Sonar : Jurnal Multidisiplin

https://jurnal.untirta.ac.id/index.php/sonar/index

e-ISSN: XXXX-XXXX

Vol. 1, No. 2, April 2024, Page 81 - 89

Forecasting rainfall against flood potential using linear regression in the case study of Serang City

Gebyta Lee Yun Sindia[1]*, Yayat Ruhiat[1], Ganesha Antarnusa[1]

[1]Department of Physics Education, Universitas Sultan Ageng Tirtayasa, Indonesia

**E-mail: 2280200068@untirta.ac.id*

(Received: XX; Accepted: YY; Published: ZZ)

ABSTRACT

In 2022, Serang City experienced 166 floods, most of which occurred due to high rainfall. To anticipate and minimize the consequences that will occur by the flood disaster in 2023, rainfall forecasting is needed to forecast the potential flood areas in Serang City. The purpose of this research is to implement multiple linear regression analysis to forecast rainfall, forecast rainfall against flood potential, and identify the relationship between rainfall and flood events. The method of forecasting rainfall is done with multiple linear regression model analysis, Geographic Information System spatial analysis to forecast potential flood areas, and Pearson correlation test to identify the relationship between rainfall and flood events. The results show that rainfall in 2023 is predicted to be 66% of Serang City in the low category with rainfall values less than 1500 mm/year and 33% of other areas in the medium category with rainfall between 1500-2500 mm/year. In the same year, 9% of Serang City is predicted to have high flood potential, 25% medium, 1% low, and 64% is predicted to be safe from flooding. The relationship between rainfall and flood occurrence shows no correlation, so further research is needed.

Keywords: rainfall, Flood, forecasting, multiple linear regression, geographic information system.

DOI:

INTRODUCTION

Serang City is one of the eight districts/cities located in Banten Province and serves as the center of Banten Provincial Government. Geographically, Kota Serang is located between $5^{\circ}99' - 6^{\circ}22'$ South latitude dan $106^{\circ}07' - 106^{\circ}25'$ East longitude with an area of $266{,}18\ Km^2$ (Badan Pusat Statistik Kota Serang, 2022). Most of the geographical conditions of Serang City are in the lowlands with an altitude of less than 500 meters above sea level and a tropical climate. Areas that have a tropical climate sometimes have high rainfall (Hidayat et al., 2018). Under these circumstances, Serang City is likely to experience high rainfall which

can be one of the factors causing potential flooding.

Based on the flood disaster report of the Regional Disaster Management Agency (BPBD) of Serang City, in the last ten years 2013-2022 almost every year Serang City experienced a flood disaster. For example, in 2022 Serang City recorded 166 flood disasters, which were generally caused by very high rainfall. In 2019-2021, the Serang City Health Office reported 393 flood victims, consisting of 4 deaths, 1 missing victim, 88 minor injuries, and 300 displaced victims (Rahmadewi, 2023).

Referring to the high rainfall that causes many flood events in Serang City, rainfall forecasting of flood potential is needed to anticipate flood conditions and minimize the consequences that will occur. Forecasting is the act of methodically predicting what is most likely to happen in the future. (Yusuf et al., 2022).

Among the many ways to do rainfall forecasting, one of them is through the multiple linear regression method (Azhari, 2015). The multiple linear regression method is used to predict the value of the dependent variable based on the independent variables (Maulita & Nurdin, 2023). Multiple linear regression algorithms have the advantage of processing forecasts with effective and efficient results (Nurdin et al., 2022).

Several studies related to rainfall prediction have been carried out by previous studies using regression analysis, including by Mulyani et al. (2019) who predicted monthly rainfall in Majalengka Regency using multiple linear regression methods which resulted in a correlation coefficient value of r = 0.90 and an RMSE of 40.96%. In addition, Yusuf et al. (2022) who analyzed rainfall predictions in Sorong City with datasets obtained from BMKG DEO Sorong for the period 2017 to 2021. A correlation coefficient value of 0.8175 was obtained. Thus, the author will conduct rainfall forecasting of flood potential in Serang City using multiple linear regression.

RESEARCH METHODS

The data used is divided into three. First, the data for forecasting rainfall uses monthly data on rainfall, air temperature and wind speed for the 10-year period 2013-2022, in addition, monthly data on air temperature and wind speed for 2023 are also used as predictor variables for the rainfall forecast. Second, data to forecast areas with flood potential using data from the 2023 rainfall forecast and 2013-2022 flood events. Third, data to identify the relationship between rainfall and flood events using annual data of maximum rainfall and flood events 2013-2022. Rainfall data comes from BMKG Serang City and NASA (National Aeronautic and Space Administration) Power Data of the United States based on two rainfall measuring coordinates in Serang City, namely Serang Class I Maritime Meteorological Station and BPTPH Banten Province. Meanwhile, data on flood events came from BPBD Serang City and BPBD Banten Province.

The method of forecasting rainfall is done with multiple linear regression analysis using SPSS 25 software. The method of forecasting potential flood areas is done with Geographic Information System (GIS) spatial analysis with overlay techniques using ArcGIS 10.8 software. And the method to identify the relationship between rainfall and flood events is done with the Pearson correlation test using SPSS 25 software.

Multiple Linear Regression

Multiple linear regression is performed to produce rainfall forecasting using several independent variables as predictors. In this case the predictor variables (X) are air temperature and wind speed data. While the dependent variable or variable to be predicted (Y) is rainfall data. The equation form of Multiple Linear Regression is as follows (Harlan, 2018).

$$Y = a + b_1X_1 + b_2X_2 \quad (1)$$

Where Y is rainfall, a is a constant value, b_1 is regression coefficient for air temperature, b_2 is regression coefficient for wind speed, X_1 is air temperature, and X_2 is wind speed.

Spatial Analysis Reclassify and Overlay

The process of mapping rainfall forecast areas and flood potential in Serang City uses GIS assistance. Mapping of rainfall forecast distribution areas is done by spatial analysis using the Inverse Distance Weighting (IDW) interpolation method and reclassify. IDW is a traditional summation technique that considers distance as a weight (Albeta, 2021), the closer the distance between close input points will have a high weight compared to distant data points (Arianti et al., 2021). Reclassify is a classification technique used in mapping to convert a quantity with a certain interval into another new interval based on the specified category (Surjoatmodjo, 2019).

Meanwhile, the mapping of potential flood areas was carried out by spatial analysis using the flood parameter scoring method and parameter overlay analysis techniques. Overlay is the ability to combine maps that produce attribute information from the merged map (Surjoatmodjo, 2019).

Correlation analysis

Correlation analysis is used to determine as well as measure how strong the relationship between rainfall and flood events that occur in Serang City. The type of correlation test used is the Pearson correlation test (*r*) which is carried out with the help of SPSS. The decision making is as follows.

H_0 : There is no relationship between rainfall and flood events

H_1 : There is a relationship between rainfall and flood occurrence

According to Sugiyono (2017) to determine how strong the relationship between variables is to correlation, it can be guided by the provisions shown in Table 1.

Table 1. Guidelines for the Relationship Between Variables Correlation Coefficient

Interval Coefficient	Level of Relationship
0,00 – 0,199	Very low
0,20 – 0,399	Low
0,40 – 0,599	Medium
0,60 – 0,799	Strong
0,80 – 1,00	Very strong

RESULTS AND DISCUSSION

Rainfall Forecasting in Serang City

Using monthly data on rainfall, air temperature and wind speed for the period 2013-2022, multiple linear regression models were obtained at each rainfall measuring coordinate point in Serang City as shown in equations (2) and (3).

Maritime Meteorological Station Class I Serang:

$$Y = 1766{,}096 - 63{,}403X_1 + 61{,}160X_2 \quad (2)$$

BPTPH Banten Province:

$$Y = 976{,}007 - 38{,}351X_1 + 126{,}426X_2 \quad (3)$$

The following values of 2023 rainfall forecasting results obtained from multiple linear regression models at each coordinate point are shown in Table 2 and the distribution of 2023 rainfall forecasts in the Serang City area is shown in Figure 1. And the error deviation as well as the correlation between the forecast value and the actual value at each coordinate point are shown in Table 3.

Table 2. 2023 Rainfall Forecasting Results at each Coordinate Point

Month	Maritime Meteorological Station Class I Serang (mm)	BPTPH Banten Province (mm)
January	108,4	335,1
February	139,6	313,7
March	121,0	193,4
April	101,3	237,4
May	20,3	113,3
June	58,3	128,6
July	96,4	160,4
August	70,8	164,3
September	58,1	148,9
October	-17,8	69,9
November	52,2	73,8
December	7,8	152,3

Table 3. RMSE and Correlation Values

	Maritime Meteorological Station Class I Serang	BPTPH Banten Province
RMSE	85,2	114,4
Correlation	0,719	0,582

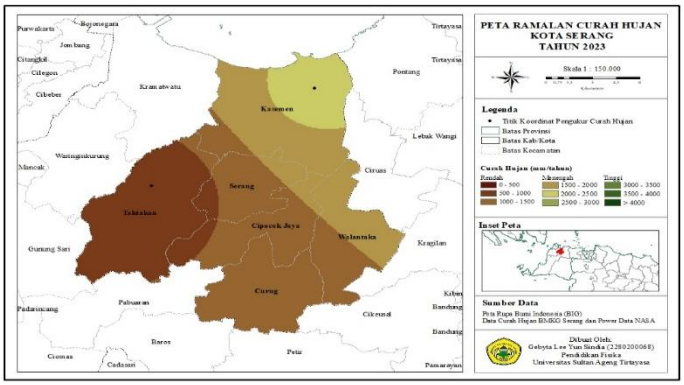

Figure 1. Map of Rainfall Forecast of Serang City in 2023

The RMSE values in Table 3 show that the multiple linear regression method for each coordinate point has an accuracy error of 85.2 and 114.4. With these RMSE values, rainfall forecasting with this method can be said to be poor because the error value between the forecast results and the actual data produced is so large. However, the correlation value shows a relationship between rainfall forecasts and actual rainfall data and is positive. Where according to Sugiyono (2017) a correlation value of 0.719 indicates a strong relationship and a correlation of 0.582 indicates a moderate relationship. And the correlation relationship is positive, stating that if the actual rainfall value increases, the rainfall forecast results also increase as well as if the actual rainfall decreases.

Badan Meteorologi Klimatologi dan Geofisika Stasiun Klimatologi Banten (2022), classifies rainfall into four categories of rainfall classes consisting of low rainfall (0 - 100 mm), medium rainfall (100 - 300 mm), high rainfall (300 - 500 mm), and very high rainfall (> 500 mm).

The highest rainfall at the coordinates of the Serang Class I Maritime Meteorological Station is predicted to occur in February at 139.6 mm. This shows conformity with the actual data. However, the forecast results show that rainfall in February is included in the medium rainfall category, while the actual is included in the high category of 309.4 mm. While the lowest rainfall is predicted to occur in October at -17.8 mm with (-) indicating that in that month no rain occurred. This also shows conformity with the actual data that in October there was no rain. But based on the actual no rain also occurred in August. Overall, the results of the monthly rainfall forecast from January 2023 to December 2023 at the coordinates of the Serang Class I Maritime Meteorological Station using multiple linear regression show lower results than the actual data except in July to October.

The highest rainfall at the coordinates of BPTPH Banten Province was predicted to occur in January at 335.1 mm with a high rainfall category. While the actual, occurred in February amounting to 555.1 mm with high rainfall category. The lowest rainfall was predicted to occur in October at 69.9 mm. While the actual, occurred in September amounting to 3.6 mm. The closest match or rainfall forecast results to the actual data occurred in June. Overall, the results of monthly rainfall forecasts from January 2023 to December 2023 at the coordinates of BPTPH Banten Province using multiple linear regression show higher results than the actual data except in February, March, May, June, and November.

The Rainfall Forecast distribution map in Figure 1 shows that most areas of Serang City

in 2023 experience a low rainfall category of 66%, which occurs in the Taktakan and Curug sub-districts, part of the Walantaka sub-district, most of the Serang sub-district, Cipocok Jaya, and a small part of the Kasemen sub-district with rainfall below 1500 mm / year. And the other 34% of the area experienced medium rainfall category, which occurred in most of the kasemen sub-district, part of the walantaka sub-district, and a small part of the serang and cipocok jaya sub-district with rainfall of 1500-2500 mm/year.

Forecasting Flood Potential Areas in Serang City

By scoring and overlaying the parameters, namely the rainfall forecast distribution map and the history of flood events shown in Figure 2, the forecast distribution of potential flood areas is obtained as shown in Figure 3.

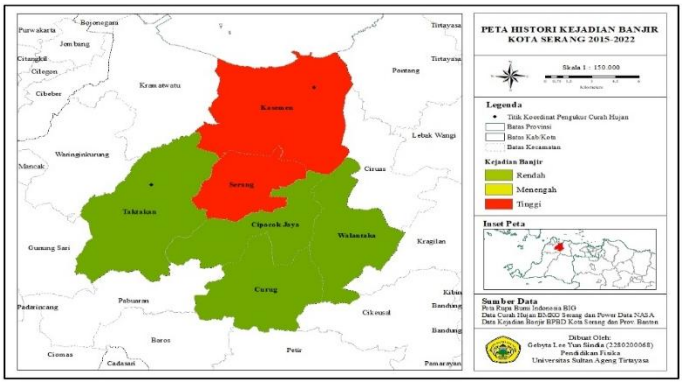

Figure 2. History of Flood Events in Serang City 2013-2022

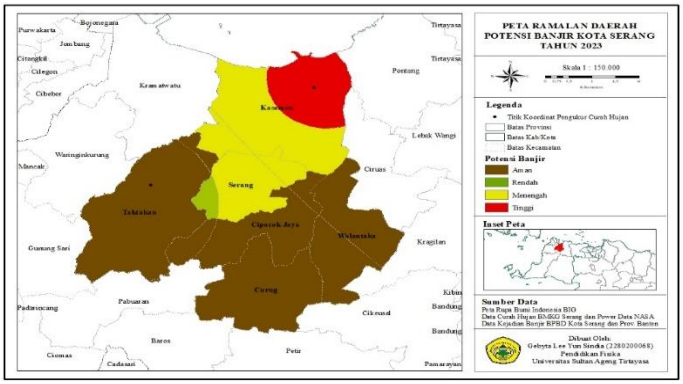

Figure 3. Forecast Map of Potential Flood Areas in Serang City in 2023

The following is the scoring of each map parameter.

Table 4. Rainfall Scoring

Number	Category	Amount of Rainfall (mm/year)	Score
1	Low	< 1000	1
2	Medium	1000 – 2000	2
3	High	2000 – 3000	3
4	Very High	> 3000	4

(Ramadhani et al., 2022 with author's modification)

Table 5. Scoring of Historical Flood Events

Number	Category	Number of Events	Score
1	Low	< 46	1
2	Medium	46 – 87	2
3	High	> 87	3

(Ramadhani et al., 2022 with author's modification)

After scoring each parameter, the parameters were combined using an overlay technique. Then the rainfall score was added with the historical flood event score to obtain the total score from the scoring (see Table 6). Furthermore, the total score was analyzed using the following equation (Bandawa Putri et al., 2023) to determine the forecast of potential flood areas based on the size of the category interval.

$$Ki = \frac{Xt - Xr}{K} \quad (4)$$

Where Ki is category interval, Xt is highest data, Xr is lowest data, and K is number of categories.

Table 6. Classification of Potential Flood Areas

Number	Category of Flood Potential	Total Score
1	Safe	3
2	Low	4
3	Medium	5
4	High	6

The forecast of flood potential is shown in Figure 3 which states that the Serang City area is divided into four categories of flood potential, namely safe potential, low potential, medium potential, and high potential. Areas with safe flood potential include Taktakan, Cipocok Jaya, Curug, and Walantaka sub-districts with an area of 17,078.5 hectares or 64% of Serang City. Areas with low flood potential include a small part of Serang sub-district with an area of 357.1 Ha or 1%. Areas with medium flood potential cover most of the Kasemen and Serang sub-districts with an area of 6,595.1 Ha or 25%. And areas with high flood potential include a small part of Kasemen sub-district with an area of 2,459.7 Ha or 9%.

Relationship between Rainfall and Flood Occurrence in Serang City

By using the annual data of maximum rainfall and flood events from 2013 to 2022, the Pearson correlation test obtained the results as shown in Table 7.

Table 7. Pearson Correlation Test Results

Correlations			
		Rainfall	Flood_Occurrence
Rainfall	Pearson Correlation	1	-,441
	Sig. (2-tailed)		,203
	N	10	10
Flood_Occurrence	Pearson Correlation	-,441	1
	Sig. (2-tailed)	,203	
	N	10	10

The correlation test results in Table 7 show that H0 is accepted and H1 is rejected, meaning that there is no relationship between rainfall and flood events in Serang City. This can be seen from the significance value (sig.), it is said that there is a relationship if the significance value is smaller than 0.05, otherwise there is no relationship if the significance value is greater than 0.05. The absence of a relationship between rainfall and flood events can be caused by many factors other than rainfall. Based on the Serang City flood disaster report, several factors cause flooding, such as overflowing water and drainage that is not functioning properly. The results of this correlation test are also in accordance with the research Effendy & Daryono (2018) that partially the rainfall variable has no relationship with flooding.

CONCLUSION

The multiple linear regression model obtained to forecast rainfall in Serang City in 2023 is $Y = 1766{,}096 - 63{,}403X_1 + 61{,}160X_2$ for the coordinates of Serang City Meteorological Station and $Y = 976{,}007 - 38{,}351X_1 + 126{,}426X_2$ for BPTPH Banten Province. Rainfall forecasting in Serang City in 2023 shows that 66% of the Serang City area mostly experiences low category rainfall with values below 1500 mm/year and 34% of the other area experiences medium category rainfall with values of 1500-2500 mm/year. In that year, it is predicted that 9% of Serang City has a high flood potential, 25% medium, 1% low, and 64% safe or no flood potential. The relationship between rainfall forecasts and flood events in Serang City through correlation analysis, shows that rainfall and flood events in Serang City have no correlation. This can be caused by many factors other than rainfall that cause flooding in Serang City, so further research is needed.

REFERENCES

Albeta, R. (2021). *Studi Sebaran Curah Hujan Menggunakan Metode Interpolasi Invers Distance Weighted (IDW) di Sumatera Selatan*. Universitas Sriwijaya.

Arianti, I., Soemarno, AW, H., & R, S. (2021). Rainfall Estimation By Using Thiessen Polygons, Inverse Distance Weighted, Spline, And Kriging Methods : A Case Study in Pontianak, West Kalimantan. *International Journal of Education and Research*, *6*(11).

Azhari, A. (2015). Simulasi Prediksi Curah Hujan Bulanan Di Kota Medan Menggunakan Metode Regresi Linier Berganda. *EINSTEIN E-JOURNAL*, *3*(2). https://doi.org/10.24114/einstein.v3i2.12346

Badan Meteorologi Klimatologi dan Geofisika Stasiun Klimatologi Banten. (2022). *Buletin Klimatologi Provinsi Banten dan DKI Jakarta Tahun 2022*. Stasiun Klimatologi Banten.

Badan Pusat Statistik Kota Serang. (2022). *Kota Serang dalam Angka Serang Municipality in Figures 2022*. BPS Kota Serang.

Bandawa Putri, S., Meirany, J., Novita Pratiwi, N., Deah Lestari, A., & Meddy Danial, M. (2023). Pemetaan Daerah Rawan Bencana Dalam Upaya Mitigasi Bencana Berbasis Gis Di Kecamatan Sungai Raya Kepulauan Kabupaten Bengkayang Provinsi Kalimantan Barat. *Zona Laut Jurnal Inovasi Sains Dan Teknologi Kelautan*, *4*(2), 171–181.

https://doi.org/10.62012/zl.v4i2.27374

Effendy, Z., & Daryono, D. (2018). Hubungan Curah Hujan dan Penggunaan Lahan Terhadap Genangan Banjir di Kecamatan Sidoarjo Kabupaten Sidoarjo. *Jurnal Mahasiswa Universitas Negeri Surabaya*, *2*(1), 1–10.

Harlan, J. (2018). *Analisis Regresi Linier*. Jakarta: Gunadarma.

Hidayat, N. M., Pandiangan, A. E., & Pratiwi, A. (2018). Identifikasi perubahan curah hujan dan suhu udara menggunakan Rclimdex Di wilayah Serang. *Jurnal Meteorologi Klimatologi Dan Geofisika*, *5*(2), 37–44.

Maulita, M., & Nurdin, N. (2023). Pendekatan Data Mining Untuk Analisa Curah Hujan Menggunakan Metode Regresi Linear Berganda (Studi Kasus: Kabupaten Aceh Utara). *IDEALIS : InDonEsiA JournaL Information System*, *6*(2), 99–106. https://doi.org/10.36080/idealis.v6i2.3034

Mulyani, E. D. S., Septianingrum, I., Nurjanah, N., Rahmawati, R., Nurhasani, S., & K, K. M. R. (2019). Prediksi Curah Hujan Di Kabupaten Majalengka Dengan Menggunakan Algoritma Regresi. *E-JURNAL JUSITI : Jurnal Sistem Informasi Dan Teknologi Informasi*, *8*(1), 67–77. https://doi.org/10.36774/JUSITI.V8I1.602

Nurdin, N., Fajriana, F., Maryana, M., & Zanati, A. (2022). Information System for Predicting Fisheries Outcomes Using Regression Algorithm Multiple Linear. *Journal Of Informatics And Telecommunication Engineering*, *5*(2), 247–258. https://doi.org/10.31289/jite.v5i2.6023

Rahmadewi, R. R. F. C. A. (2023). *Pengaruh Edukasi Penanggulangan Bencana Banjir Terhadap Tingkat Pengetahuan Kesiapsiagaan Banjir Pada Masyarakat Kota Serang*. Universitas Sultan Ageng Tirtayasa.

Ramadhani, D., Hariyanto, T., & Nurwatik, N. (2022). Penerapan Metode Analytical Hierarchy Process (AHP) dalam Pemetaan Potensi Banjir Berbasis Sistem Informasi Geografis (Studi Kasus: Kota Malang, Jawa Timur). *Geoid*, *17*(1), 72–80. https://doi.org/10.12962/j24423998.v17i1.10250

Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

Surjoatmodjo, M. I. (2019). *Analisis Potensi Pergerakan Arah Aliran Banjir Lumpur untuk Menunjang Kegiatan Mitigasi Bencana (Studi Kasus: Lumpur Sidoarjo Kecamatan Porong dan Tanggulangin)*. Institut Teknologi Sepuluh Nopember.

Yusuf, M., Setyanto, A., & Aryasa, K. (2022). Analisis Prediksi Curah Hujan Bulanan Wilayah Kota Sorong Menggunakan Metode Multiple Regression. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, *6*(1), 405–417. https://doi.org/10.30645/J-SAKTI.V6I1.455