Journal of Applied Geology, vol. 8(1), 2023, pp. 98–121
DOI: http://dx.doi.org/10.22146/jag.92417

Flood Disaster Prediction Model Using Long Short-Term Memory (LSTM) in Pekalongan, Central Java

Muhammad Asrofi[1], Muhammad Rizqy Septyandy[1,3], and Tito Latif Indra*[2]

[1]*Program Study of Geology, Departement of Geoscience, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia*
[2]*Department of Geography, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia*
[3]*Program Study of Geological Engineering, Faculty of Engineering, Universitas Mulawarman, Samarinda, Indonesia*

Received: January 1, 2024 | Accepted: April 16, 2024 | Published online: May 31, 2024

ABSTRACT. Pekalongan is located in the northern part of Java Island, directly adjacent to the sea in the north. Natural disasters that often occur in Pekalongan are floods, especially in the north of the area, which has a height of 0 meters above sea level. In addition, Pekalongan also has a relatively low land slope of around 0–5%, which makes it challenging to distribute water and construct drainage. This study aims to be able to perform predictive modeling of flood-prone areas for the next five years. This study used eight parameters: rainfall, elevation, slope, distance to the river, distance to the sea, groundwater table to surface, soil type, and land use. This research used the Long Short-Term Memory (LSTM) method to predict rainfall parameters using the Python programming language with Jupyter Notebook software. Later, the data will be used as training and test data. Training data testing and tests are conducted to find the minimum failure or error value. The weight scoring method is carried out on each parameter to indicate areas with a high flood vulnerability level. Then, an overlay is carried out on each parameter. The accuracy results of the LSTM model show that the r2 value in Zone 1 is 92 %, and the r2 value in Zone 2 is 97.4 %. Pekalongan has a medium to very high vulnerability level, with a dominant high vulnerability level. The very high level of vulnerability is prevalent in the northern part of the research area, which is directly adjacent to the sea or in the North Pekalongan District. Floods that occur in the northern part of the study area are not only due to high levels of rainfall but can also occur due to the inflow of seawater towards the mainland resulting from high tides and high sea waves. The southern region of the study area has a smaller vulnerability level than the northern region, which has a medium to high vulnerability level.

Keywords: Flood · Hazard · Precipitation · LSTM · Rainfall.

1 INTRODUCTION

Pekalongan has rainfall that tends to fluctuate. These rainfall fluctuations can affect daily life. Besides disrupting daily activities and life, high rain can cause natural disasters like floods. Pekalongan is one of the areas in Indonesia where floods often occur, especially in the northern area, which has a height of 0 meters above sea level. In addition, Pekalongan also has a relatively low land slope of around 0–5 %, which makes it challenging to distribute water and construct drainage. Not only does it come from rain, but flooding in Pekalongan can also be caused by rising sea levels, which inundate some of the settlements on the coast and are commonly referred to as tidal floods (Salim, 2018).

*Corresponding author: T.L. INDRA, UI-Campus, Beji, Depok, West Java 16424, Indonesia. E-mail: tito.latif@sci.ui.ac.id

In this study, historical rainfall data for the last 40 years (1981–2021) will be modeled using the Long Short-Term Memory (LSTM) method to predict the intensity of rainfall that will occur in the future. The novelty of this research is identifying a flood prediction model for the research object and analyzing the accuracy of the LSTM model based on rainfall data, which will later be used to analyze flood-prone areas in Pekalongan City. This model approach has never been carried out on research objects. It is hoped that the results of this research can be used as a reference for the surrounding community in carrying out practical activities. Besides that, it is also used as information for the authorities in carrying out sustainable flood disaster management. Based on this research, the author wants to make a flood risk map in Pekalongan to determine the distribution of areas at risk of floods. The objectives of this study are: 1) Identify and determine the LSTM model related to flood prediction based on rainfall data in the research object. 2) Identify and determine the accuracy of the LSTM model in predicting floods in the research object. 3) Generate vulnerability areas using flood predictions with the LSTM model in the research object in the next few years. 4) Identify the advantages and disadvantages of using the LSTM model in predicting flood disasters in the research object.

2 Geological Setting

Pekalongan is located in the northern part of Java Island, directly adjacent to the sea in the north. The geographical position of Pekalongan is between 6°50′42"–6°55′44" South Latitude and 109°37′55"–109°42′19" East Longitude (Figure 1). Pekalongan is administratively divided into four sub-districts: North Pekalongan, East Pekalongan, West Pekalongan, and South Pekalongan. The population of Pekalongan is 307,150 people or around 6,787.85 people/km2 of population density (BPS Pekalongan, 2021).

Physiographically, Pekalongan is located on the island of Java, precisely in the province of Central Java. Pekalongan is in the northern part of the island of Java and is directly adjacent to the ocean. Based on the physiographical location (Figure 2), Pekalongan is located in the North Coast Alluvial Plain Zone, the area is around the Quaternary Volcano Zone and is located in the northern part of the Bogor Anticlinorium Zone (Bemmelen, 1949).

Based on the geological map of Banjar and Pekalongan (Condon, 1975). The stratigraphy comprises Pekalongan, whose surroundings are sedimentary rock, volcanic rock, and intrusive igneous rock. In the Pekalongan area, on the dominant surface, there are alluvium deposits (Figure 3), which are composed of river deposits and coastal deposits and morphologically have low elevations (Van Zuidam, 1985).

3 Methodology

This research uses parameters that can influence the occurrence of flood disasters in Pekalongan City. Analysis of the level of vulnerability to flood disasters in Pekalongan City using parameters: precipitation, land use, groundwater level, slope, soil type, elevation, river buffer, and marine buffer (Bretschneider & Wybro, 1977; Kusumo & Nursari, 2016). The rainfall analysis used in this study uses precipitation data covering a period of about 40 years from the period 1981–2021. In addition, secondary data, including solar radiation intensity, wind speed, temperature, and humidity, correlate with precipitation. The data that has been collected is then processed using a program with the Long Short-Term Memory (LSTM) method using the Python programming language with the Jupyter Notebook. The results of each parameter will be overlaid using ArcGIS software to obtain areas prone to flooding in Pekalongan City.

3.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a variant or part of a Recurrent Neural Network (RNN). Long Short-Term Memory (LSTM) was first discovered by Hochreiter and Schmidhuber in 1997, which then, over time, made LSTM even more refined. In general, LSTMs can learn patterns using data sets that were sequential in the past. Conceptually, LSTM is more complicated than RNN because, in RNN, each cell only contains one layer of neurons using the tanh activation function. Meanwhile, in LSTM, each cell can have more than one layer of neurons, and in LSTM, there are four layers of neurons called gates11. In processing data, the LSTM

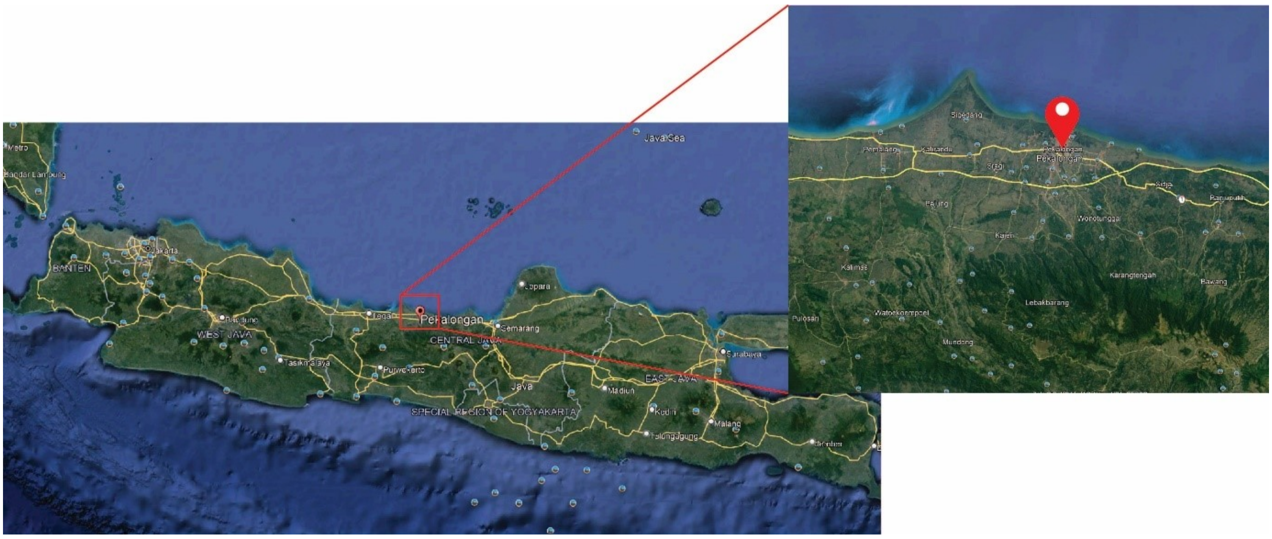

FIGURE 1. The study area of Pekalongan, Central Java (Google Earth).

cell will produce two outputs. The first output will reveal the actual result of the input made to that cell, while the other is the output result of the forward data going to the next cell or cell state (Ct). The LSTM method focuses on correlating cell state with element-wise operations performed on each cell.

LSTM has three gates in its application: Forget Gate, Input Gate, and Output Gate. The initial stage carried out in pre-processing the data is to enter the program library, which will later be used in processing LSTM data, then input data that is visualized in graphical form, which will subsequently make it easier to read the model, the data used in this study is historical data precipitation in Pekalongan for 40 years (January 1981 – December 2021). The data that has been input is then normalized using the min-max scaler from sklearn. This technique changes the original value of the precipitation data into a range interval. In this case, the data value is changed to 0–1. The next stage is to train the LSTM model, which is used to train experimental data that has been input. The LSTM model was trained using 700 epochs in 64 batches, and the author adjusted the number of epochs to get the best results. The result of running the LSTM model is to get a relatively small loss value. After training the LSTM model and getting a good model, the predicted values are denormalized in the form of a range interval. The prediction values obtained after training the experimental data then need to be evaluated to determine the level of accuracy obtained by the predicted value against the actual value. After the LSTM model is run and performs well, the predicted value for the coming year can be used. This study predicts precipitation for the next five years from 2022–2026. The prediction values for the coming year produce 60 precipitation data, each representing one month. The prediction results will be visualized as a map using the IDW interpolation technique in ArcGIS software. The prediction results are expected to predict the precipitation that will occur in Pekalongan for the next five years.

3.2 Scoring and Weighting

This study used eight parameters in analyzing the level of flood hazard in the study area, which consisted of parameters of precipitation, land use, groundwater level, slope, soil type, elevation, river buffer, and marine buffer (Bretschneider & Wybro, 1977; Kusumo & Nursari, 2016). The rainfall analysis used in this study uses precipitation data, which covers a time span of about 40 years from 1981 to 2021. The level of rainfall can affect the occurrence of floods, where the higher the rainfall in an area, the higher the potential for flooding. The slope can affect the movement of water, where water will move on a high slope towards a low slope. Flat slopes will have a higher potential for flooding than areas with steeper slopes. Soil type or lithology can affect water absorption or infiltration in the soil layer. The finer the texture of the soil, the more difficult it is for the soil to absorb water compared to soil with a coarse texture, so areas with fine soil characteristics will have a higher potential for flooding. Land use can affect the rate of water from rain. Areas that easily allow rainwater to flow into rivers can have a higher potential for flooding due to the lack of infiltration of rainwater in the area, and areas that are predominantly filled with plant vegetation have a lower potential for flooding because much rainwater will infiltrate. Areas with higher elevations tend to have smaller flood potential than those with lower elevations because water moves from high areas to lower areas. Areas close to the river will have a greater potential for flooding than areas further away from the river. This condition can occur when the volume of water in the river exceeds its capacity so that the water can overflow. Areas closer to the coast will have a greater potential for flooding than areas farther away. The condition can occur when there is a rise in sea level or high sea waves. Areas with a small distance between the groundwater table and the surface have a higher potential for flooding than areas with a relatively large distance between the groundwater table and the surface. The closer the groundwater table is to the surface, the lower the holding capacity of the soil layer (Table 1).

Based on the eight parameters, weighting is carried out, where the weighted value of each parameter is carried out qualitatively. The weighting results in this study refer to the weighting done by Kusumo & Nursari (2016), which the authors then modified to adjust the parameters used in this study. The greater the influence of the parameter on flooding, the higher the weight will be (Table 2).

This analysis was carried out to classify the level of flood vulnerability based on the total score of all parameters, where each parameter used will be overlayed and the sum of the score values (Equation 1). Based on the total score obtained from all parameters, normalization is carried out to be applied to all flood vulnerability level maps (Equation 2). Areas with a high final total value tend to have high flood vulnerability and vice versa. They are then grouped into five categories based on the value of flood vulnerability (Table 3).

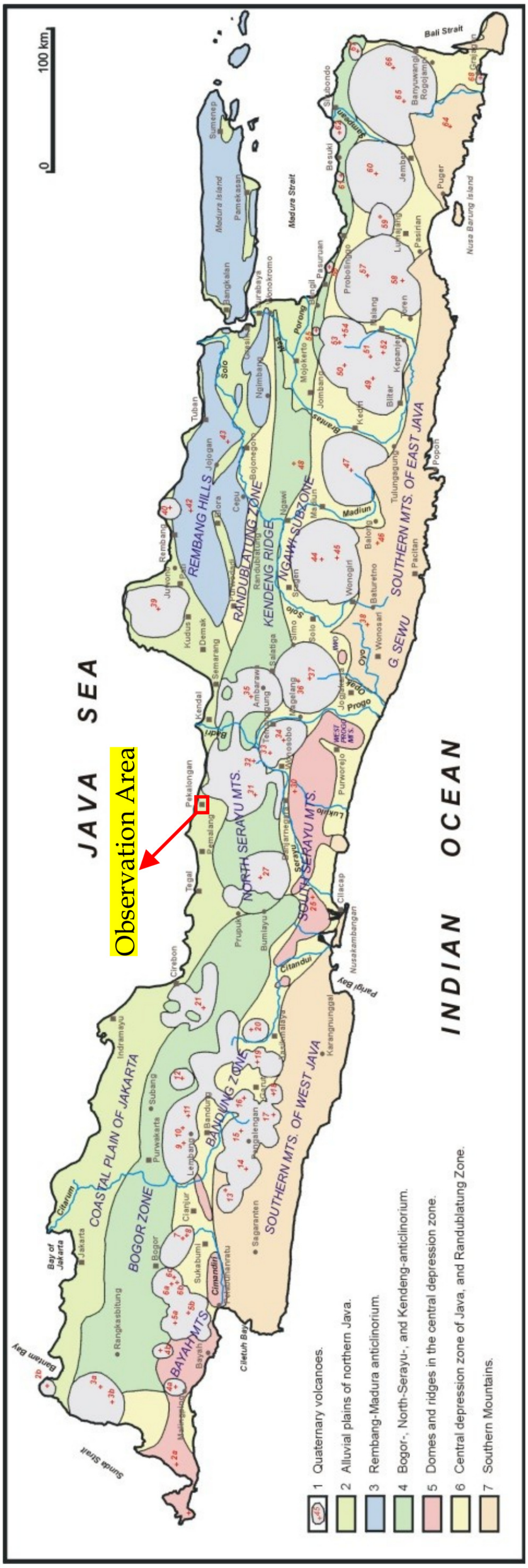

FIGURE 2. Physiography of Central Java and East Java (Bemmelen, 1949) (Setiaji *et al.*, 2016).

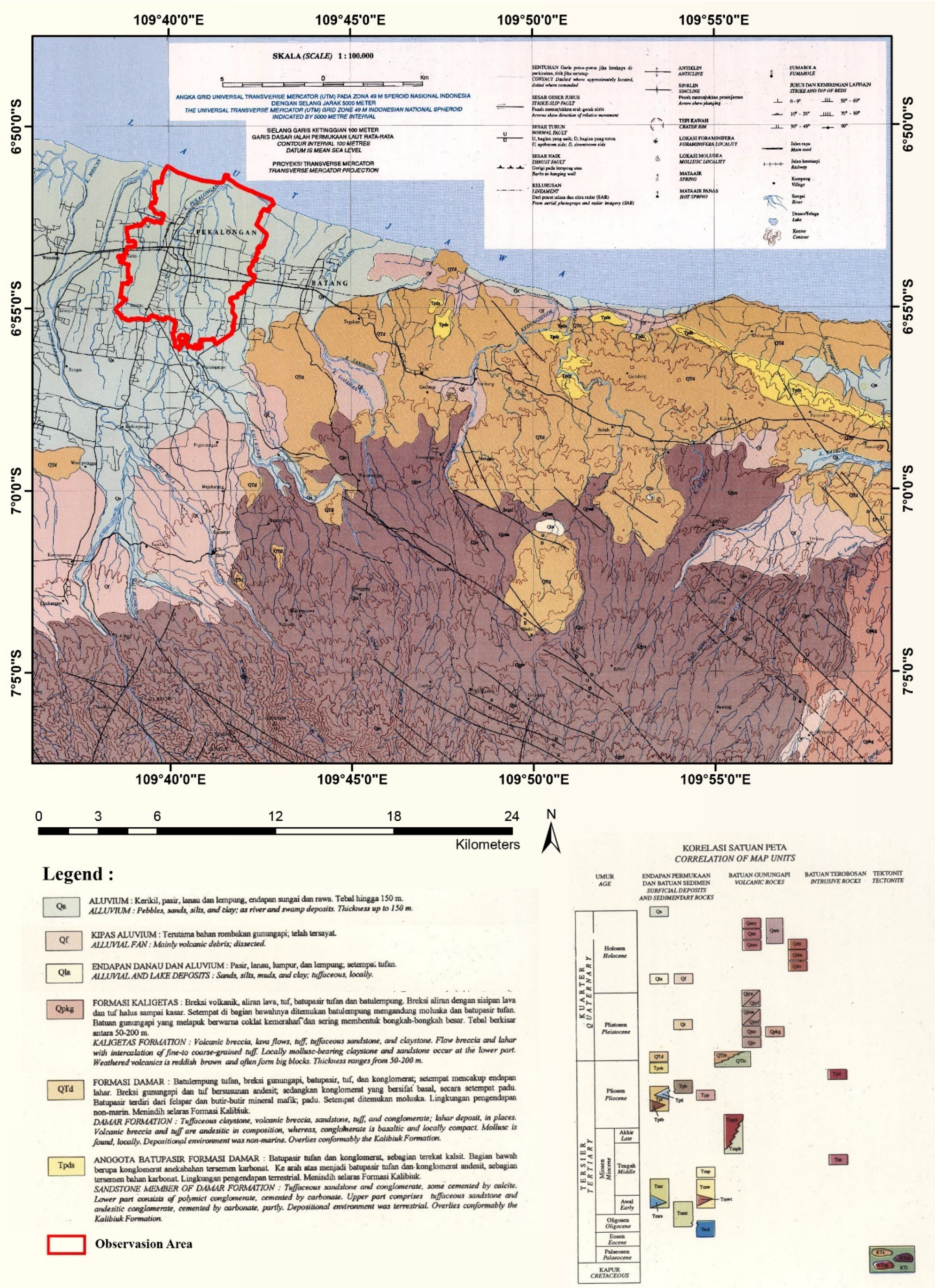

FIGURE 3. Pekalongan on the Sheet Map of Banjarnegara and Pekalongan (Condon, 1975; modified by the author).

TABLE 1. Flood vulnerability parameters (Kusumo & Nursari, 2016; modified by the authors).

Precipitation (mm/month)	Score	Vulnerability
>500	9	Very High
300 – 500	7	High
100 – 300	5	Medium
50 – 100	3	Low
0 – 50	1	Very Low
Slope (%)		
0 – 8	9	Very High
8 – 25	7	High
15 – 25	5	Medium
25 – 40	3	Low
>40	1	Very Low
Soil Type		
Vertisol, Oxisol, Fluvisol	9	Very High
Alfisol, Ultisol, Molisol	7	High
Inceptisol	5	Medium
Histosol	3	Low
Spodosol, Andisol	1	Very Low
Landuse		
Open land - Water bodies - Ponds	9	Very High
Settlement - Rice Field	7	High
Plantation - Farmland	5	Medium
Mixed garden - Shrubs	3	Low
Forest	1	Very Low
Elevation (m)		
0 – 25	9	Very High
21 – 50	7	High
51 – 100	5	Medium
101 – 300	3	Low
>300	1	Very Low
River Buffer (m)		
0 – 25	9	Very High
25 – 50	7	High
50 – 75	5	Medium
75 – 100	3	Low
>100	1	Very Low
Marine Buffer (m)		
0 – 556	9	Very High
556 – 1400	7	High
1400 – 2404	5	Medium
2404 – 3528	3	Low
>3528	1	Very Low
Ground Water Level		
0 – 1	9	Very High
1 – 2	7	High
2 – 6	5	Medium
6 – 10	3	Low
>10	1	Very Low

TABLE 2. The weight for each flood vulnerability parameter (Kusumo & Nursari, 2016; modified by the authors).

Parameters	Point Weight
Precipitation	15
Slope	10
Elevation	15
River Buffer	10
Marine Buffer	10
Ground Water Level	10
Soil Type	10
Landuse	20
Total	100

TABLE 3. Flood hazard class (National Disaster Management Agency (NDMA), 2019; modified by the authors).

Flood Vulnerability Level	Total Value
Very High	801 – 1000
High	601 – 800
Medium	401 – 600
Low	201 – 400
Very Low	0 – 200

$$Ts = \sum_{i=1}^{n} W_i \times X_i \tag{1}$$

where Ts = Total score; W_i = Parameter score; and X_i = Parameter weight

$$Kb = \frac{Ts}{900} \times 1000 \tag{2}$$

where Kb = Flood hazard level; and Ts = Total score.

4 RESULTS AND DISCUSSION

Precipitation is one of the main parameters determining an area's flood vulnerability level. This research uses historical precipitation data in Pekalongan with a 40-year history of precipitation (January 1st, 1981 – December 31st, 2021). In general, there are eight observation points in the research area because the research area is relatively small, so the possibility of getting different precipitation values will be smaller. The total of 8 observation points is divided into two zones with different precipitation values. The observation points (1, 3, and 8) are called Zone 1, and observation points (2, 4, 5, 6, and 7) are called Zone 2 (Figure 4).

The relationship between parameters shows the correlation or relationship between one parameter and another. In this case, the relationship between the parameters used is precipitation, intensity of solar radiation, minimum temperature, and humidity. The results of the correlation between parameters are visualized in the form of a heat map. Two relationships in the heat map reading show a strong correlation (-1 and 1). The relationship between parameters that are correlated and aligned with each other is indicated to have a value close to 1 or 1.

Conversely, the relationship between correlated parameters opposite each other indicates a value close to -1 or -1. The humidity parameter concerning precipitation shows the best parameter correlation. The correlation of these two parameters shows a value of 0.67 (positive, close to 1). This condition can indicate that the relationship between precipitation and humidity parameters is aligned or that the higher the level or value of humidity, the higher the precipitation value. In addition, the relationship between precipitation parameters and minimum temperature also shows a consistent correlation with a value of 0.66 (positive, close to 1). The condition can indicate that the relationship between precipitation parameters and minimum temperature is aligned; the lower the minimum temperature, the higher the precipitation value. The humidity parameter shows the best parameter correlation in zone 2 to precipitation. The correlation of these two parameters shows a value of 0.7 (positive, close to 1). This condition can show that the relationship between precipitation and humidity parameters is in harmony, or the higher the level or value of humidity, the higher the precipitation value. In addition, the relationship between precipitation parameters and minimum temperature also shows a correlation that is aligned by showing a value of 0.59 (positive, close to 1). The condition shows that the relationship between precipitation parameters and minimum temperature is aligned; the lower the minimum temperature, the higher the precipitation value (Figure 5).

LTSM model performance is based on the results of loss and validation loss values. Loss

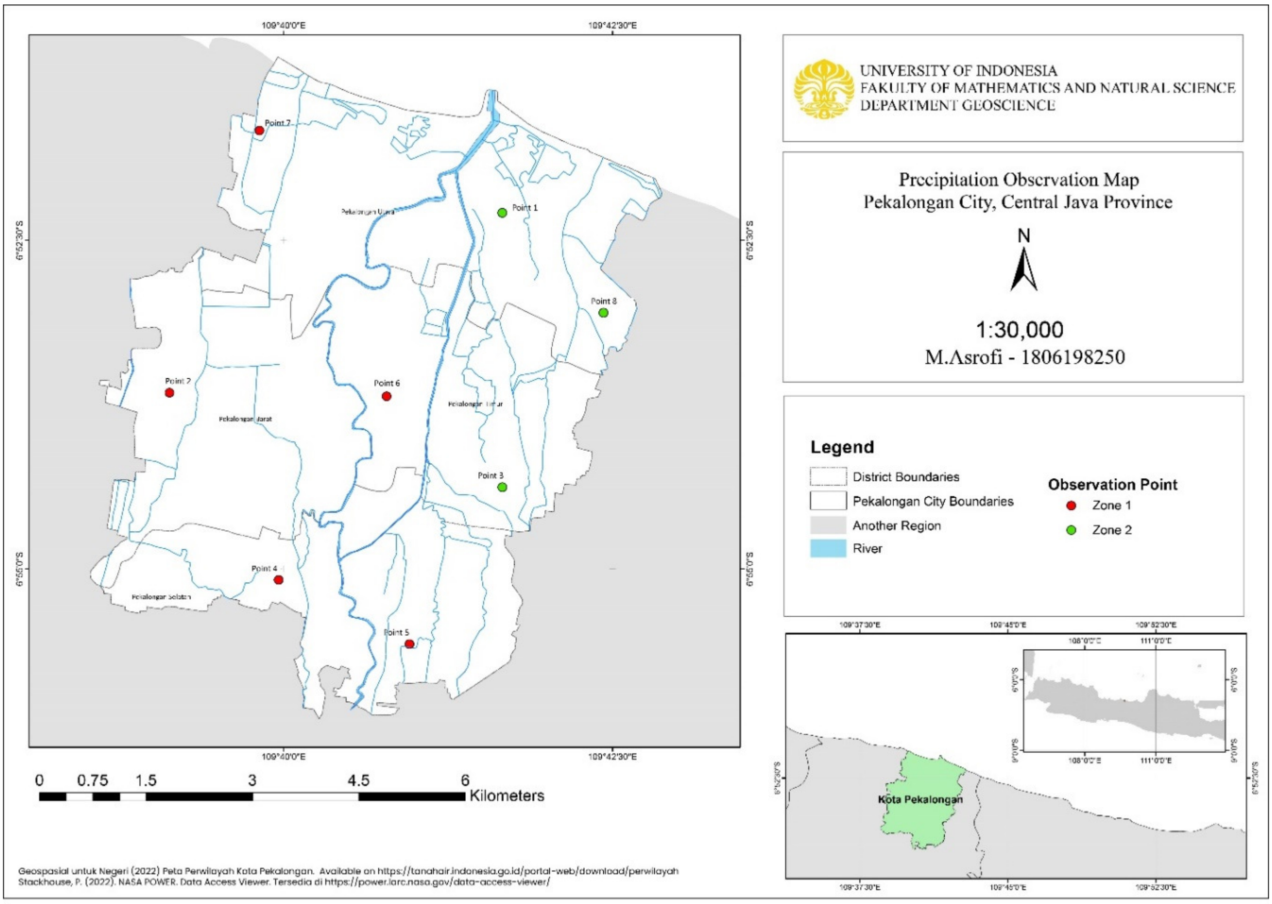

FIGURE 4. Precipitation observation point.

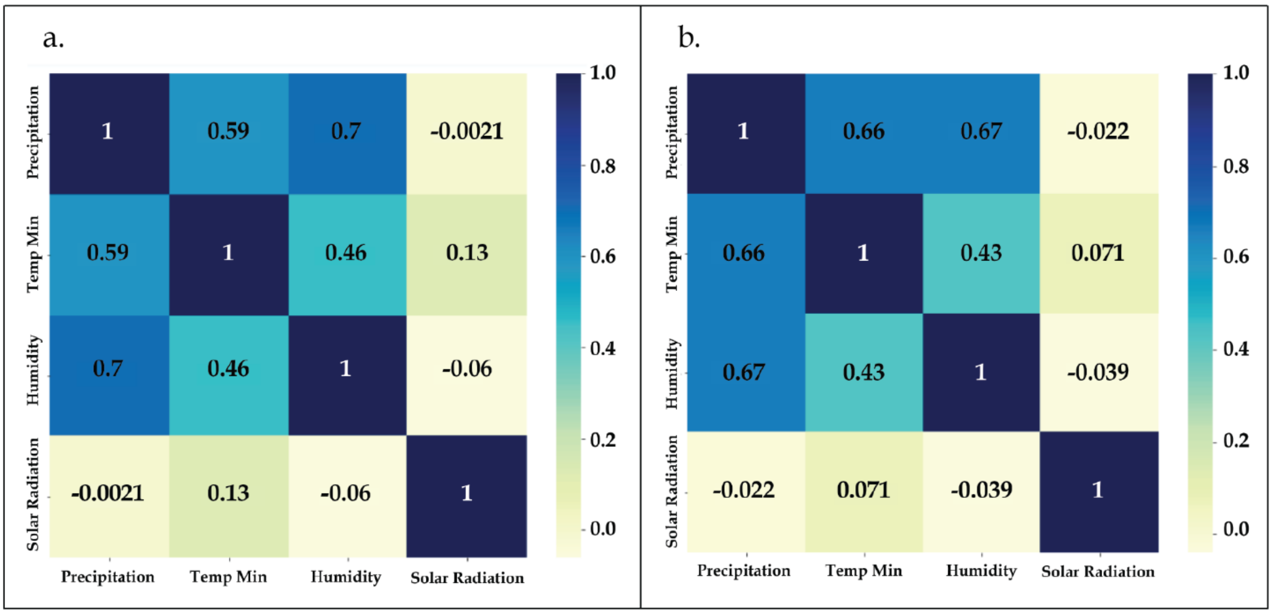

FIGURE 5. (a) Heat map zone 1, (b) Heat map zone 2.

and validation loss values are obtained after running on experimental data, and the results of loss and validation loss values in LSTM model training are then visualized in graphical form (Figure 6). The lower the loss value obtained in training the LSTM model, the better the model will be, and vice versa. The graph of loss and validation loss values shows that the increasing epoch will cause the loss value to be smaller, whereas in this study, using displays a relatively small loss and validation loss value, this can indicate the performance of the model run is good.

The prediction results for the next five years are based on the rainfall zones in the study area, which are divided into Zone 1 and Zone 2. The results of the rainfall prediction in Zone 1 and Zone 2 are then displayed in graphical form to make it easier to read. The graph shows the predicted values for the two zones in the study area. It shows that there is a similar pattern of trends in the history of previous rainfall (1981 – 2021) (Figure 7; Table 4). The result of the model accuracy was obtained using the coefficient of determination (r2). The prediction results for the actual data show that the value of r2 in zone 1 shows the independent variable (predicted value) 92,1% can explain the dependent variable (actual value), and the value of r2 in zone 2 shows the independent variable (predicted value) 97,4% can explain the dependent variable (actual value) (Figure 8).

TABLE 4. Evaluation of the LSTM model in Zone 1 and Zone 2.

Metric Type	Zone 1	Zone 2
RMSE	1.26	0.69
MAE	0.79	0.39
MSE	1.59	0.47

The results of rainfall predictions for the next five years (January 2022 - December 2026) use a prediction model that has been made before, whereas the LSTM model that has been made before can produce predicted values for the next five years (January 2022 - December 2026). The prediction results are based on the rainfall zone in the research area, which is divided into Zone 1 and Zone 2 (Figure 9). The results of rainfall prediction (January 2022 - December 2026) are then processed to produce rainfall prediction maps in the study area. Four months will be taken for processing each year (January, June, October, and December). The predicted rainfall value is processed using the IDW interpolation technique on ArcGIS to display rainfall distribution in the study area (Figure 10). The interpolated rainfall map from the predicted value will later be used for precipitation parameters, which are then overlaid with other parameters such as land use, groundwater level, slope, soil type, elevation, river buffer, and marine buffer (Figure 11) in determining flood hazard areas in the research area. (Figure 12).

The actual rainfall value in Pekalongan in Zone 1 and Zone 2 is then validated using rainfall data at the nearest observation station. The actual value of zone 1 uses rainfall data at the Tegal Maritime Meteorological Station, and zone 2 uses rainfall data at the Ahmad Yani Meteorological Station. This validation was carried out to determine the effect of the rain in the Tegal Regency and the Semarang on the rainfall in the Pekalongan. Comparing the actual rainfall value of Pekalongan (Zone 1) with the rainfall value in Tegal Regency measured by the Meteorological, Climatological, and Geophysical Agency (BMKG) at the Tegal Maritime Meteorological Station shows a similar rainfall pattern (Figure 13a). The highest rainfall occurs in February, and the lowest occurs from May to October (Table 5). The actual rainfall value of Pekalongan in Zone 1 and the rainfall value of the Tegal Maritime Meteorological Station after evaluation and visualization in the graph, resulting in a coefficient of determination (r2) of 0.901 or 90.1%, these results indicate that the occurrence of rain that occurs in Tegal Regency can affect the event of rain in Pekalongan (Zone 1) by 90.1% (Figure 13b; Table 6).

Meanwhile, the comparison of the actual rainfall value of Pekalongan (Zone 2) with the rainfall value of Semarang measured by the Meteorological, Climatological, and Geophysical Agency (BMKG) at Ahmad Yani Meteorological Station shows a similar rainfall pattern. The highest rainfall occurs in February, and the lowest in July (Table 5). The actual rainfall value of Pekalongan in Zone 2 and the rainfall value of Ahmad Yani Meteorology after evaluation and

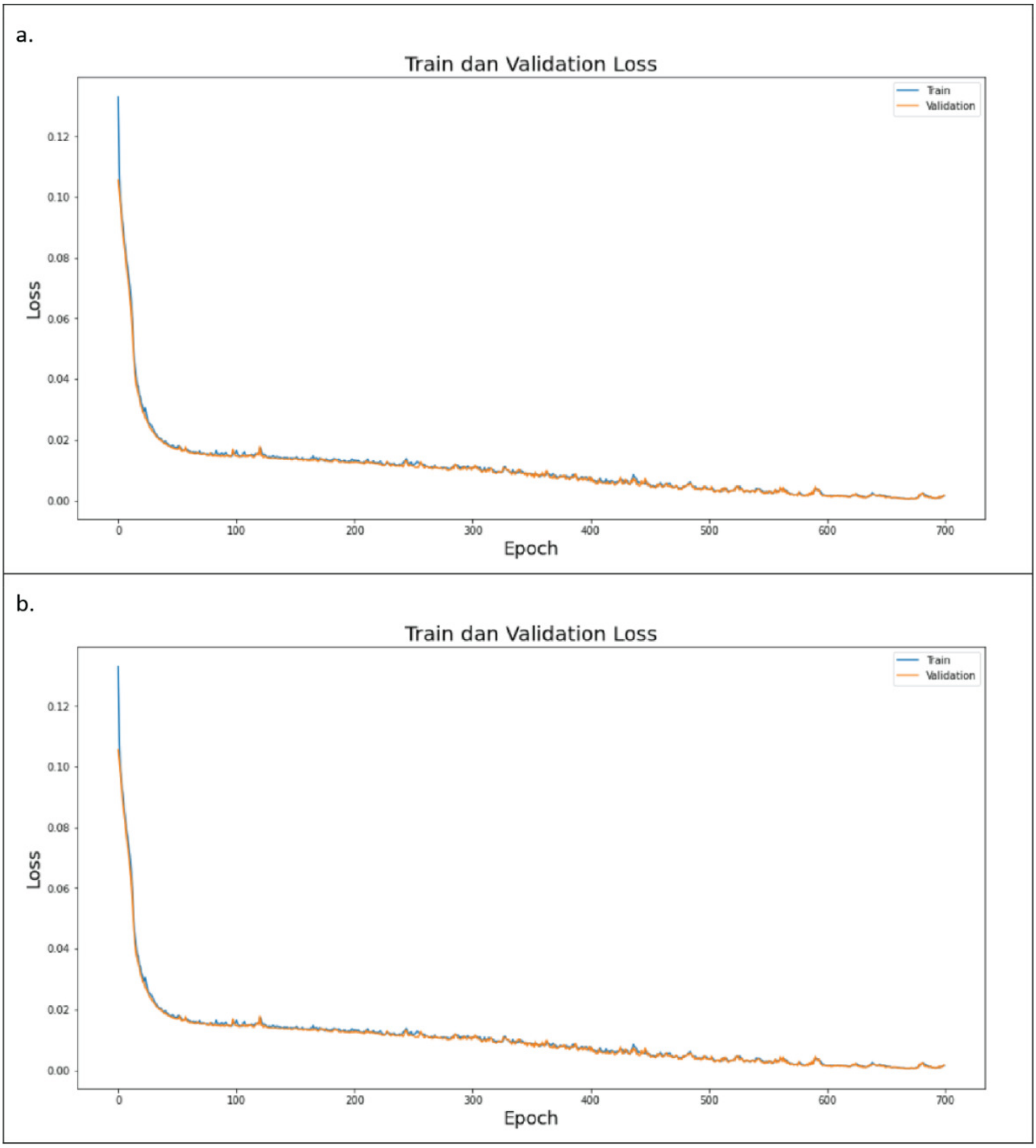

FIGURE 6. (a) Train and validation loss zone 1, (b) Train and validation loss zone.

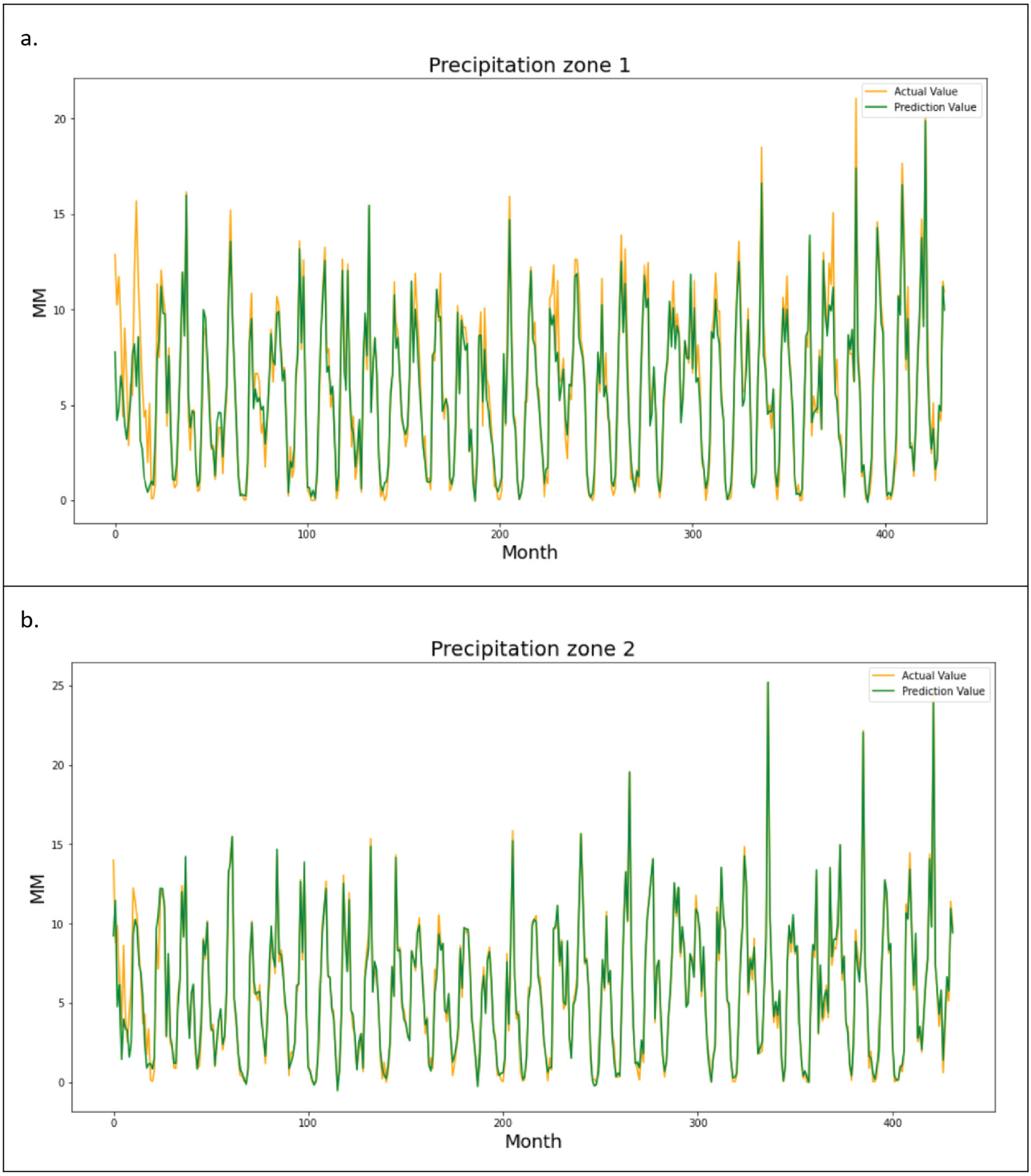

FIGURE 7. (a) LSTM model performance in Zone 1; (b) LSTM model performance in Zone 2.

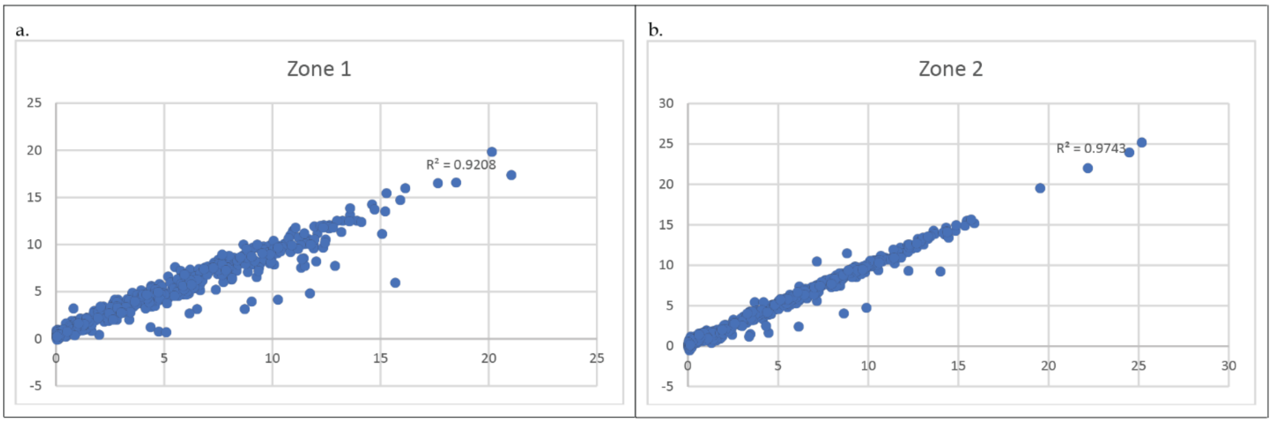

FIGURE 8. (a) The coefficient of determination (r2) of the LSTM model in zone 1; (b) The coefficient of determination (r2) of the LSTM model in zone 2.

TABLE 5. Comparison of rainfall values (January to December 2021) for Tegal (Tegal Maritime Meteorological Station) with actual rainfall values for Pekalongan Zone 1 and for Semarang (Ahmad Yani Meteorological Station) with actual rainfall values for Pekalongan Zone 2.

Month	Tegal Regency (mm/day)	Pekalongan Zone 1 (mm/day)	Semarang (mm/day)	Pekalongan Zone 2 (mm/day)
January	7.6	9.53	6.95	9.87
February	16.9	20.15	25.73	24.48
March	11.56	9.36	4.72	7.31
April	5.64	4.76	4.39	5.52
May	1.16	2.46	5.97	3.54
June	1.78	5.12	5.1	5.53
July	1.17	1.04	0.59	0.61
August	1.5	2.63	5.56	3.27
September	2.7	4.7	5.65	6.04
October	1.75	4.18	4.13	5.12
November	8.74	11.46	11.38	11.42
December	9.14	10.9	5.26	9.66

TABLE 6. Evaluation of the rainfall value for Tegal with the actual rainfall value for Pekalongan Zone 1 and Semarang with actual rainfall value for Pekalongan Zone 2 (from January to December 2021 and from January to October 2022).

Metric Type	January 2021 – December 2021		January 2022 – October 2022	
	Tegal vs Pekalongan Zone 1	Semarang vs Pekalongan Zone 2	Tegal vs Pekalongan Zone 1	Semarang vs Pekalongan Zone 2
RMSE	2.13	2.04	2.47	2.47
MAE	1.92	1.57	6.10	6.10
MSE	4.54	4.16	2.18	2.18

visualized in the graph, resulting in a coefficient of determination (r2) of 0.8966 or 89.66%, these results indicate that the occurrence of rain that occurs in Semarang can affect the occurrence of rain in Pekalongan (Zone 2) by 89.66% (Figure 13c; Table 6).

The results of the predicted value in 2022 (January to October) were validated using rainfall data at the nearest observation station. Pre-

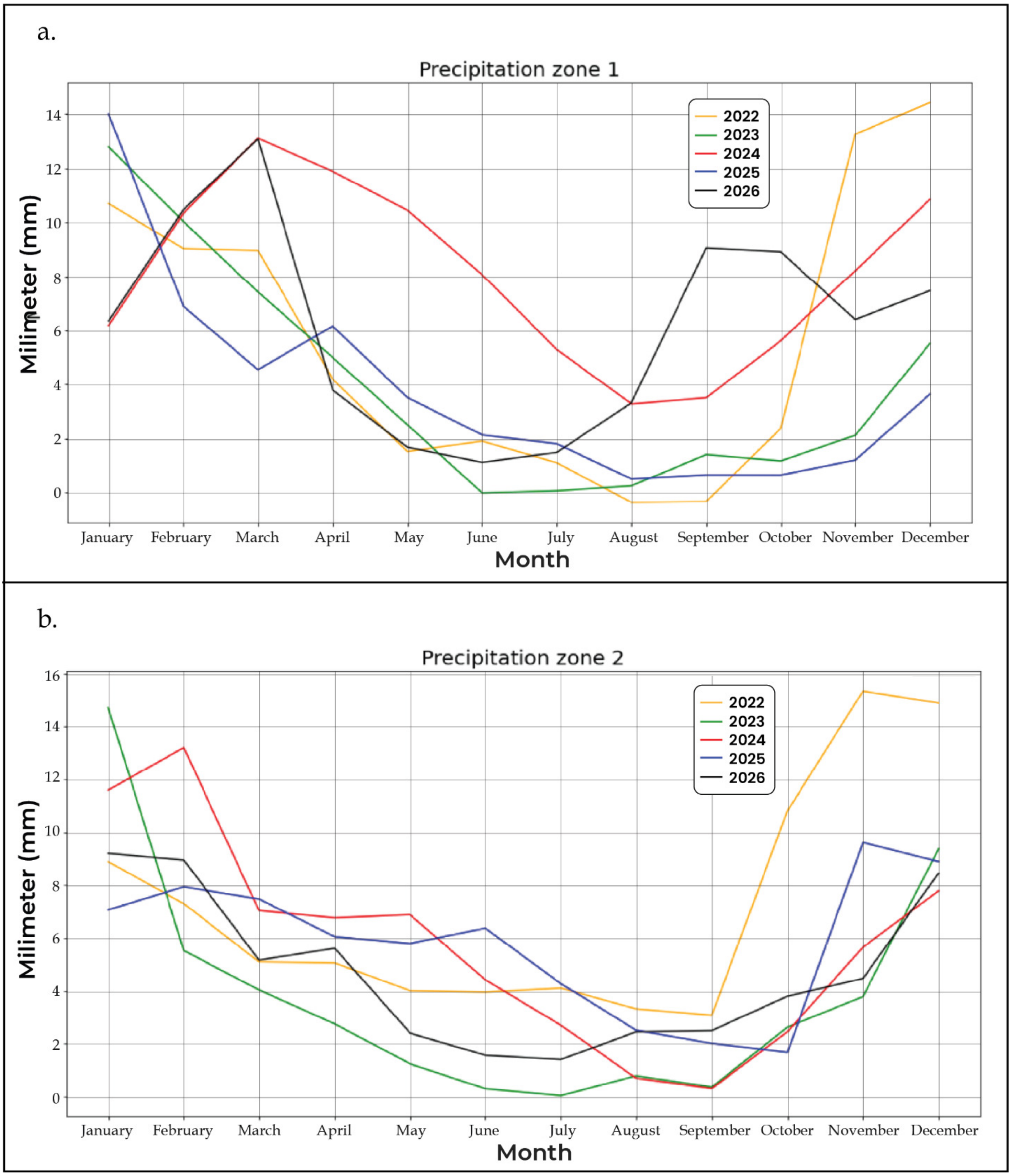

FIGURE 9. (a) Prediction graph of the LSTM model in Zone 1; (b) Prediction graph of the LSTM model in Zone 2.

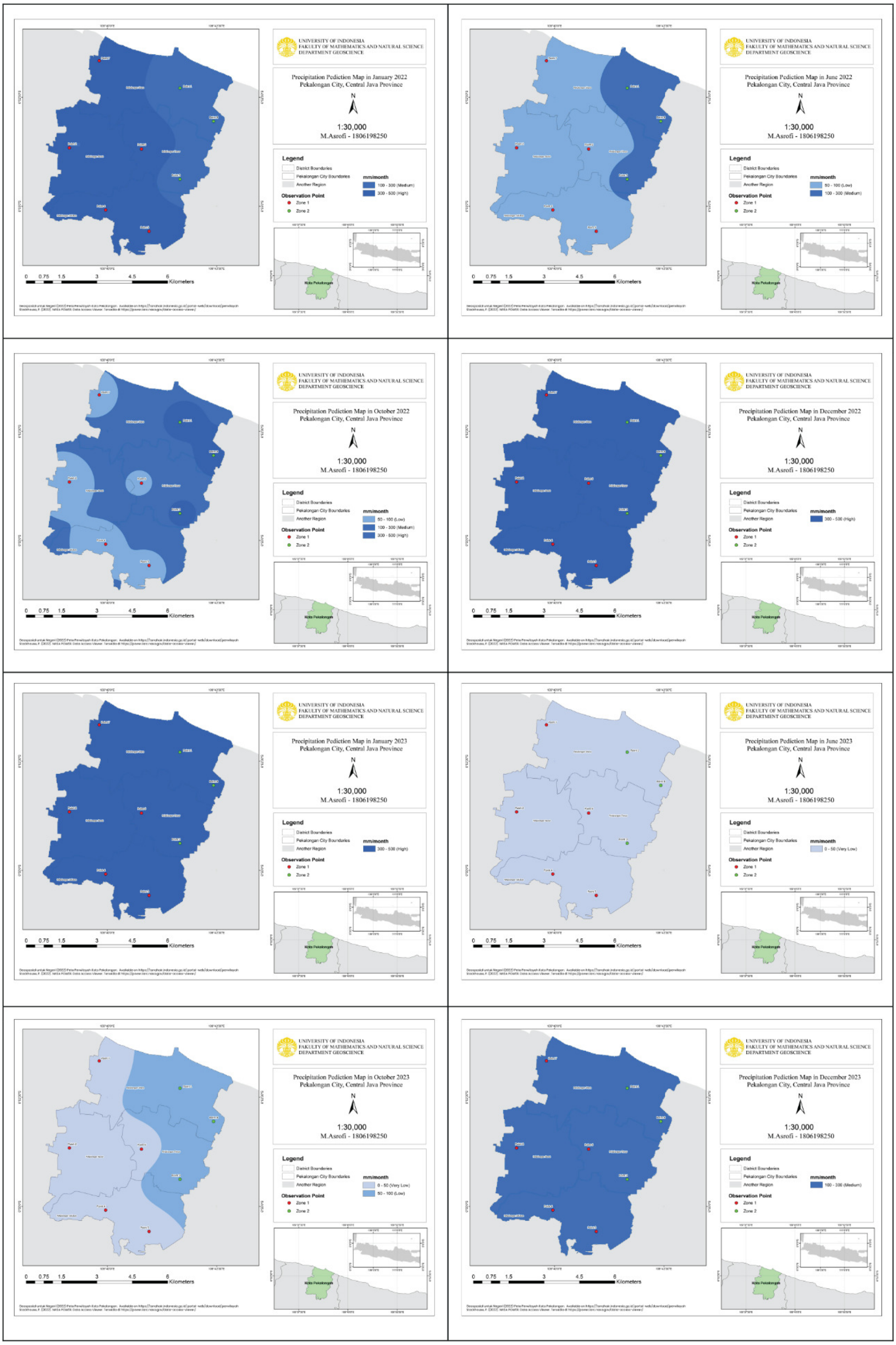

FIGURE 10. Precipitation prediction map (2022–2026).

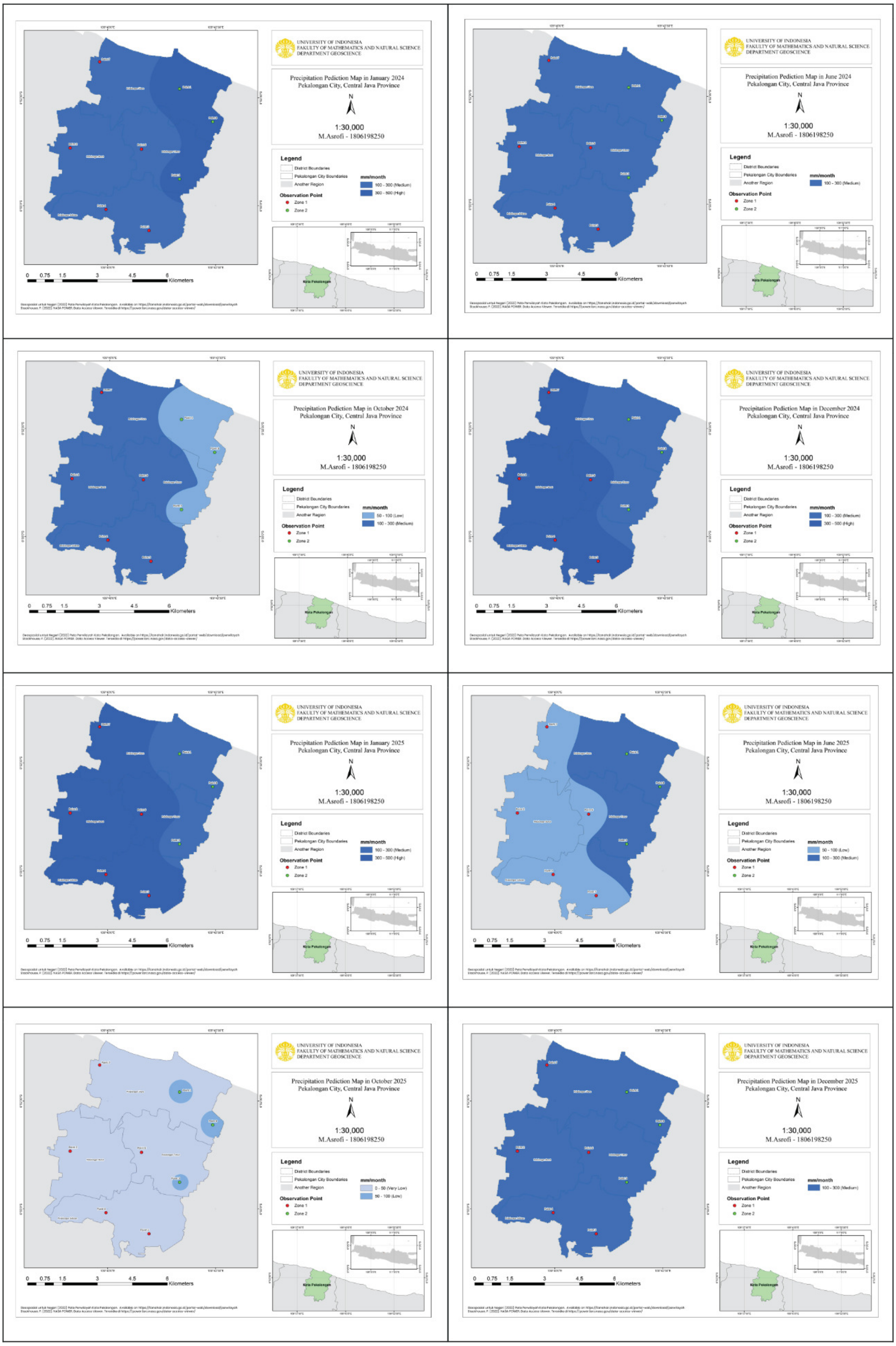

FIGURE 10. Continue...

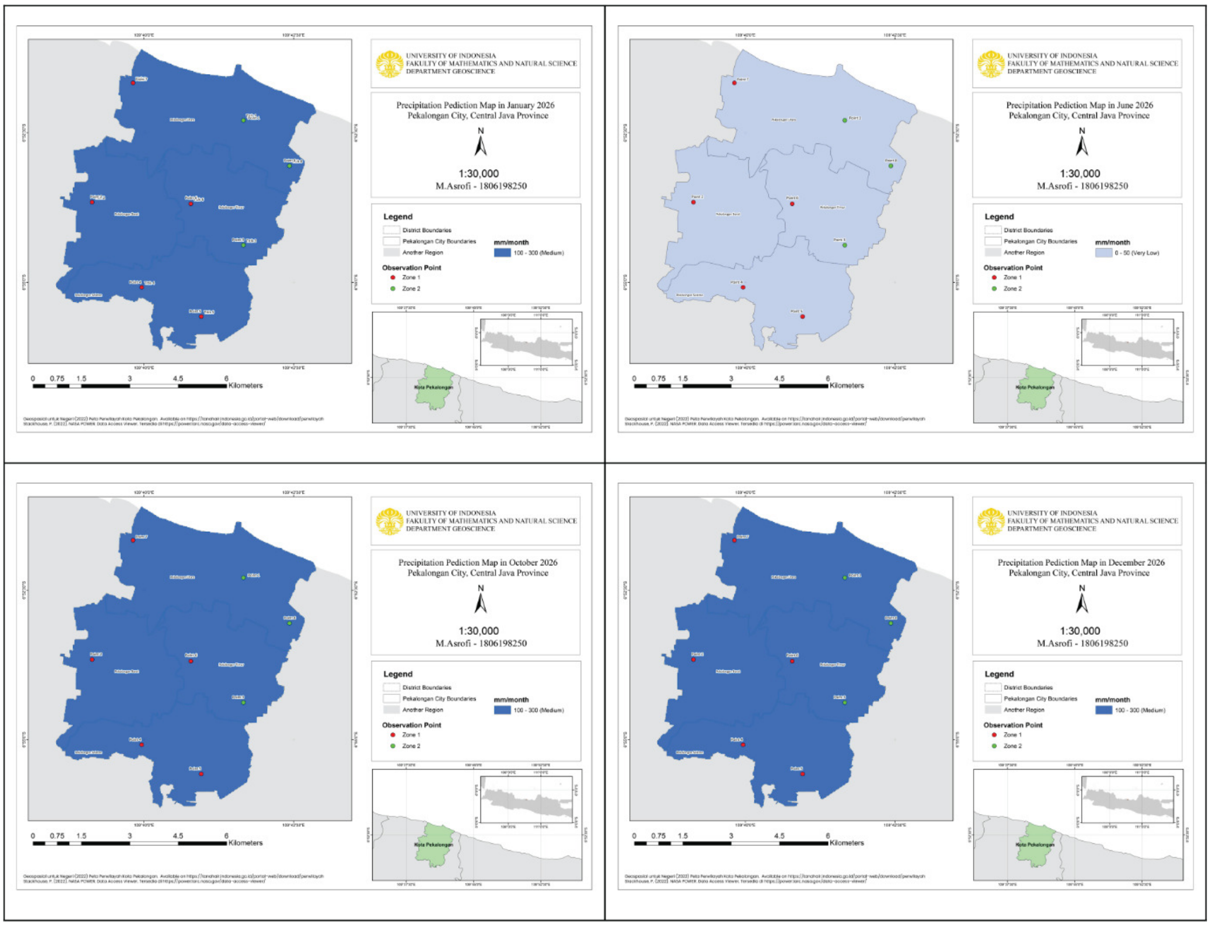

FIGURE 10. Continue...

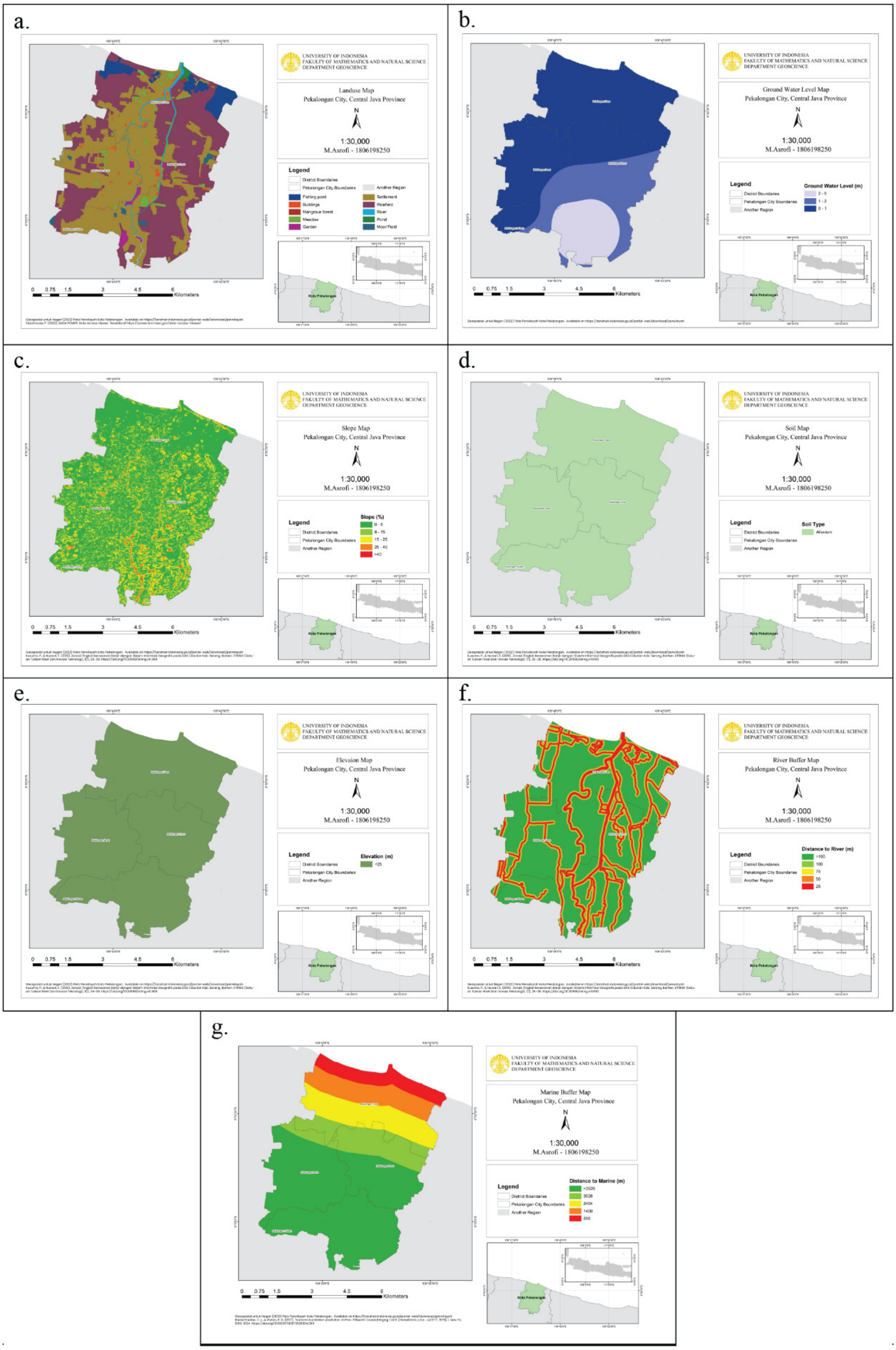

FIGURE 11. (a) Land use map, (b) Groundwater level map, (c) Slope map, (d) Soil map, (e) Elevation map, (f) River buffer map, (g) Marine buffer map.

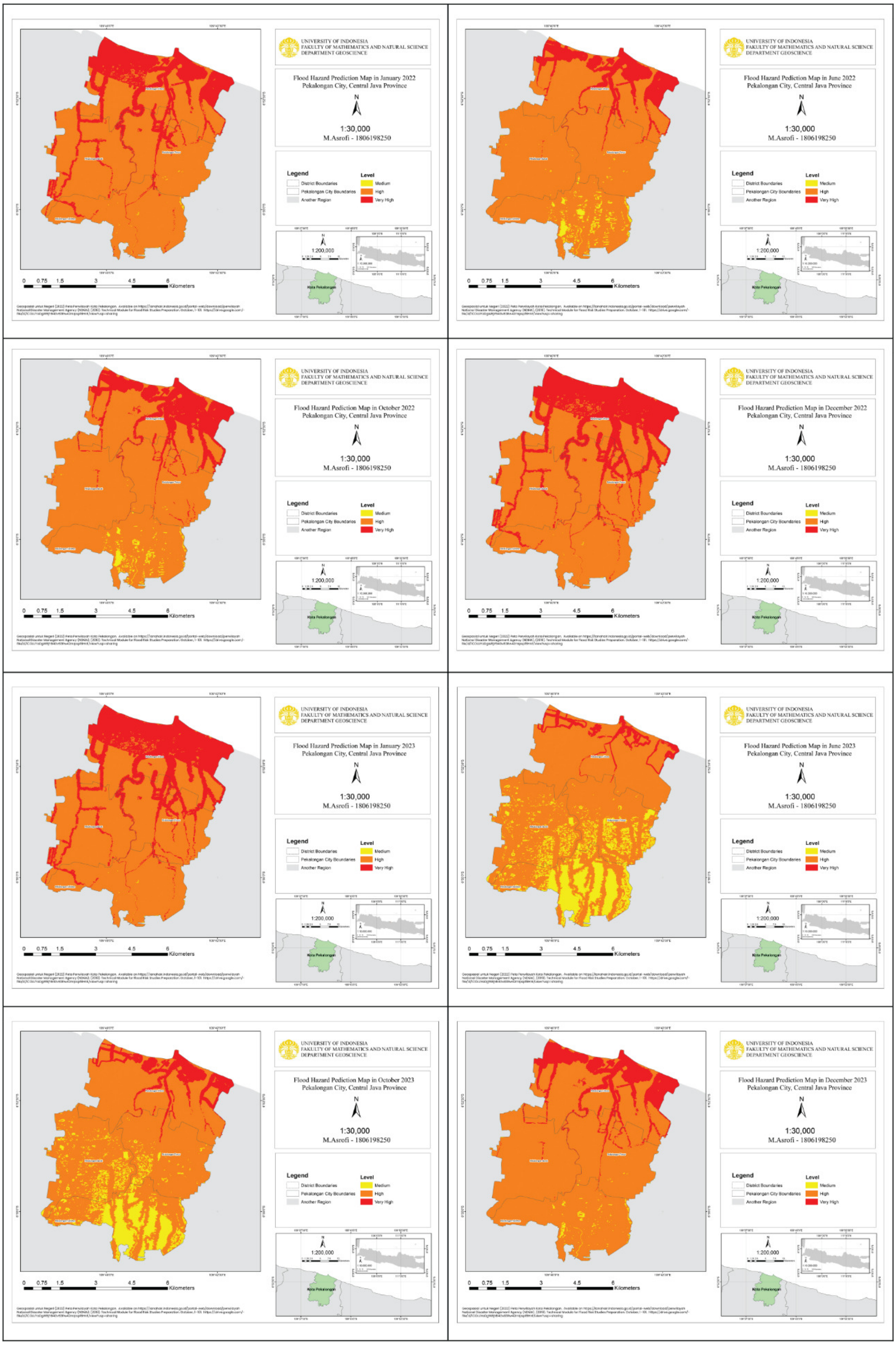

FIGURE 12. Flood hazard prediction map (2022–2026).

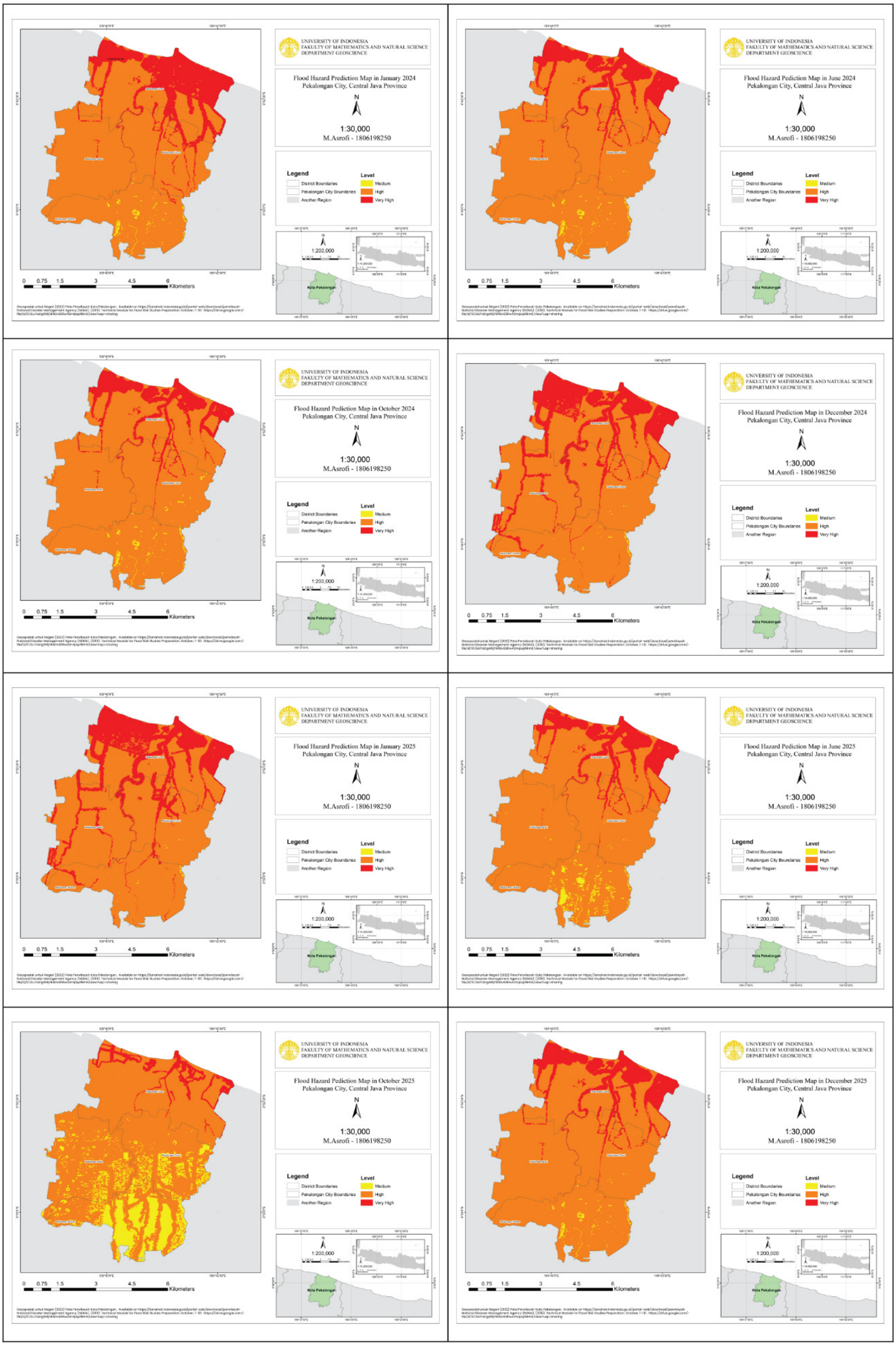

FIGURE 12. Continue...

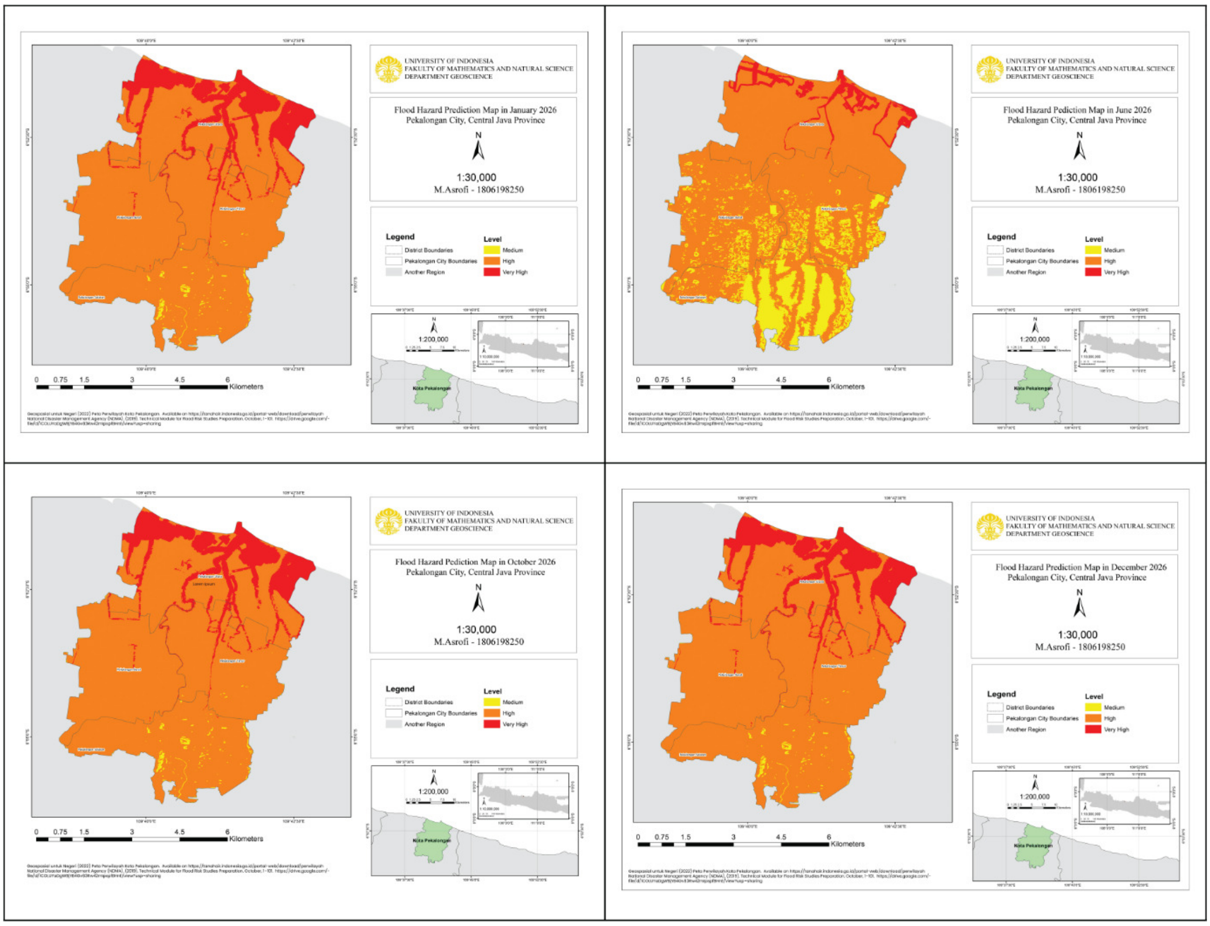

FIGURE 12. Continue...

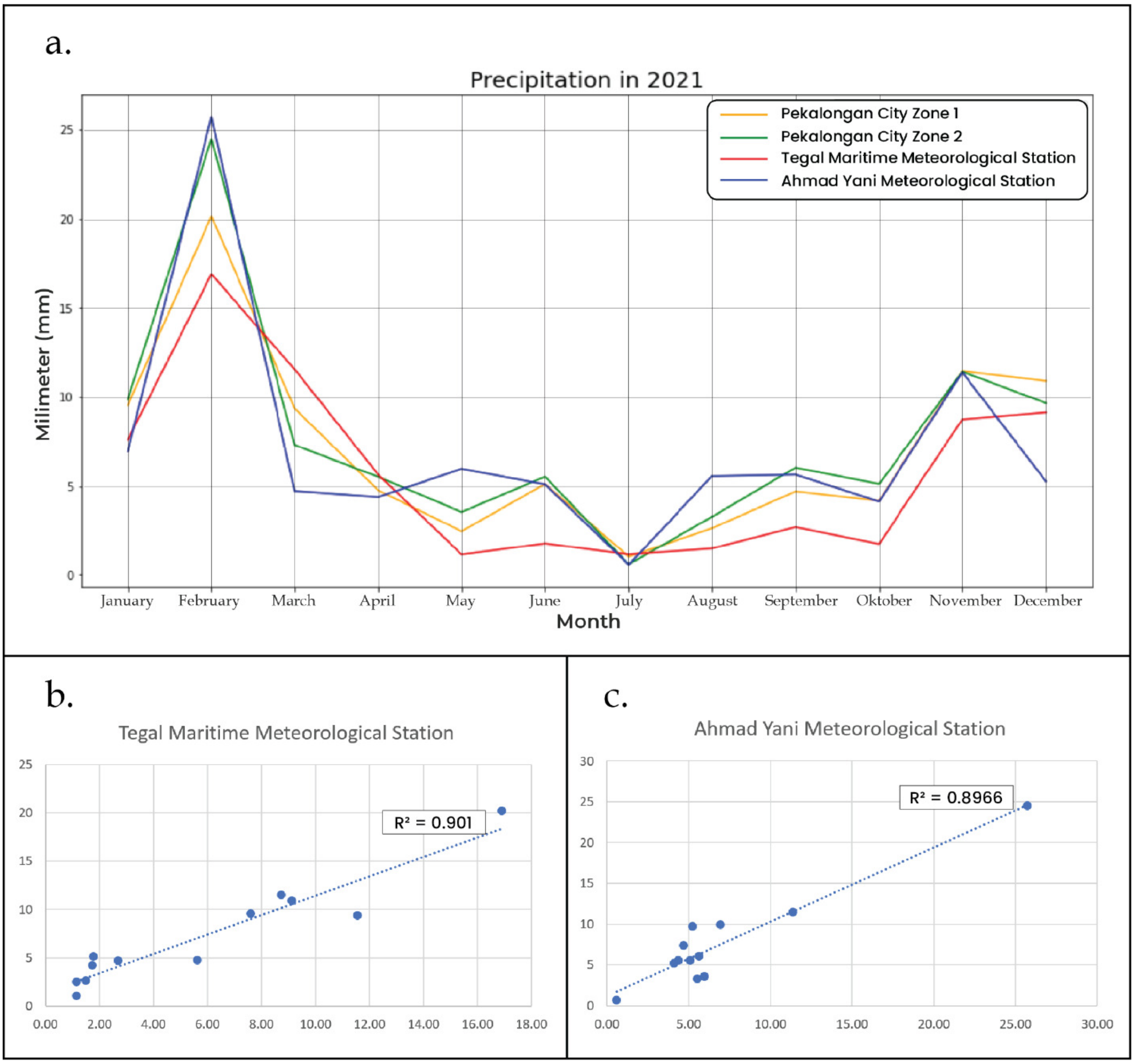

FIGURE 13. (a) Comparison of the actual rainfall value for Pekalongan (Zone 1 & Zone 2) with the rainfall value for Tegal (Tegal Maritime Meteorological Station) and Semarang (Ahmad Yani Meteorological Station) from January to December 2021; (b) The value of the coefficient of determination (r2) for validating the comparison of the rainfall value for Tegal (Tegal Maritime Meteorological Station) with the actual rainfall value for Pekalongan Zone 1; (c) The coefficient of determination (r2) value is used to validate comparing the rainfall value for Semarang (Ahmad Yani Meteorological Station) with the actual rainfall value for Pekalongan Zone 2.

diction results of Zone 1 using rainfall data at Tegal Maritime Meteorological Station and Zone 1 using rainfall data at Ahmad Yani Meteorological Station (Figure 14; Table 6). The prediction results for 2022 (January to October) in both zones have values and trend patterns that are close to the actual rainfall values from each of the closest observation stations. The predicted value for Zone 1 and the rainfall value from the Tegal Maritime Meteorological Station have a coefficient of determination (r2) of 0.7751 or 77.51%. The predicted value of Zone 2 and the rainfall value from the Ahmad Yani Meteorological Station have a coefficient of determination (r2) of 0.7575 or 75.75 % (Figure 15).

5 CONCLUSION

This research uses data on precipitation history in Pekalongan, which has a 40-year precipitation history (January 1st, 1981 – December 31st, 2021). There are eight observation points in the study area, divided into two zones with different precipitation values. Namely, observation points (1, 3, and 8) are called Zone 1, and observation points (2, 4, 5, 6, and 7) are called Zone 2. Long Short-Term Memory (LSTM) modeling uses precipitation data for the study area for 40 years (January 1981 – December 2021). This study uses the LSTM model analysis to predict precipitation/rainfall in the study area for the next five years (January 2022 – December 2026). The accuracy of the LSTM model shows the value of r2 in Zone 1, shows the independent variable (predicted value) 92 % can explain the dependent variable (original value), and the value of r2 in Zone 2, shows the independent variable (predicted value) 97.4 % can explain the dependent variable (original value). Validation carried out on rainfall data at the nearest rainfall observation station shows the predicted value of Zone 1, and the rainfall value from the Tegal Maritime Meteorological Station has a coefficient of determination (r2) of 0.7751 or 77.51 %.

The predicted value for Zone 2 and the rainfall value from the Ahmad Yani Meteorological Station have a coefficient of determination (r2) of 0.7575 or 75.75 %. The predicted results of the flood vulnerability level in Pekalongan show medium to very high levels, with a dominant high vulnerability level. Very high vulnerability levels are dominant in North Pekalongan District or coastal areas directly adjacent to the sea because flooding can also occur from the entry of seawater into the mainland. North Pekalongan District has the potential for tidal flooding because its area is directly adjacent to the sea. Ribs originating from the sea generally occur when sea waves are more than one meter. The advantage of using the LSTM model is that it is well used in projecting predictive values using datasets with a fairly long period. The drawbacks of the LSTM model are that changes in the actual values that are significant in the dataset make the error in the predicted value even greater.

Acknowledgements The Authors would like to thank Program Study Geology and the Department of Geography Universitas Indonesia for supporting this research.

REFERENCES

Ari Setiaji, Romy & Chandra, Andhika & Nahli, Kiddy & Radityo, Daniel & Arifullah, Ery. (2016). *The Stratigraphic Significance of Glossifungites Ichnofacies in Cipari Area, Central Java.*

Azis, M. F. A., Darari, F., & Septyandy, M. R. (2020). *Time series analysis on earthquakes using EDA and machine learning*. 2020 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2020, 405–412.

Bemmelen, V. R. W. (1949). *The Geology of Indonesia.* General Geology of Indonesia and Adjacent Archipelagoes. In Government Printing Office, The Hague (pp. 545–547; 561–562).

Badan Pusat Statistik (BPS) Pekalongan. (2021). *Pekalongan in Numbers. BPS Kota Pekalongan.*

Bretschneider, C. L., & Wybro, P. G. (1977). *Tsunami inundation prediction*. In: Proc. Fifteenth Coastal Engng. Conf. (Hawaii Univ., U.S.A.: Jul.11-17, 1976), 1, New Yo, 1006–1024.

Condon, W. H. (1975). *Geological Map of Banjarnegara and Pekalongan, Java*. Scale 1: 100000.

Hall, T., Brooks, H. E., & Doswell, C. A. (1999). Precipitation forecasting using a neural network. *Weather and Forecasting*, 14(3), 338–345.

Hardiyawan, M. (2012). *Kerentanan Wilayah Terhadap Banjir Rob Di Pesisir Pekalongan. Kerentanan Wilayah Terhadap Banjir Rob Di Pesisir Kota Pekalongan*, 20294641.

Hartono, G. (2010). *Produk Batuan Gunung Api Tersier Jawa Tengah the Role of Paleovolcanism in the Tertiary Volcanic Rock Product Setting At Mt . Gajahmungkur, Wonogiri, Central Jawa the Role of Paleovolcanism in the Tertiary Volcanic Rock Product Setting.*

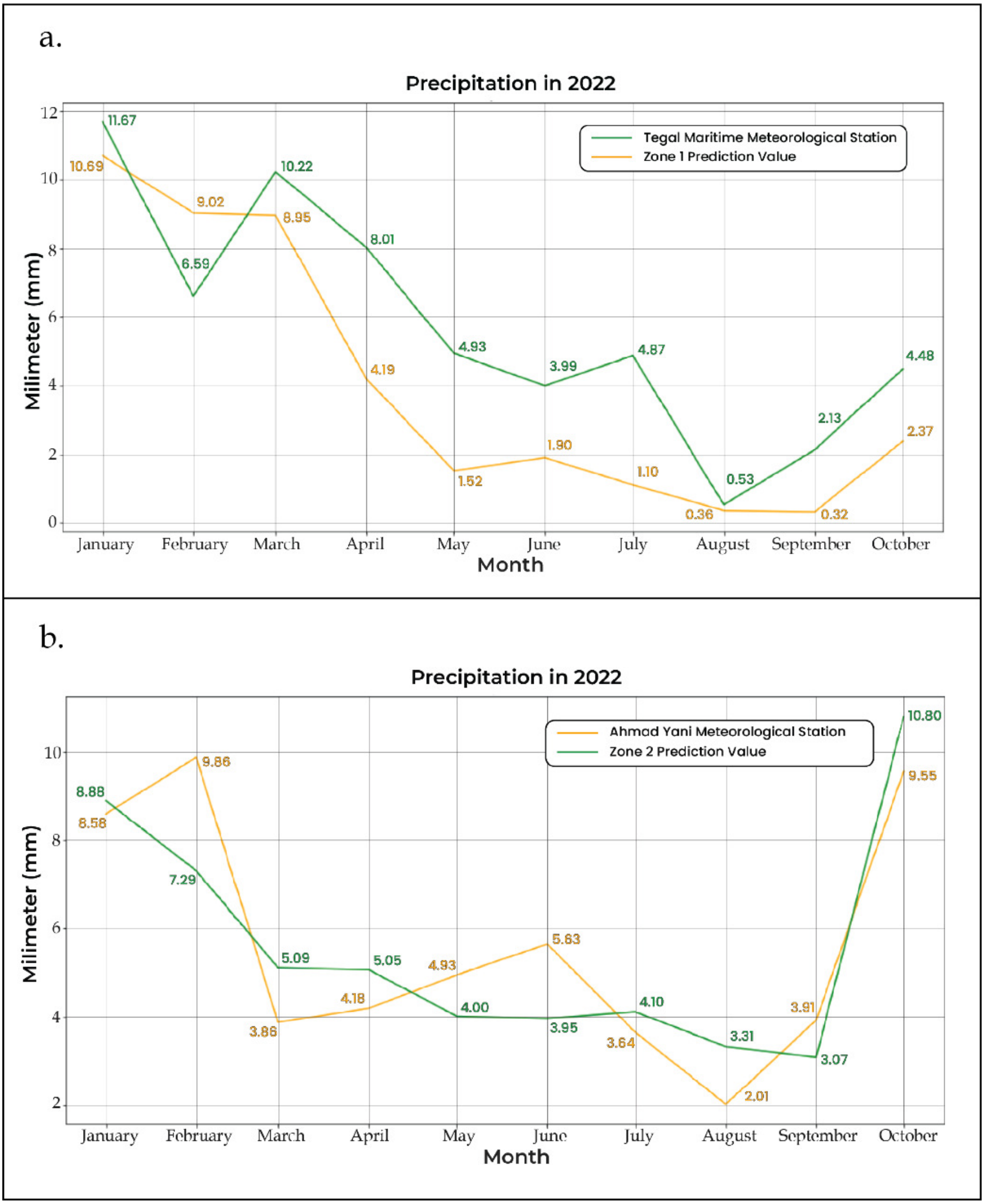

FIGURE 14. (a) Validation of the predicted value of Zone 1 with rainfall data from the Tegal Maritime Meteorological Station; (b) Validation of the predicted value of Zone 2 with rainfall data from the Ahmad Yani Meteorological Station, Semarang.

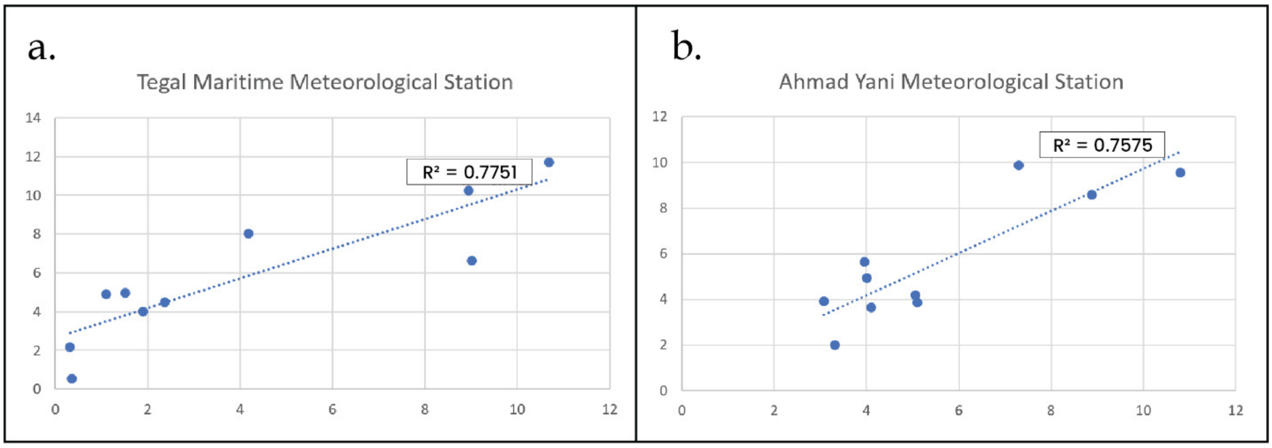

FIGURE 15. (a) The coefficient of determination (r2) for validating the predicted value of zone 1 with rainfall data from the Tegal Maritime Meteorological Station; (b) The coefficient of determination (r2) for validating the predicted value of zone 2 with rainfall data from the Ahmad Yani Meteorological Station, Semarang.

Hilmi, F., & Haryanto, I. (2008). Pola Struktur Regional Jawa Barat. *Bulletin of Scientific Contribution*, 6(1), 57–66.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory [J]. *Neural Computation*, 9(8), 1735–1780.

Kementrian Pekerjaan Umum dan Perumahan Rakyat. (2017). *Modul Geologi dan Hidrogeologi Pelatihan Perencanaan Air Tanah 2017*. Pusat Pendidikan Dan Pelatihan Sumber Daya Air Dan Konstruksi, 76.

Kuligowski, R. J., & Barros, A. P. (1998). Experiments in short-term precipitation forecasting using artificial neural networks. *Monthly Weather Review*, 126(2), 470–482.

Kusumo, P., & Nursari, E. (2016). Zonasi Tingkat Kerawanan Banjir dengan Sistem Informasi Geografis pada DAS Cidurian Kab. Serang, Banten. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 1(1), 29–38.

Meena, B. L., & Agrawal, J. D. (2015). Tidal level forecasting using ANN. *Procedia Engineering*, 116, 607–614.

Moustris, K. P., Larissi, I. K., Nastos, P. T., & Paliatsos, A. G. (2011). Precipitation forecast using artificial neural networks in specific regions of Greece. *Water Resources Management*, 25, 1979–1993.

National Disaster Management Agency (NDMA). (2019). Technical Module for Flood Risk Studies Preparation. October, 1–101.

Nugroho Adji, T., Nurjani, E., & Wicaksono, D. (2014). *Zonasi Potensi Airtanah Dengan Menggunakan Beberapa Parameter Lapangan dan Pendekatan SIG di Daerah Kepesisiran*. Laporan Akhir Penelitian, 1–30.

Rizki, M., Basuki, S., & Azhar, Y. (2020). Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory(LSTM) Untuk Prediksi Curah Hujan Kota Malang. *Jurnal Repositor*, 2(3), 331.

Rosyidie, A. (2013). Banjir: Fakta dan Dampaknya, Serta Pengaruh dari Perubahan Guna Lahan. *Journal of Regional and Planning*, 24(3), 241.

Salim, M. A. (2018). Penanganan banjir dan rob di wilayah Pekalongan. *Jurnal Teknik Sipil*, 11, 15–23.

Supriyadi, E. (2021). Prediksi Parameter Cuaca Menggunakan Deep Learning Long-Short Term Memory (Lstm). *Jurnal Meteorologi Dan Geofisika*, 21(2), 55.

Van Zuidam, R. A. (1985). *Aerial photo-interpretation in terrain analysis and geomorphologic mapping*. The Netherlands. Smith Publishers.

Wiranda, L., & Sadikin, M. (2019). Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 8(3), 184–196.