E-ISSN : 2540 - 8984

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Volume 07, Nomor 04, Desember 2022 : 1244 - 1253

PREDICTION AND MAPPING RAINFALL CLASSIFICATION USING NAÏVE BAYES AND SIMPLE KRIGING

Indra Kusuma Yoga[1], Sri Suryani Prasetyowati[2], and Yuliant Sibaroni[3]
[1, 2, 3]Informatics, Scholl of Computing, Telkom University
Jl. Telekomunikasi No. 1, Bandung
e-mail: indrakusumayoga@student.telkomunversity.ac.id[1], srisuryani@telkomuniversity.ac.id[2], yuliant@telkomuniversity.ac.id[3]

ABSTRAK

Penelitian ini menggunakan naïve bayes untuk klasifikasi dan simple kriging untuk interpolasi peta curah hujan pulau Jawa. Hujan merupakan fenomena alam yang tidak bisa dikendalikan dan sulit untuk diprediksi secara langsung. Curah hujan yang tinggi yang berlangsung lama dapat mengakibatkan banjir dan dampak negatif lainnya. Dengan memanfaatkan machine learning untuk prediksi curah hujan, memudahkan masyarakat untuk lebih berantisipasi. Hasil dari penelitian ini adalah prediksi dan pemetaan curah hujan pulau Jawa tanggal 1 April 2022 sampai 7 April 2022 dan bulan April 2022 sampai September 2022. Data yang digunakan adalah data curah hujan di setiap stasiun cuaca di pulau Jawa periode 2010-2021. Digunakan dua pemodelan yaitu model harian dan bulanan. Model harian didapatkan nilai akurasi sebesar 67% sedangkan pada model bulanan didapatkan nilai akurasi sebesar 88%. Nilai akurasi harian lebih kecil dikarenakan data yang digunakan lebih banyak dibanding data bulanan, sedangkan metode naive bayes kurang cocok dengan data banyak. Namun, evaluasi tidak hanya berfokus pada akurasi, tetapi juga hasil perhitungan precision, recall, dan f1-score. Semivariogram yang digunakan pada model harian adalah spherical dengan rata-rata RMSE 1,021. Untuk model bulanan menggunakan semivariogram gaussian dengan rata-rata RMSE 0,34. Hasil pemetaan penelitian ini dapat digunakan masyarakat untuk mengantisipasi dampak dari curah hujan yang akan terjadi.

Kata Kunci*: Naïve bayes, simple kriging, interpolasi, klasifikasi.*

ABSTRACT

This study discusses the development of a prediction model for the classification of rainfall based on time in Java. The method used in this research is naive bayes and simple kriging. Naive Bayes is used for classification prediction, while simple kriging is an interpolation method used for mapping. There are two scenarios used, that is building a prediction model for daily and monthly rainfall classification, with data taken from 27 weather stations on the island of Java from 2010 to 2021. The results obtained in the classification process are an accuracy value of 67% for the daily model and 88% for the monthly model. The daily model data uses a spherical semivariogram with an average RMSE of 1,021. For the monthly model data using a Gaussian semivariogram with an average RMSE of 0,34. Then interpolation using simple kriging for mapping rainfall. The results of this study are predictions for the classification and mapping of daily rainfall models from April 1 to April 7 2022 and monthly models from April to September 2022. The contribution of this research is to provide predictive information and mapping of future rainfall so that public people can anticipate more.

Keywords*: Naïve bayes, simple kriging, interpolation, classification.*

I. INTRODUCTION

WEATHER and climate are natural phenomena whose existence is important in determining human activities. For example, weather and climate information is used as a reference in agriculture for determining the planting period and transportation for aircraft departure schedules. Weather is a natural phenomenon that tends to be uncontrollable[1]. One form of weather and climate is rainfall. High rainfall usually occurs in December-March every year. One of the impacts of high rainfall is flooding. One of the areas experiencing flooding is Jakarta. In 2007 Jakarta was hit by a big flood[2]. It is reported that almost 70% of the Jakarta area is inundated by water.

Weather and climate information increases as unusual natural phenomena increase [3]. The impact of high rainfall can be minimized by providing predictive information on the occurrence of unusual rainfall in the future. One of the Government Institutions that provide rainfall prediction data is BMKG. According to BMKG, predictions for 2022-2023 will be more accurate due to changes in the seasonal zone data used. Previously used the 1981-2010 season zone, now updated to 1991-2020. However, it is possible that the predictions made are incorrect.

The size of the intensity of rainfall cannot be regulated by human intervention but can be predicted. Accurate rainfall prediction is needed to anticipate the negative impact of high rainfall. Thus, it is possible to conduct research

E-ISSN : 2540 - 8984

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Volume 07, Nomor 04, Desember 2022 : 1244 - 1253

on rainfall prediction. Research[4] are air temperature, wind, exposure time, and relative humidity. used the attributes of air temperature, wind, exposure time, and relative humidity. Prediction is done by classification using machine learning on the data provided. Based on this research, there are still shortcomings, namely the attributes used are still few and the results obtained do not display a mapping only predictive values.

One of the well-known machine learning methods for the classification process is Naive Bayes. Naïve Bayes was discovered in the 18th century by Thomas Bayes using Bayes' formula[5]. Research conducted by Slamet Triyanto et al in 2021 predicts flood disaster classification based on rainfall using naive Bayes and obtains an accuracy of 79.16%[6]. In 2018 research was also conducted by Nabila and Siti to predict rainfall in the city of Subang[7]. The study compares several classification methods with evaluation using RMSE and for the nave Bayes model, the RMSE value is 0.37 which is a fairly good value. However, in this study, only 5 parameters were used. In 2018, A. K. Sharma et al compared 4 methods for the prediction of rainfall in India[8]. One of the methods used is nave Bayes which obtains accuracy above 96% for each different area. Research[9] conducted by A. U. Uzmi et al in 2021 carried out a classification of rainfall using Naive Bayes and obtained an accuracy of 96%. Based on several previous studies, the nave Bayes method is a fairly good method for classifying rainfall.

The availability of rainfall prediction maps will greatly help the community, especially farmers, to more easily understand the results of rainfall predictions. However, in several previous studies, there has been no research that produces output in the form of predictive maps. S. S. Prasetyowati et al conducted research on air pollution in 2020 in the Bandung area and produced output in the form of a prediction map of air pollution in the next few years using the Simple Kriging method[10]. In the same year, the research of M. Hassim et al. compared several interpolation methods to map rainfall in the Langat river area, Malaysia and the final result obtained is the simple kriging method is the most optimal method because it has the smallest RMSE value[11]. In 2021, [12] research will interpolate rainfall in Pakistan. One of the interpolation techniques used is Simple Kriging. The average RMSE value is 0.365 for each semivariogram model used and includes the smallest RMSE compared to other interpolation techniques. In the same year, an interpolation study using simple kriging was also carried out and the RMSE value was 0.93[13].

In some of the rainfall prediction studies above, the attributes used are still few and there is no output in the form of mapping. In research on interpolation, the resulting output still does not produce a mapping according to the existing map form. So, from these several factors, it is an opportunity to conduct research on the prediction of rainfall classification using Naive Bayes and mapping with simple kriging interpolation. The addition of attributes and outputs that adjust the shape of the map will be added to this research. The results of this study are predictions for the next few days and months according to the data used. It is hoped that this research can be useful for the general public to better anticipate the impact of rainfall.

II. Research Method

The first time this research was carried out was to take datasets from the BMKG website in the period 2010 to 2021 for weather stations located on the island of Java. Next is data processing so that the data can be processed. Then perform classification predictions using the Naïve Bayes method. After knowing the prediction results using nave Bayes, kriging interpolation is carried out to produce a prediction map for rainfall classification for the next few years using Simple Kriging. For a clearer flow can be seen in Figure 1.

E-ISSN : 2540 - 8984

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Volume 07, Nomor 04, Desember 2022 : 1244 - 1253

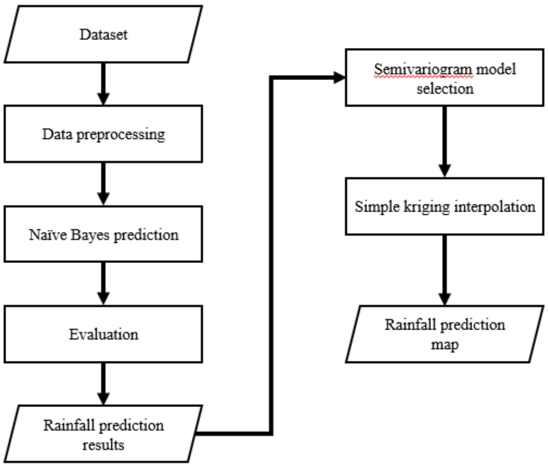

Figure 1 Classification and mapping process

A. Dataset

The dataset used in this study was taken from several BMKG observation stations on the island of Java. There are 9 parameters used in this study, namely, wind direction at maximum speed, rainfall, maximum final speed, average wind speed, average humidity, irradiation time, maximum temperature, minimum temperature, and average temperature.

B. Preprocessing

Preprocessing is the process of converting raw data into data that is ready to be processed for modeling. There are several stages in data preprocessing, namely overcoming null or empty values, labeling rainfall classes, normalizing, and dividing training data and test data.

C. Naïve Bayes

Nave Bayes is one of the most popular classification methods in the machine learning field. Nave Bayes models are useful for large datasets because they are easy to build and easy to understand [7]. The nave Bayes method yields the validity of probability theory, but cannot perform the inferentiality of many closely related rules. The advantage of the nave Bayes method is that it produces a fairly high accuracy value[14]. The nave Bayes equation can be seen in (1).

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{1}$$

Where H is hypothesis and X is sample data. While P(H) is observed probability and P(X) is previous probability.

From the beginning, the nave Bayes method classifies the probability values of a dataset. This can be done if the dataset used is discrete data. However, this study uses continuous data which if using the usual nave Bayes model will not work. Several methods have been introduced to overcome the problems in ordinary nave bayes.

Gaussian Naïve Bayes classification is a nave Bayes method to overcome research in the form of continuous data. Gaussian nave Bayes uses the assumption that it has Gaussian distribution on values labeled as a class or can be called continuous data[15]. The formula used is as follows:

$$P(hj) = \frac{1}{\sqrt{2\pi\sigma^2}} exp \frac{(hj-\mu)^2}{2\sigma^2} \tag{2}$$

Where hj is likelihood, σ is standard deviation, and μ is mean for the data.

D. Evaluation

After the classification, it is necessary to know the performance of the built model. To determine the value of performance evaluation is carried out. The evaluation that will be used is the Confusion Matrix. By knowing the

E-ISSN : 2540 - 8984

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Volume 07, Nomor 04, Desember 2022 : 1244 - 1253

total prediction is correct, it is possible to see the accuracy value[14]. The confusion matrix can be seen in Table I.

TABLE I
CONFUSION MATRIX

Correct classification \ Classification	+	-
+	True Positives(TP)	False Negative(FN)
-	False Positives(FP)	True Negatives(TN)

Based on table I , TP is positive and correct predictive data, FP is negative and true predictive data, FN is positive and false predictive data, and TN is negative and incorrect predictive data. Parameters evaluated are accuracy, precision, recall, and F1-score.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\% \tag{3}$$

$$Precision = \frac{TP}{(TP+FP)} \times 100\% \tag{4}$$

$$Recall = \frac{TP}{(TP+FN)} \times 100\% \tag{5}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \tag{6}$$

E. Semivariogram Model

The process of preparing the prediction map starts with determining the best semivariogram model, which is then followed by simple kriging interpolation. The determination of the semivariogram model is done by matching the pattern between the experimental and theoretical semivariogram, based on the smallest RMSE value[10].

Below are some of the most popular semivariogram models [16] which may be used in this research:

1. Spherical

$$y(h) = \begin{cases} C_0 + C & h > r \\ C_0\left(\frac{3h}{2r} - \frac{1}{2}\left(\frac{h}{r}\right)^3\right) & 0 < h \le r \end{cases} \tag{7}$$

2. Exponential

$$y(h) = \begin{cases} 0 & h = 0 \\ C_0\{1 - exp(-|h|/r)\} & 0 < h \end{cases} \tag{8}$$

3. Gaussian

$$y(h) = C_0\left(1 - exp\left(\frac{-|h|}{r}\right)^2\right) \tag{9}$$

Where C is sill, r is range, and h is lag

F. Simple Kriging

After determining the best semivariogram model, then kriging interpolation is carried out to produce a predictive map. One of the kriging methods is Simple Kriging. Simple kriging uses a known and constant population means[17]. Simple kriging is the method with the simplest calculations, but the fewest parameters to vary [18].

E-ISSN : 2540 - 8984

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Volume 07, Nomor 04, Desember 2022 : 1244 - 1253

Simple kriging using affine linear equation [19]:

$$Z^*(x) = m + \sum_{i=1}^{n} \lambda_i [Z(x_i) - m] \tag{10}$$

Where m is mean and n is amount of data.

The error estimator between the unknown truth and the kriging estimator is $\varepsilon = Z(x) - Z*(x)$. Kriging interpolation is obtained by minimizing squad error (ϵ 2). Can be described in the formula:

$$\sum_{i=i}^{n} \lambda_i C_{ij} = C_{0j} \tag{11}$$

Where C is covariance.

G. Mapping

After the kriging interpolation will produce a rainfall prediction map for the island of Java on April 1, 2022 - April 7, 2022 and April 2022 - September 2022.

III. Result and Analysis

A. Data

In this study, the data class consists of 6 categories. The purpose of giving this category is to represent classes with a range of rain values according to BMKG guidelines. The categories are 0 is cloudy, 1 light rain, 2 moderate rain, 3 heavy rain, 4 very heavy rain, 5 extreme rain.

TABLE II
RAINFALL CLASS LABEL

Value Range	Category	Class
RR < 0	Cloudy	0
0 < RR < 20	Light rain	1
20 < RR < 50	Moderate rain	2
50 < RR < 100	Heavy rain	3
100 < RR < 150	Very heavy rain	4
RR > 150	Extreme rain	5

B. Testing

Testing is divided into two models, namely daily and monthly. Each model is classified using nave Bayes and simple kriging interpolation.

1. Naïve Bayes Classification

In the classification process the confusion matrix is used to measure the performance of each model. The parameters used are accuracy, precision, recall, and f1-score.

TABLE III
CLASSIFICATION EVALUATION RESULTS

Model	Accuracy	Precision	Recall	F1-score
Daily	67%	64%	67%	65%
Monthly	88%	92%	88%	90%

Based on the test results, it was found that the difference in accuracy values was quite high between the daily and monthly models. The daily model got 67% accuracy, 64% precision, 67% recall, and 65% f1-score. The monthly model got 88% accuracy, 92% precision, 88% recall, and 90% f1-score. The difference in accuracy values between the daily model and the monthly model is because the daily model uses more data than the monthly model, while the Nave Bayes method is more suitable for fewer data. So for the daily model, the nave Bayes method is not suitable.

2. Semivariogram Model Selection

The best semivariogram model was selected for simple kriging interpolation. There are 3 models tested, namely spherical, exponential, and Gaussian. Model selection is based on the smallest RMSE value. The table below is the test results.

TABLE IV
DAILY MODEL RMSE CALCULATION

Daily	1 April 2022	2 April 2022	3 April 2022	4 April 2022	5 April 2022	6 April 2022	7 April 2022
Spherical	0,845	1,061	1,147	1,034	0,867	1,053	1,142
Exponential	0,888	1,098	1,149	1,036	0,87	1,046	1,179
Gaussian	0,853	1,035	1,146	1,037	0,881	1,169	1,143

The daily model RMSE calculation obtained three minimum values for the spherical model on April 1, 4, 5, and 7, the exponential model for one on April 6, and two for the Gaussian model on April 2 and 3. So, the semivariogram model used for mapping is a spherical model because it has the highest minimum RMSE value compared to other models.

TABLE V
MONTHLY MODEL RMSE CALCULATION

Monthly	April	May	June	July	August	September
Spherical	0,271	0,215	0,449	0,283	0,4	0,427
Exponential	0,271	0,215	0,449	0,282	0,4	0,427
Gaussian	0,271	0,215	0,445	0,282	0,4	0,427

The monthly RMSE calculation model obtained the minimum value in June and July for the gaussian model and the same minimum value in July for the exponential model. For other months, the RMSE value obtained is the same. So, the selection of the semivariogram model for the monthly model is the Gaussian model. Based on the analysis of the RMSE semivariogram test. In the daily model, the best semivariogram is the spherical model, while in the monthly model, the best semivariogram is the gaussian model.

C. Testing Result

There are 2 results in this study, namely prediction and mapping results in both daily and monthly models.

1. Prediction Result

Based on the nave Bayes classification, the prediction results are obtained for 7 days starting from April 1, 2022 - April 7 2022 and prediction results for 6 months starting from April 2022 - September 2022.

TABLE VI
DAILY MODEL PREDICTION

Location	01-Apr-22	02-Apr-22	03-Apr-22	04-Apr-22	05-Apr-22	06-Apr-22	07-Apr-22
Geophysic_Tangerang	2	5	2	2	2	2	5
South_Tangerang_Climatology	1	1	1	1	1	1	1
Meteorology_Budiarto	2	2	2	2	2	2	5
Meteorology_Maritim_Serang	2	2	2	2	2	2	2
Meteorology_Soekarno_Hatta	1	1	1	1	1	1	1
Geophysic _Sleman	1	1	1	1	1	2	2
Meteorology_Kemayoran	1	1	1	1	1	1	2
Meteorology_Maritim_Tanjung_Priok	1	1	1	1	2	2	1
Geophysic _Bandung	1	1	1	1	1	1	1
Climatology_Bogor	1	1	1	1	1	1	1
Meteorology_Citeko	1	1	1	1	1	1	0
Meteorology_Kertajati	1	1	1	1	1	2	2
Meteorology_Tegal	0	0	1	1	1	1	1
Meteorology_Tunggul_Wulung	5	5	5	5	5	5	2

Meteorology_Maritim_Tanjung_Emas	1	1	0	0	1	1	1
Meteorology_Ahmad_Yani	1	1	1	1	2	5	1
Climatology_Semarang	1	1	1	1	1	2	2
Meteorology_Sangkapura	1	1	1	1	1	1	1
Meteorology_Perak_I	1	1	1	1	1	1	2
Meteorology_Maritim_Tanjung_Perak	1	1	1	1	1	1	0
Meteorology_Kalianget	1	1	1	1	1	1	0
Meteorology_Juanda	1	1	1	1	1	1	1
Meteorology_Banyuwangi	1	1	5	1	1	1	1
Climatology_Malang	1	1	1	1	1	1	1
Geophysic _Pasuruan	1	1	1	1	1	2	2
Geophysic _Nganjuk	1	1	1	1	2	4	1
Geophysic _Malang	1	1	2	4	1	1	1

TABLE VII
MONTHLY MODEL PREDICTION

Location	April	May	June	July	August	September
Geophysic _Tangerang	1	1	1	1	0	0
South_Tangerang_Climatology	2	1	1	1	1	0
Meteorology_Budiarto	1	1	1	1	0	0
Meteorology_Maritim_Serang	1	1	0	1	0	0
Meteorology_Soekarno_Hatta	1	1	1	1	1	1
Geophysic _Sleman	1	1	1	1	0	1
Meteorology_Kemayoran	2	1	1	1	1	1
Meteorology_Maritim_Tanjung_Priok	1	0	0	0	0	0
Geophysic _Bandung	1	1	1	1	1	1
Climatology_Bogor	1	1	1	1	1	1
Meteorology_Citeko	1	1	1	1	1	1
Meteorology_Kertajati	1	1	0	1	1	1
Meteorology_Tegal	1	1	1	1	1	0
Meteorology_Tunggul_Wulung	1	1	2	2	1	1
Meteorology_Maritim_Tanjung_Emas	1	1	1	1	1	1
Meteorology_Ahmad_Yani	1	0	1	1	1	1
Climatology_Semarang	1	1	1	1	1	1
Meteorology_Sangkapura	1	1	1	1	1	1
Meteorology_Perak_I	1	1	1	1	1	1
Meteorology_Maritim_Tanjung_Perak	1	1	1	1	1	1
Meteorology_Kalianget	1	1	1	1	1	1
Meteorology_Juanda	1	1	1	1	1	1
Meteorology_Banyuwangi	1	1	0	1	1	1
Climatology_Malang	1	1	1	1	1	1
Geophysic _Pasuruan	1	1	1	1	1	1
Geophysic _Nganjuk	1	1	1	1	1	1
Geophysic _Malang	1	1	1	1	1	1

2. Mapping Result

Mapping using kriging interpolation is used to make the mapping results easier to read by the public. Because it is clearly visible areas that experience rainfall. Simple kriging performs rainfall interpolation for each point location of the existing weather station.

The results of the daily model mapping are as follows:

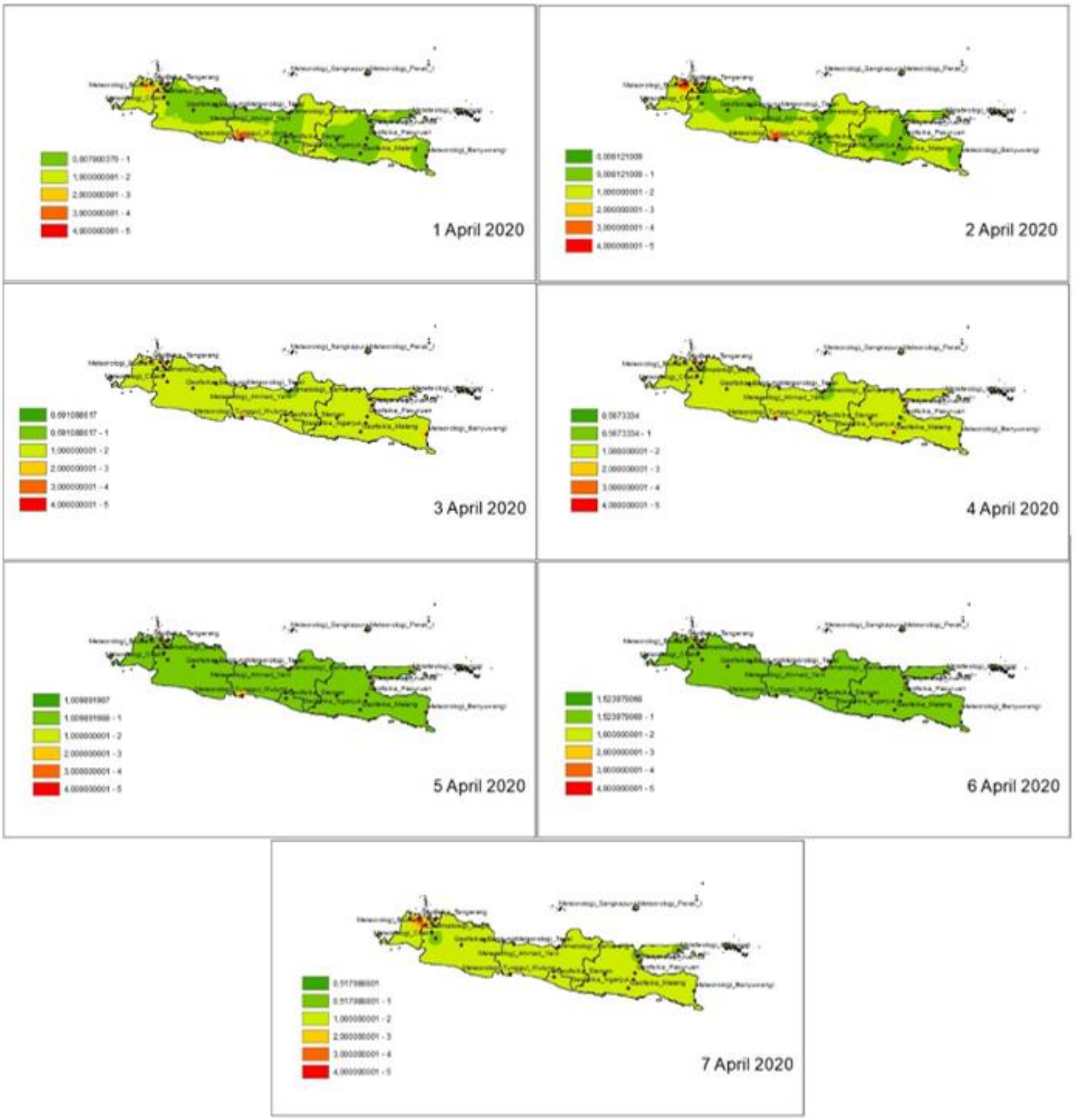

Figure 2 Daily mapping

Figure 2 shows the results of daily model predictions. In the daily model, the data obtained is more varied. On April 5 and 6, there was light rain for almost the entire island of Java. On April 3, 4, and 7, almost the entire island of Java had moderate rain and only about 5% of the area had light rain and about 5% heavy rain. On April 1, about 60% of the area had light rain, 35% moderate rain, and 5% heavy rain. On April 2, about 40% of the area had light rain, 55% moderate rain, and 7% heavy rain. There are several areas that experience high rainfall, for example in the southwest part of West Java on April 1 and April 2 there is a red zone that indicates high rainfall. For in other areas, the rainfall is relatively normal.

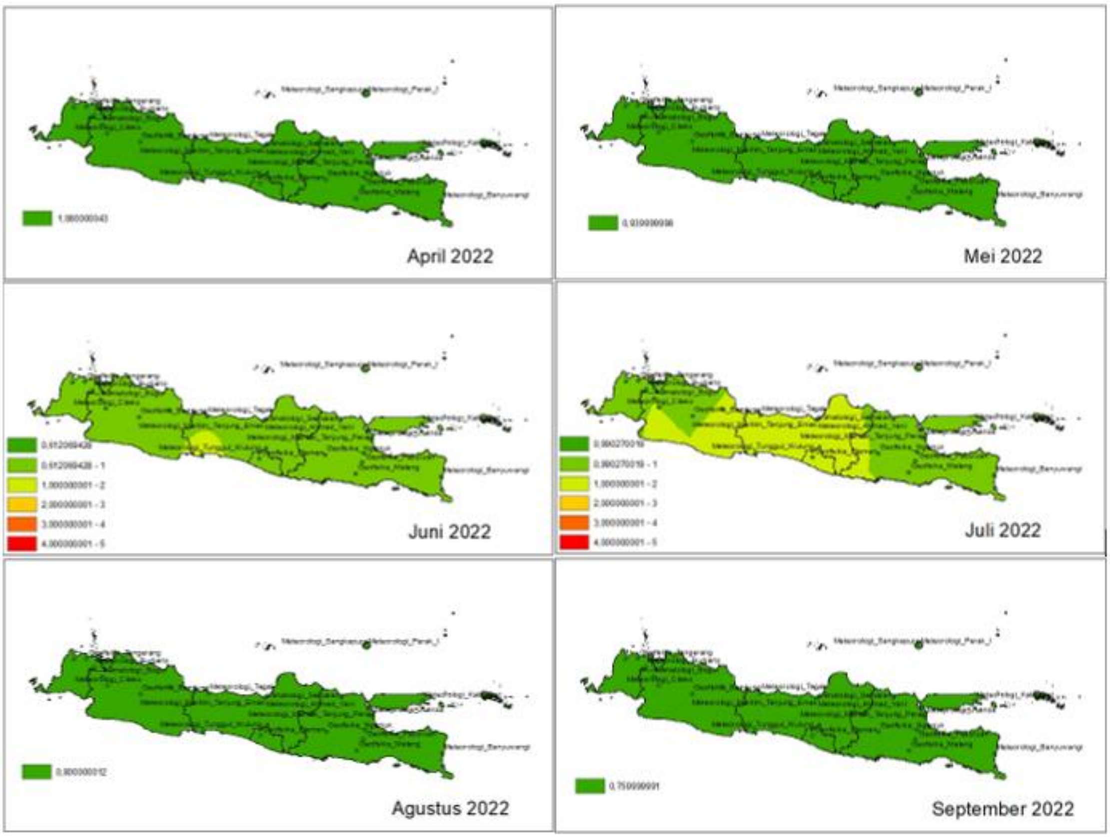

Figure 3 Monthly mapping

Figure 3 shows the results of monthly model predictions. In the monthly model, the data obtained are not very varied. In April, May, August, and September almost 100% of the area experiences light rain. In June there is about 7% moderate rain and the rest is light rain. In July around 55% moderate rain and 45% light rain.

The daily accuracy value is 67% and the monthly accuracy value is 88% with mapping results that are easy to read. Compared to previous studies[6], [7], [8], and [9] it only show prediction results, and does not display mapping. While research [10] and [11] only shows mapping that does not adjust the shape of the map. In comparison, this study produces predictive and mapping values that adjust the shape of the map.

IV. Conclusion

This study combines the nave Bayes method for classification and simple kriging for interpolation, especially for Java. The daily model has an accuracy of 67% and the monthly model has an accuracy of 88%. It can be concluded that the nave Bayes method is not suitable for the daily model but is good for the monthly model. This can happen because the daily model has a lot of data, while the Naive Bayes model is devoted to small data. The final result of the research on the daily model shows that in the southwestern part of Central Java province and the northern part of Banten province, rainfall is higher than other areas. Meanwhile, in the monthly model, Central Java province experiences higher rainfall than other provinces. Based on the results of the daily mapping, early April experienced quite high rainfall in some areas and during monthly mapping in April the rainfall was relatively normal and the same for each area, this shows that after April 7 the rainfall experienced by each area was relatively normal or low.

References

[1] Y. Andrian and E. Ningsih, “Prediksi Curah Hujan Di Kota Medan Menggunakan,” *Semin. Nas. Inform.*, pp. 184–189, 2014.

[2] D. Eko Nuryanto, “KARAKTERISTIK CURAH HUJAN ABAD 20 DI JAKARTA BERDASARKAN KEJADIAN IKLIM GLOBAL,” *Meteorogi dan Geifisika*, p. 11, 2013.

[3] T. D. Mirawati, H. Yasin, and A. Rusgiyono, “Prediksi Curah Hujan dengan Metode Kalman Filter (studi kasus di Kota Semarang Tahun 2012),” *J. Gaussian*, vol. 2, no. 3, pp. 239–248, 2013, [Online]. Available: https://media.neliti.com/media/publications/137182-ID-prediksi-curah-hujan-dengan-metode-kalma.pdf

[4] i gede aris gunadi, “Klasifikasi Curah Hujan di Provinsi Bali Berdasarkan Metode Naïve Bayesian,” *Wahana Mat. dan Sains J. Mat. Sains, dan*

Pembelajarannya, vol. 12, no. 1, pp. 14–15, 2018, [Online]. Available: https://ejournal.undiksha.ac.id/index.php/JPM/article/view/pril2018-2

[5] I. G. W. Sena, J. W. Dillak, P. Leunupun, and A. J. Santoso, "Predicting Rainfall Intensity using Naïve Bayes and Information Gain Methods (Case Study: Sleman Regency)," *J. Phys. Conf. Ser.*, vol. 1577, no. 1, pp. 0–6, 2020, doi: 10.1088/1742-6596/1577/1/012011.

[6] S. Triyanto, A. Sunyoto, and M. R. Arief, "Analisis Klasifikasi Bencana Banjir Berdasarkan Curah Hujan Menggunakan Algoritma Naïve Bayes," *JOISIE (Journal Inf. Syst. Informatics Eng.*, vol. 5, no. 2, pp. 109–117, 2021, doi: 10.35145/joisie.v5i2.1785.

[7] N. W. Zamani and S. S. M. Khairi, "A comparative study on data mining techniques for rainfall prediction in Subang," *AIP Conf. Proc.*, vol. 2013, 2018, doi: 10.1063/1.5054241.

[8] A. K. Sharma, S. Chaurasia, and D. K. Srivastava, "Supervised Rainfall Learning Model Using Machine Learning Algorithms," *Adv. Intell. Syst. Comput.*, vol. 921, 2020.

[9] A. U. Azmi, A. F. Hadi, D. Anggraeni, and A. Riski, "Naive bayes methods for rainfall prediction classification in Banyuwangi," *J. Phys. Conf. Ser.*, vol. 1872, no. 1, 2021, doi: 10.1088/1742-6596/1872/1/012028.

[10] S. S. Prasetiyowati, Y. Sibaroni, and S. Carolina, "Prediction and Mapping of Air Pollution in Bandung Using Generalized Space Time Autoregressive and Simple Kriging," *2020 Int. Conf. Data Sci. Its Appl. ICoDSA 2020*, 2020, doi: 10.1109/ICoDSA50139.2020.9212820.

[11] M. Hassim, A. Yuzir, M. N. Razali, F. C. Ros, M. F. Chow, and F. Othman, "Comparison of Rainfall Interpolation Methods in Langat River Basin," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 479, no. 1, pp. 0–9, 2020, doi: 10.1088/1755-1315/479/1/012018.

[12] G. Ali *et al.*, "Spatial–temporal characterization of rainfall in Pakistan during the past half-century (1961–2020)," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, 2021, doi: 10.1038/s41598-021-86412-x.

[13] B. Province, F. F. Firoozabad, and H. A. Ardakani, "Evaluation and Comparison of Interpolation and Linear Regression Methods to Determine the Spatial Distribution of Precipitation in Chaharmahal and Bakhtiari Province, Iran," vol. 8, no. 1, pp. 218–232, 2022, doi: 10.22034/JEWE.2021.291533.1584.

[14] M. Muthmainnah, M. Ashar, I. M. Wirawan, and T. Widiyaningtyas, "Time Series Forecast for Rainfall Intensity in Malang City with Naive Bayes Methodology," *3rd Int. Conf. Sustain. Inf. Eng. Technol. SIET 2018 - Proc.*, pp. 137–141, 2018, doi: 10.1109/SIET.2018.8693171.

[15] M. T. Ali Haghpanah Jahromi, "A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features," *2017 Artif. Intell. Signal Process. Conf.*, no. 1, pp. 134–139, 2017.

[16] A. M. Oliver and R. Webster, *Basic Steps in Geostatistics:The Variogram and Kriging*. 2015.

[17] F. R. Kodong and J. Fajar, "Prediction of wind disaster using kriging spatial interpolation and internet of things," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 620, no. 1, 2019, doi: 10.1088/1757-899X/620/1/012098.

[18] B. Rebholz and M. Almekkawy, "Efficacy of Kriging Interpolation in Ultrasound Imaging; Subsample Displacement Estimation," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2020-July, pp. 2137–2141, 2020, doi: 10.1109/EMBC44109.2020.9175457.

[19] Y. Z. Ma, *Quantitative Geosciences: Data Analytics, Geostatistics, Reservoir Characterization and Modeling*. 2019. doi: 10.1007/978-3-030-17860-4.