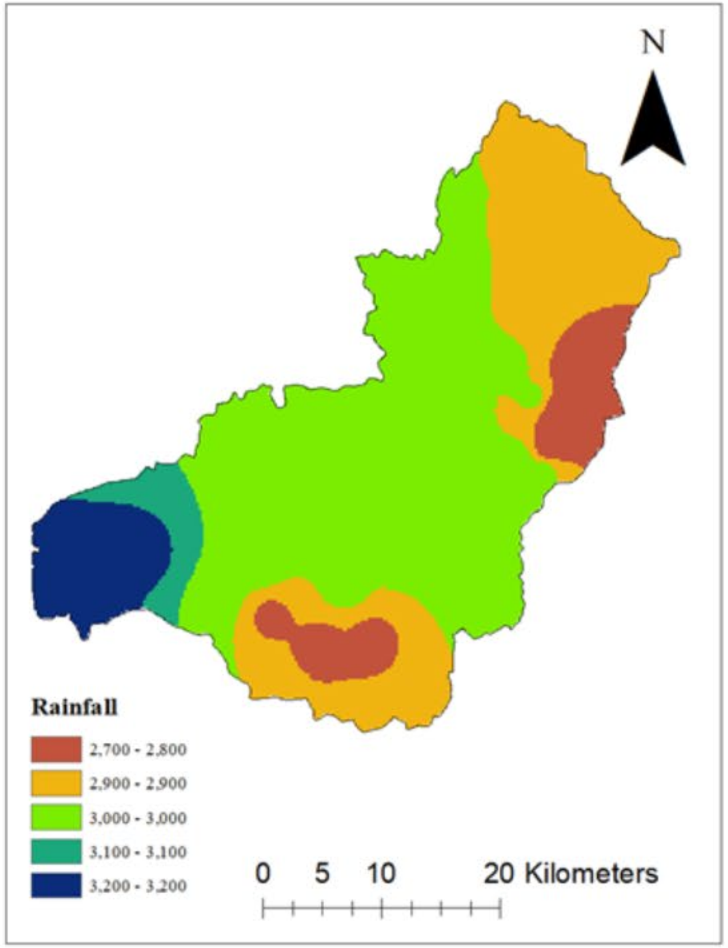


LAND USE

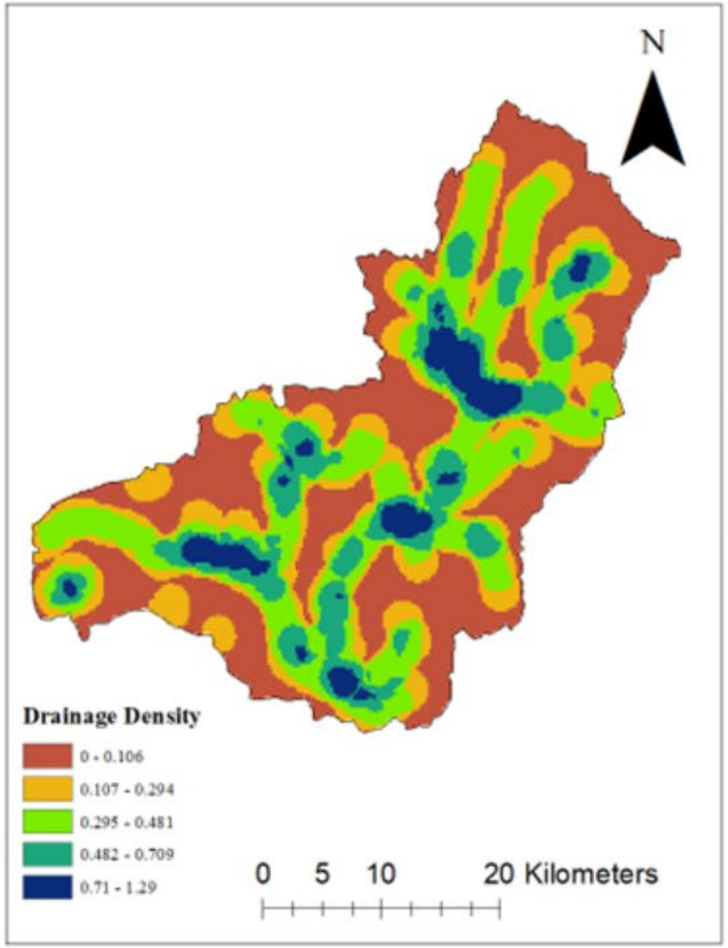


SLOPE

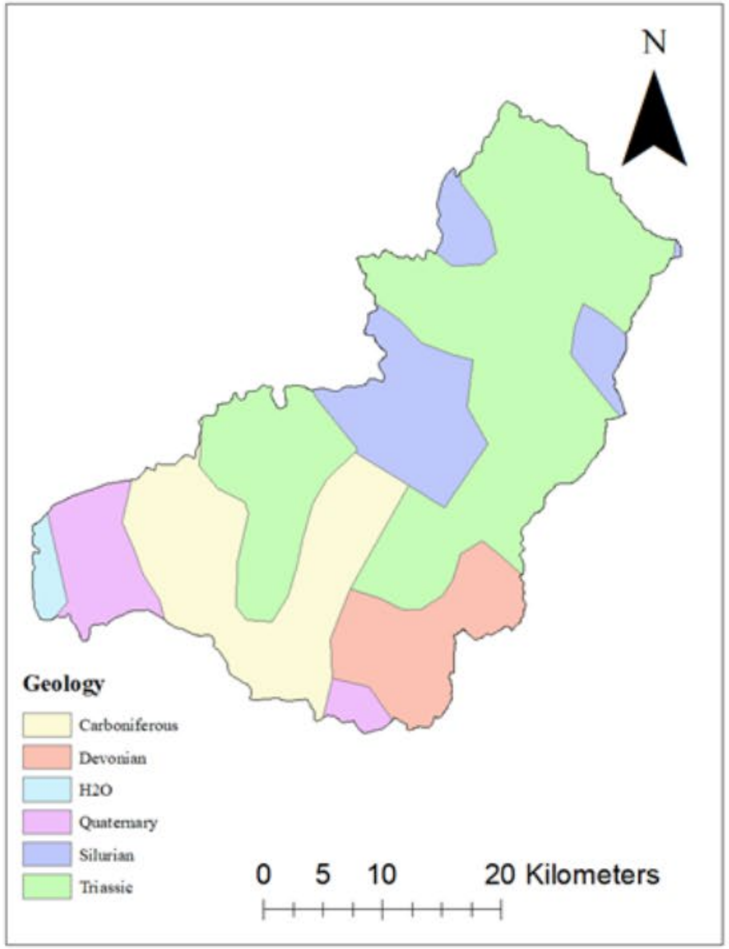
FLOOD CAUSATIVE FACTORS



RAINFALL

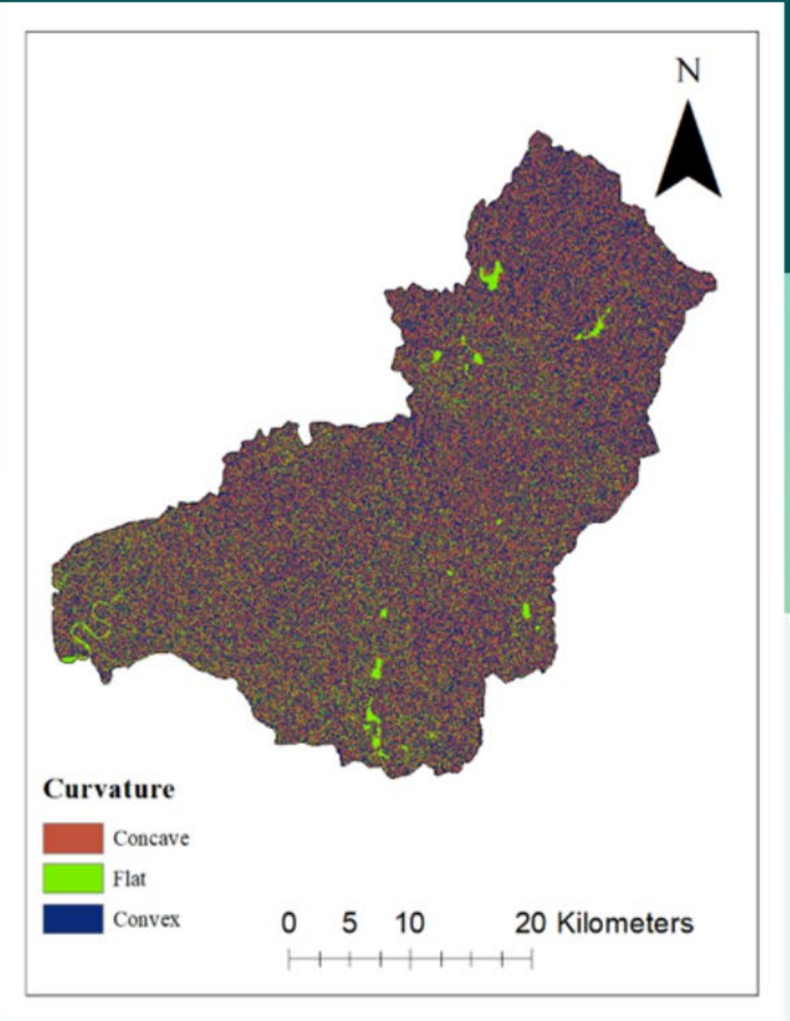


DRAINAGE DENSITY

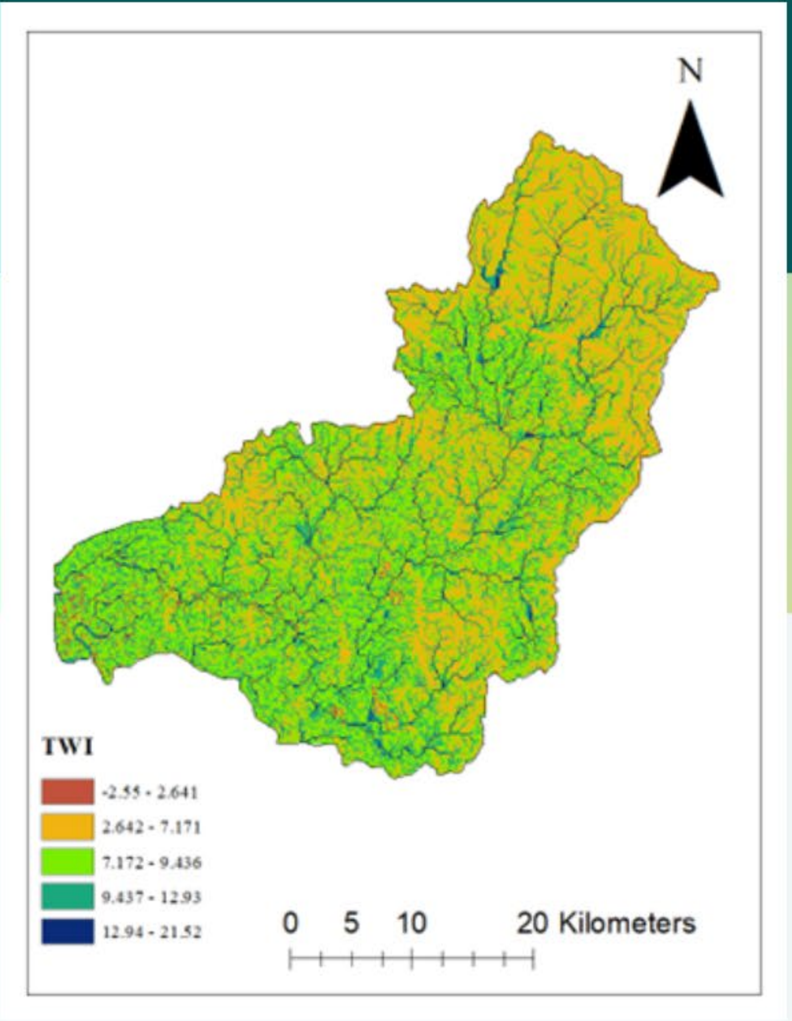


GEOLOGY

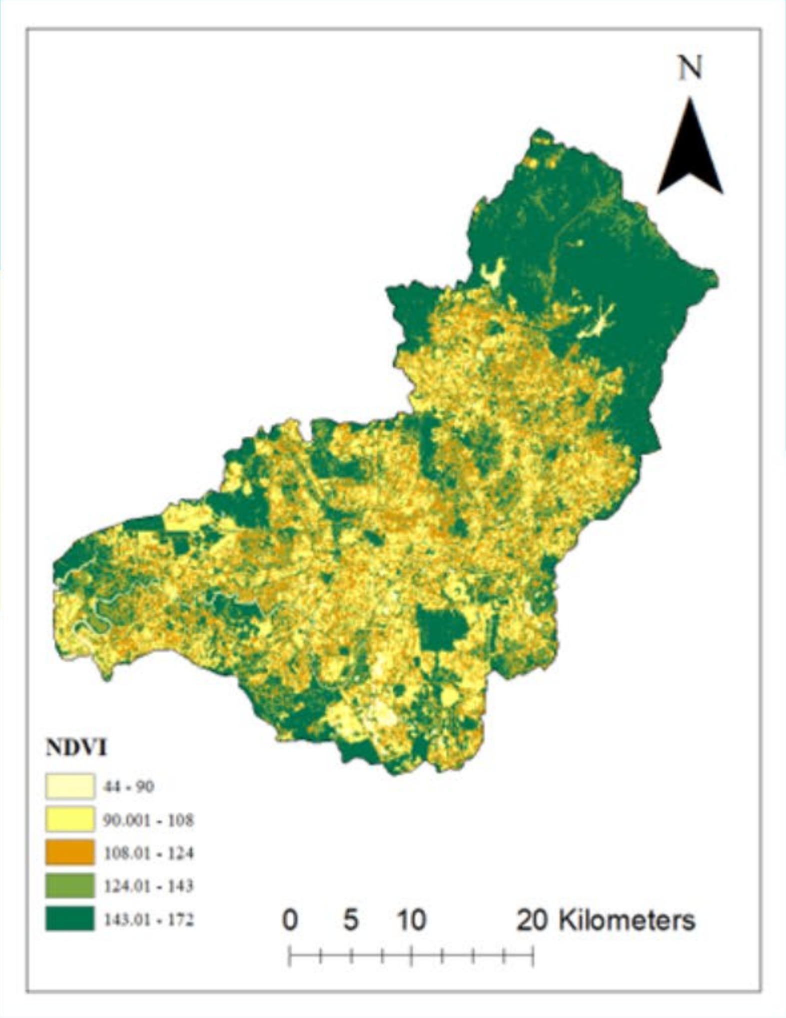
FLOOD CAUSATIVE FACTORS



CURVATURE

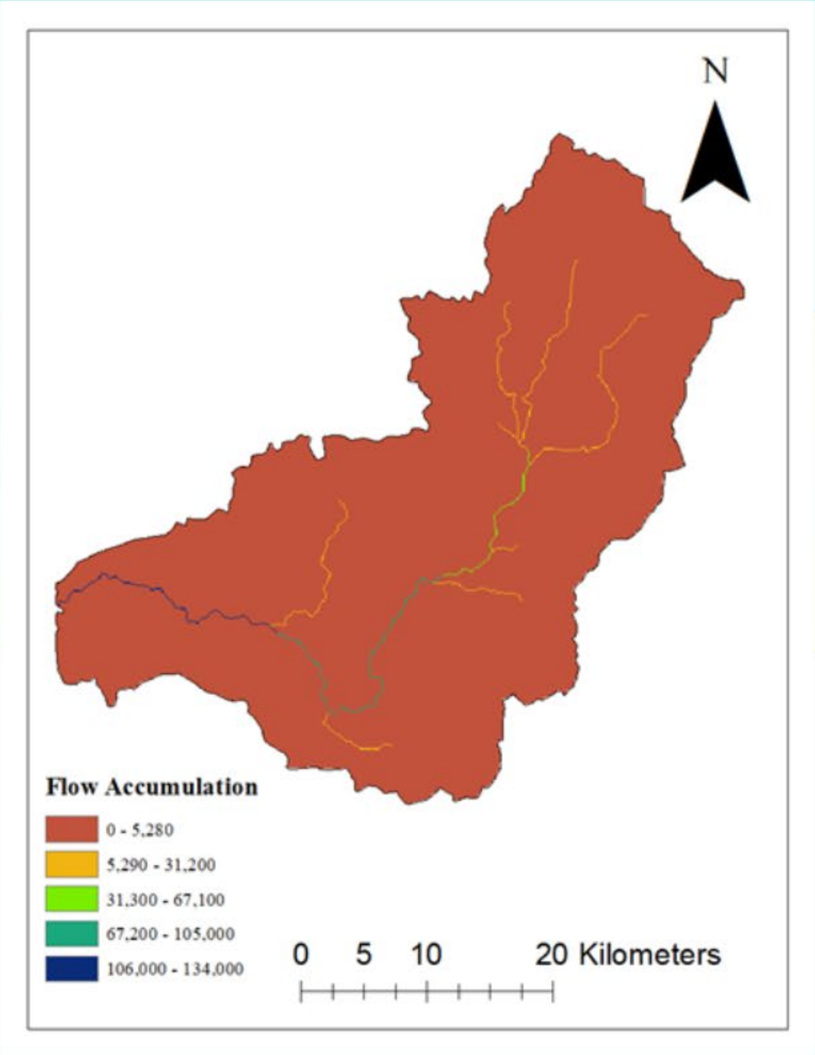


TOPOGRAPHICAL WETNESS INDEX (TWI)



NORMALIZED DIFFERENCE VEGETATION INDEX (NDVI)

FLOOD CAUSATIVE FACTORS



FLOW ACCUMULATION

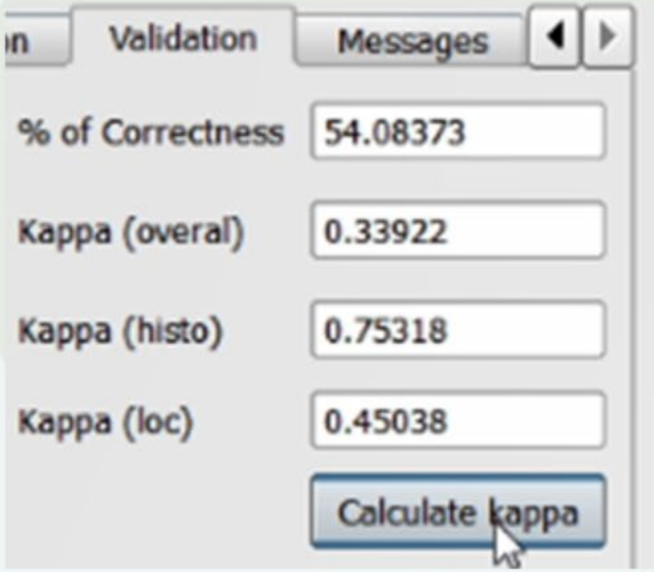
FLOOD CAUSATIVE FACTORS

- Different sub-factors have different levels in causing flood

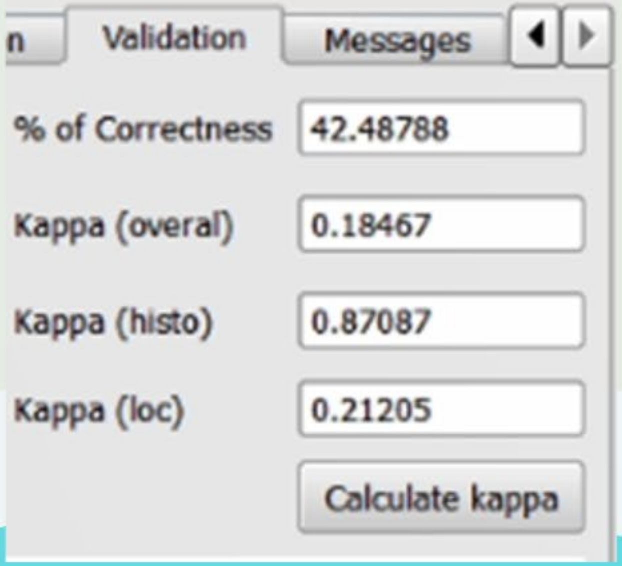
Rank	Factors	Weight (%)	Sub-factors	Descriptive Level	Rating
1	Rainfall (mm/year)	29.1	2700 - 2800	Very Low	1
			2900 - 2900	Low	2
			3000 – 3000	Medium	3
			3100 – 3100	High	4
			3200 – 3200	Very High	5
2	Drainage Density (km/km2)	20.9	0 – 0.106	Very Low	1
			0.107 – 0.294	Low	2
			0.295 – 0.481	Medium	3
			0.492 – 0.709	High	4
			0.71 – 1.29	Very High	5
3	Flow Accumulation	15.3	0 – 5280	Very Low	1
			5290 - 31200	Low	2
			32300 - 67100	Medium	3
			67200 - 105000	High	4
			106000 - 134000	Very High	5
4	Land Use	10.1	Water	Very High	5
			Agricultural	Very Low	1
			Vegetation	Very Low	1
			Residential	High	4
			Industrial	Medium	3
5	Slope (degree)	6.7	0 – 5.644	Very High	5
			5.645 – 13.4	High	4
			13.41 – 23.28	Medium	3
			23.29 – 58.91	Low	2
			58.92 – 89.95	Very Low	1
6	Elevation (m)	6.5	0 - 127	Very High	5
			127.1 - 344	High	4
			344.1 - 678	Medium	3
			678.1 - 1418	Low	2
			1419 - 65540	Very Low	1
7	Topographical Wetness Index (TWI)	4.5	-2.55 – 2.641	Very Low	1
			2.642 – 7.171	Low	2
			7.172 – 9.436	Medium	3
			9.437 – 12.93	High	4
			12.94 – 21.52	Very High	5
8	Curvature (rad/m)	3.1	Concave	Medium	3
			Flat	Very Low	1
			Convex	Very High	5
9	Geology	2.2	Carboniferous	Low	2
			Devonian	Medium	3
			H2O / Quaternary	Very High	5
			Silurian	Very Low	1
			Triassic	High	4
10	Normalized Difference Vegetation Index (NDVI)	1.7	44 - 90	Very High	5
			90.001 - 108	High	4
			108.01 - 124	Medium	3
			124.01 - 143	Low	2
			143.01 - 172	Very Low	1

LAND USE PREDICTION RESULT

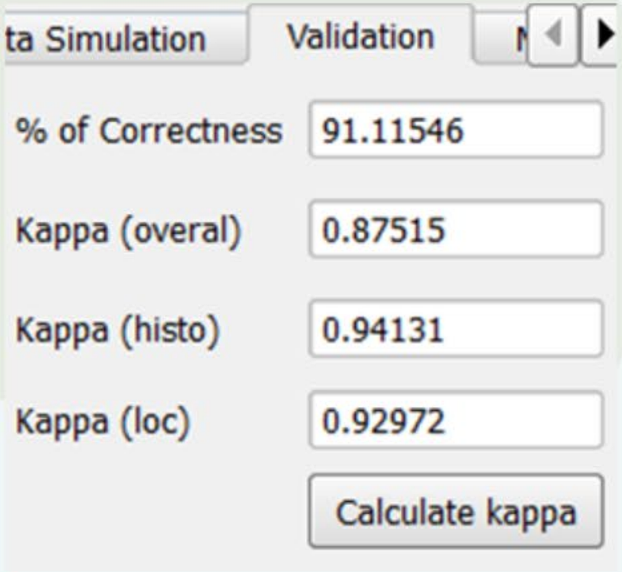
ANN EVALUATION



FIRST TRIAL



SECOND TRIAL



THIRD TRIAL

LAND USE PREDICTION RESULT

Conducted using QGIS with MOLUSCE

Flood prediction was done based on the land use changes predicted for future years

A total of 3 trials of the ANN – MLP evaluations were conducted

First trial

- Accuracy: 54.08%
- Kappa: 0.45

Second trial

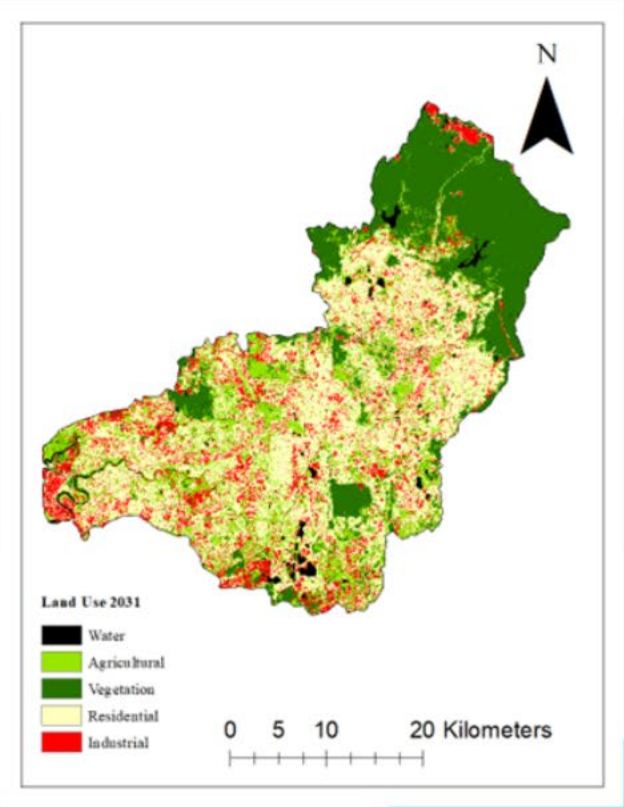
- Accuracy: 42.49%
- Kappa: 0.21

Reclassification of the land use in ArcGIS was done

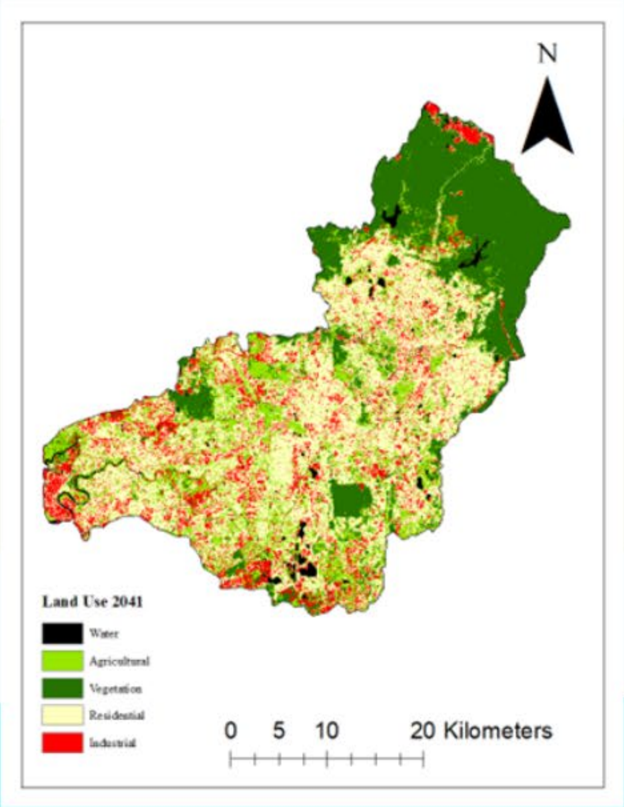
Third trial

- Accuracy: 91.12%
- Kappa of 0.93

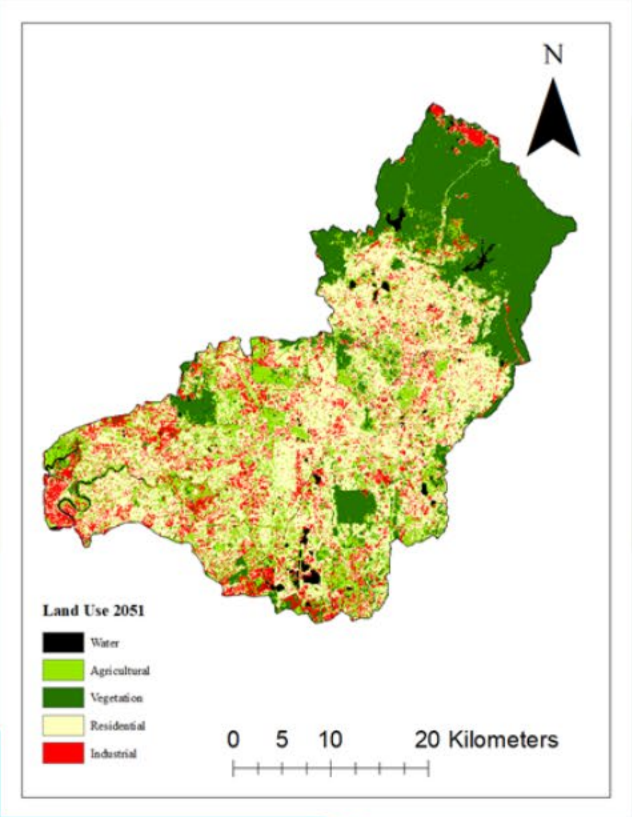
LAND USE PREDICTION RESULT



LAND USE 2031

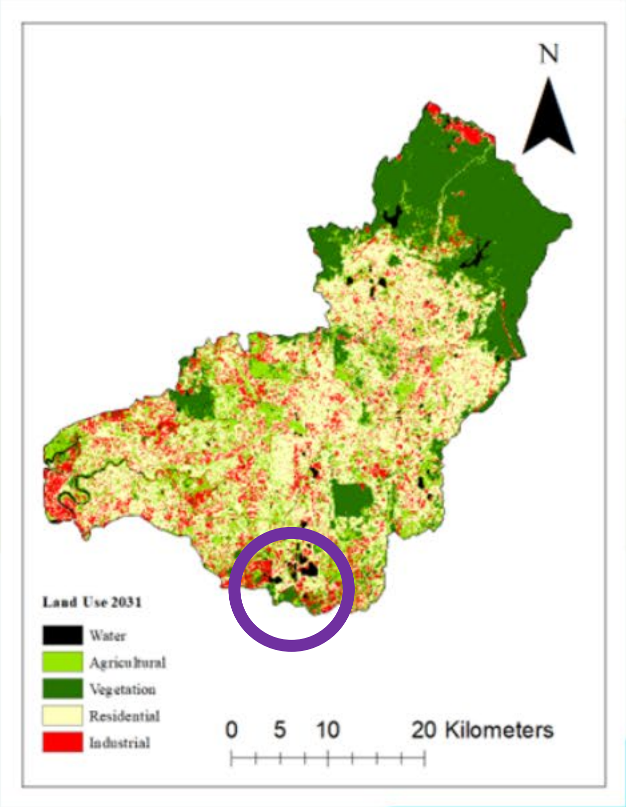


LAND USE 2041

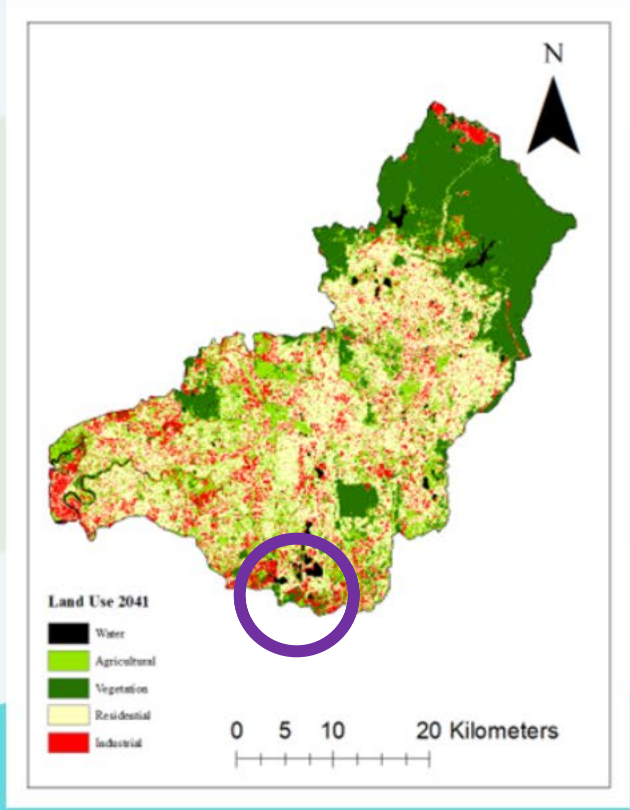


LAND USE 2051

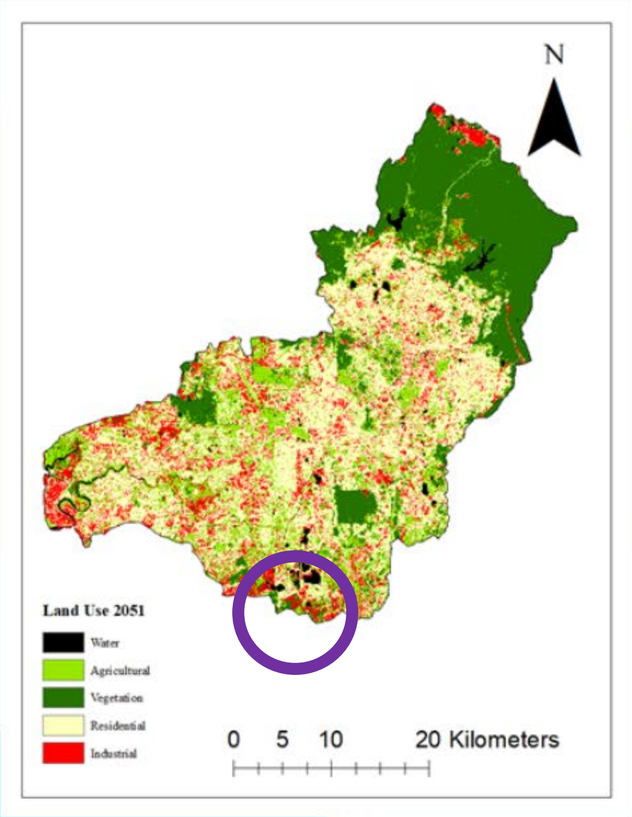
LAND USE PREDICTION RESULT



LAND USE 2031

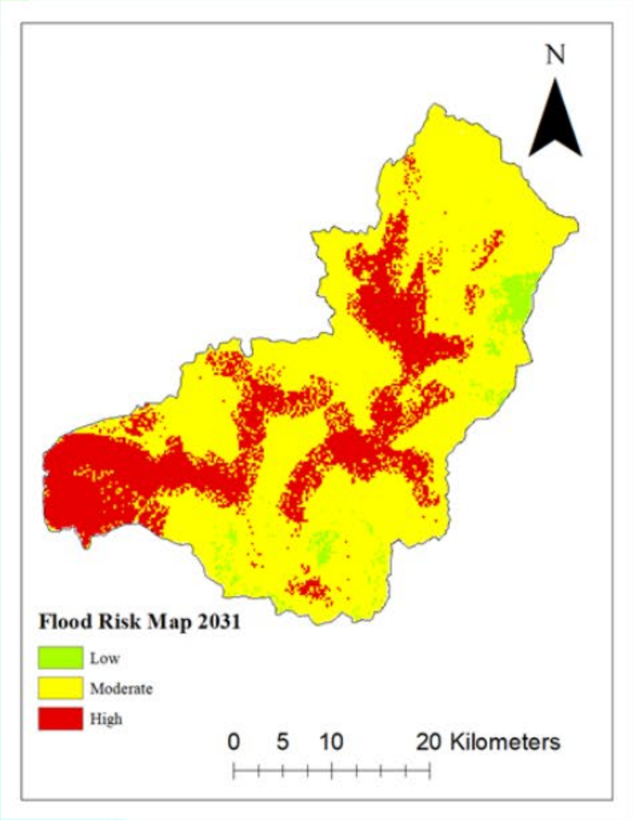


LAND USE 2041

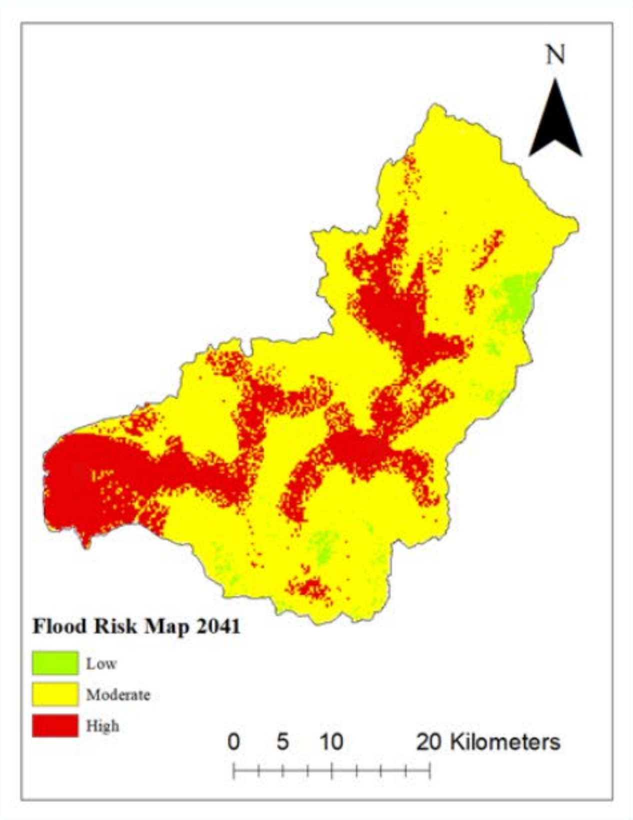


LAND USE 2051

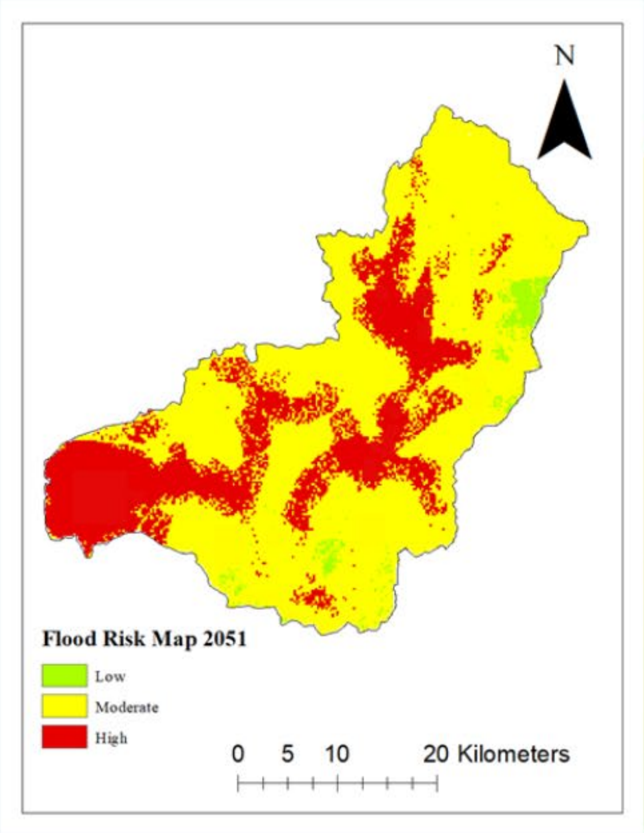
FLOOD PREDICTION RESULT



FLOOD RISK MAP 2031

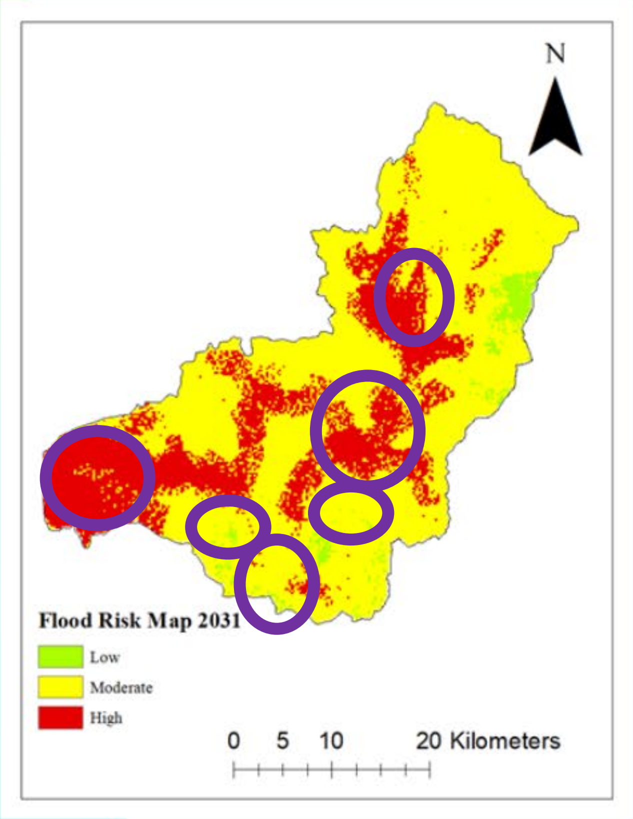


FLOOD RISK MAP 2041

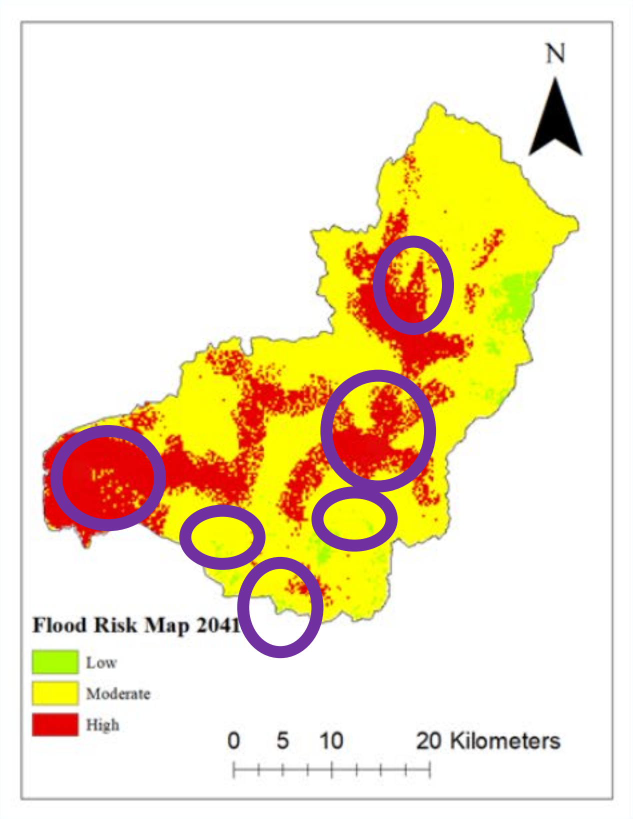


FLOOD RISK MAP 2051

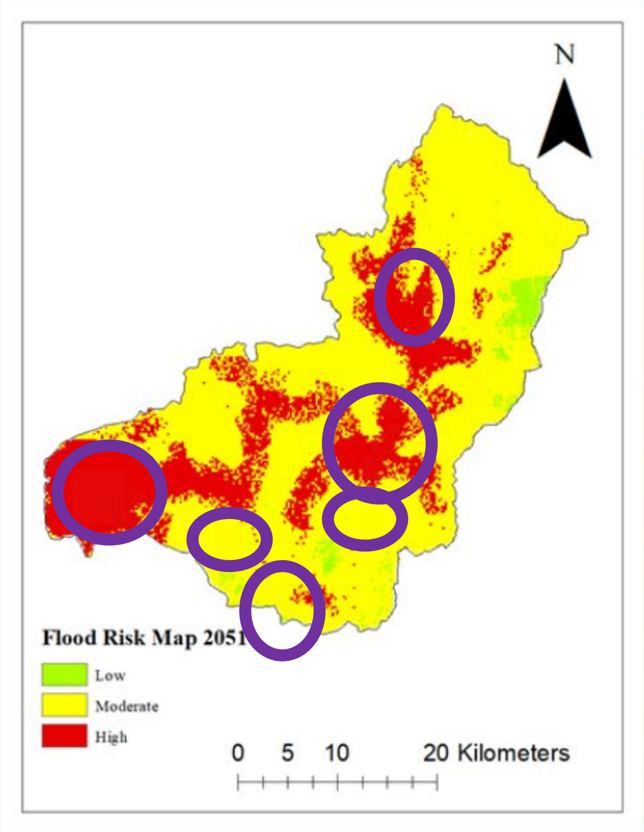
FLOOD PREDICTION RESULT



FLOOD RISK MAP 2031



FLOOD RISK MAP 2041



FLOOD RISK MAP 2051

FLOOD PREDICTION RESULT

Flood magnitude of the basin :-

- **Green areas (the lowest risk flood zone):**
    - these parts of the basin are far from the river and its attributes
    - located near to the upstream of the attributes of the river

- **Yellow areas (moderate risk flood zone):**
    - Majority of areas have 50% of chances that might be affected by flooding
    - near to the main river and its attributes but not along the river

- **Red areas (the highest risk flood zone):**
    - very close to and along the main river and its attributes
    - mostly located at the middle and downstream of the Klang River

FLOOD PREDICTION RESULT

The results have shown that

- Majority of the high flood risk zones are allocated at the surrounding of the downstream of Klang River
- Flood risk area follows the path of the Klang River and its attributes
- The river channels are the major reason and major path for flooding events to occur in a tropical region

CONCLUSION

- The results of this study were the investigation and spatial analysis of flood causative factors, the land use prediction for future years using ANN – MLP technique, and the development of flood risk maps for the year 2031, year 2041, and year 2051.

- The flood risk maps developed shown the prone flood area in the basin through the compilation of the spatial layers of all 10 flood causative factors
  - elevation, slope, land use, rainfall, drainage density, geology, curvature, Topographical Wetness Index (TWI), Normalized Difference Vegetation Index (NDVI), and flow accumulation.

RECOMMENDATIONS

1

Flood Causative Factors:

- decrease number of factors
- only choose the factors that have higher effects in causing flooding at that study area

2

Data Collected:

- use a shorter range of pass year data
- 10 years interval (2011-2021) to 5 years interval (2016-2021)
- help to determine small changes that happened within a few years and produce a more precise result

3

Software Used:

- use MATLAB for ANN modelling will perform a more precise calculations in predicting future conditions and hence produce more accurate result
- big and complicated data can be handled easily

THANK YOU

QNA SESSION