

Accuracy assessment of spatial interpolations methods using ArcGIS

Yahya Darmawan^{1*}, Munawar¹, Dwiki Anugerah Atmojo¹, Huda Wahyujati¹, and Lamtupa Nainggolan^{2,3,4}

¹Climatology Department, School of Meteorology, Climatology, and Geophysics (STMKG), Banten, Indonesia

²Taiwan International Graduate Program (TIGP), Earth System Science Program, Academia Sinica, Taipei, Taiwan

³Graduate Institute of Applied Geology, National Central University (NCU), Taoyuan City, Taiwan

⁴Education and Culture Board of Medan City Government (PEMCO MEDAN), North Sumatra, Indonesia

Abstract. The Inverse Distance Weight (IDW), Spline, and Ordinary Kriging (OK) methods are widely used for interpolating spatial data. Many GIS software options are available, particularly for spatial interpolation. This study aims to assess the accuracy of interpolation approaches based on Climate Hazards Group Infrared Precipitation with Station (CHIRPS) data in ArcGIS 10.8. The research used monthly precipitation data extracted from 147 rain gauge station locations in North Sumatra Province. Additionally, eight stations from BMKG's network were excluded from the interpolation process and instead used as validation data. Validation was conducted by calculating error statistics between the predicted and actual data at these excluded points. The predicted data was generated by interpolating values from the 147 rain gauge stations, excluding the eight points corresponding to the validation data locations. The interpolation was done using ArcGIS 10.8 with the IDW, Spline, and Kriging toolbox. A comparison was performed by assessing error parameters and Pearson correlation coefficients. According to the accuracy assessment, both IDW and OK show comparable performance; both are better than spline interpolation. In addition, the box plot chart demonstrated that the spline technique had a tendency to generate outlier values along the perimeters of the study area.

1 Introductions

The rainfall observed in the North Sumatra region shows a significant variation, primarily influenced by the area's topography and geography [1]. Generally, there is a noticeable decrease in average rainfall as a trend from west to east, resulting in pronounced disparities between the western and eastern sectors. Specifically, in the east segment, covering the east coast and slopes of North Sumatra, the peak of the rainy season occurs in October, a month ahead of the typical rainy season [2]. Topographical characteristics and localized weather systems substantially shape rainfall distribution within a specific locality. These factors significantly impact the volume of rain and how it is spread across the area[3].

In disaster management, by examining the patterns of rainy and dry seasons, it becomes essential to have accurate and detailed rainfall data to generate reliable insights [4]. However, due to the restricted number of rain gauge stations, obtaining observation data for every point on North Sumatra Island is not always possible. In this case, spatial interpolation comes into play, enabling the estimation of values for unobserved attributes. Moreover, this technique is also employed to predict variables that cannot be directly measured at numerous locations[5].

Spatial interpolation involves transforming a collection of point data into comprehensive surface data. Point data sets consist of values limited to specific locations, often corresponding to fieldwork sites within the designated study area [6]. There are several spatial interpolation methods available in the literature and GIS software. Accuracy assessment and evaluation from an interpolated area is challenging and mostly ignored. An accuracy assessment was the crucial method used to assess the final interpolation in the spatial interpolation process. Most quantitative methods use the statistical approach to evaluate the overall performance.

What are the characteristics of the three spatial interpolation methods being compared? Interpolated output cannot be judged only by a numeric index because many characteristics of spatial value are limited to be evaluated by quantitative assessments. Nowadays, there is no clear guideline that can explain the best estimation method that is appropriate for all situations [4]. This research aims to assess the accuracy among the Inverse Distance Weight (IDW), Ordinary Kriging (OK), and Spline interpolation methods in ArcGIS 10.8 to understand the characteristics of the interpolation methods.

* Corresponding author: yahya.darmawan@bmkg.go.id

2 Materials and methodology

2.1 Study area

This research was conducted in North Sumatra Province, located between 1°-4° North Latitude and 98°-100° East Longitude, with a total area of 72,981.23 km². The study consisted of two activities: data collection and data processing. The research activities used monthly rainfall data from 147 rainfall stations in North Sumatra Province to be interpolated and monthly rainfall data from 8 MKG observation stations in North Sumatra Province as test data. The monthly rainfall data was acquired from CHIRPS satellite data from 2017 to 2021.

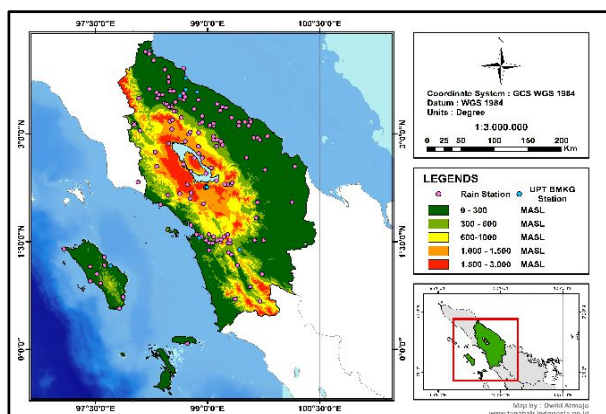


Fig. 1. Study area: North Sumatra Province.

2.2 Data used

Monthly rainfall data was obtained from the CHIRPS satellite, which can be downloaded from the website <https://data.chc.ucsb.edu/products/CHIRPS-2.0>. The obtained image was then downloaded in GeoTIFF (.tiff) data format. After downloading, the data was extracted into ArcGIS 10.8 software, which generates the data into a table. The data was then trimmed using Ms. Excel software to obtain data specific to the desired area. The interpolation process was done using ArcGIS 10.8 software, using the data cut to the specified location. After interpolation, the data could be analysed using error parameters and Pearson's correlation.

The use of CHIRPS image data is based on the collaboration of USGS Earth Resources Observation and Science (EROS) scientists to provide a collection of data that focuses on merging satellite-based rainfall improvement models from NASA and NOAA that have high-resolution grids of 0.05°. CHIRPS is suitable for accurate and precise data needs and can eliminate biases. Unlike the observation data results recorded and processed by observer stations, the availability of the data is difficult to obtain, and there are errors in recording and processing.

The accuracy of CHIRPS data can be proven by testing the correlation between CHIRPS data and MKG observer station data using Pearson correlation. Pearson correlation measures the strength and direction of the linear relationship between two variables. Two variables

are said to be correlated if changes in one variable cause changes in the other, either in the same or opposite direction.

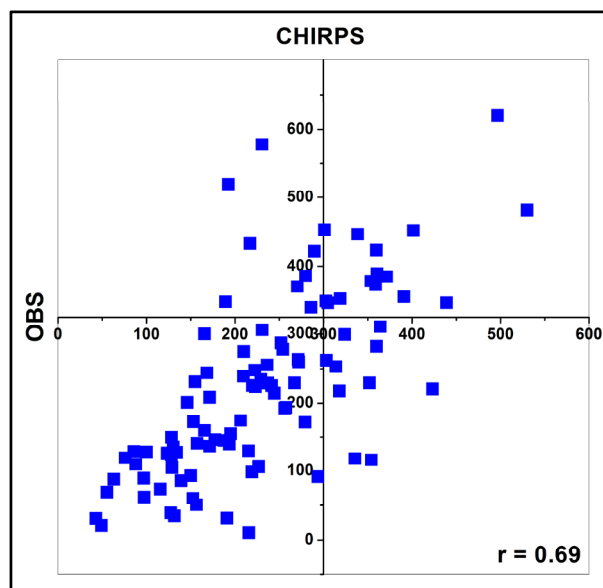


Fig. 2. Correlation coefficient (r): Rainfall Obs. vs CHIRPS.

The correlation coefficient (r) value in Figure 2 stands at 0.69, signifying a robust positive linear connection between the two datasets. The correlation coefficient (r) suggests that the CHIRPS data effectively represents the monthly rainfall data sourced from the observation stations. Consequently, the monthly rainfall data utilized in this study relies on the CHIRPS dataset. The correlation coefficient is a gauge for assessing the extent of association between variables [7]. The correlation coefficient's value ranges between $-1 < 0 < 1$, where an r of -1 indicates a complete negative correlation, implying a negligible influence of variable X (observed) on variable Y (CHIRPS). Conversely, an r of 1 signifies an absolute positive correlation, indicating a substantial impact of variable X (observed) on variable Y (CHIRPS) [8].

2.3 Methods

The schematic works for achieving the objective of this study can be seen in Figure 3. These steps encompass tasks like implementing spatial interpolation techniques, conducting accuracy assessment through statistical methods, and performing method validation.

2.3.1 Spatial interpolation

Every interpolation method employs different statistical techniques to estimate points surrounding the sample data. Below are the statistical analyses specific to each interpolation method.

a. Inverse Distance Weight (IDW)

The IDW technique is a traditional interpolation approach that employs distance as a weighting factor. This distance pertains to the separation between the data point (sample) and the target area for estimation. Consequently, when the sampling point and the target area for estimate are closer

in distance, a higher weight is assigned, and conversely [9]. The equation for the IDW interpolation model is presented as follows:

$$Z_0 = \frac{\sum_{i=1}^S Z_i \frac{1}{d_i^k}}{\sum_{i=1}^S \frac{1}{d_i^k}} \quad (1)$$

Description:

Z_0 = Estimated value at point 0

Z_i = Value of Z at control point i

d_i = Distance between point i and point 0

k = The larger the k , the greater the influence of neighbouring points

S = Number of S points used

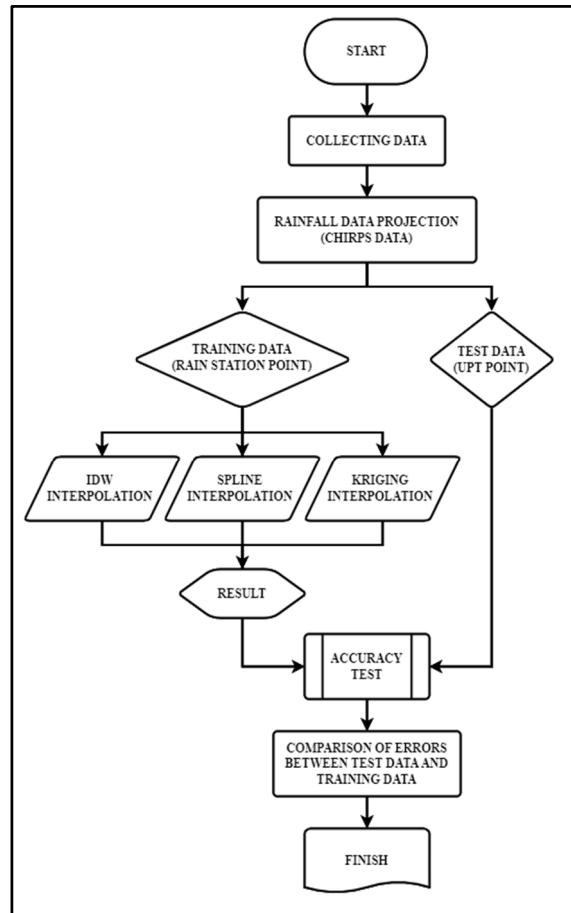


Fig. 3. Research Scheme Works.

b. Ordinary Kriging (OK)

The Kriging technique is a widely employed approach for examining geographic data using collected sample data. These sampled data points are typically gathered from outliers or specific points. This technique is applied to approximate the values of regional variables at un-sampled points, utilizing information from neighbouring sampled points while accounting for spatial correlation in the data. $Z(S_i)$ denotes a random variable at point S_i , where i assumes values of 1, 2, 3, and so forth up to n . The Kriging estimator $Z'(s)$ for $Z(s)$ can be formulated as follows [6]:

$$Z'(s) - m(s) = \sum_{i=1}^N \lambda_i [Z(S_i) - m(S_i)] \quad (2)$$

Description:

S = Location for estimation

S_i = One of the nearby data locations

$m(s)$ = Expected value of $Z(s)$

$m(S_i)$ = Expected value of $Z(S_i)$

λ_i = Weighting factor

n = Number of sample data used for estimation

If at each location $Z(s)$ there is an estimator error $e'(s)$, then the difference between the estimated value $Z'(s)$ and the value of $Z(s)$ is defined as follows [10]:

$$e'(s) = Z'(s) - Z(s) \quad (3)$$

c. Spline

The spline technique involves making value estimates through a mathematical function that minimizes the overall surface curvature. Within ArcGIS, spline interpolation is encompassed within radial basis functions (RBF). Although frequently applied in GIS, this method is suitable primarily when dealing with data featuring limited variability. RBF finds extensive use in forecasting seasonal time series data like rainfall, river flow, and agricultural production. The spline interpolation method is capable of projecting both minimum and maximum values while introducing a data stretching effect [9]. The equation utilized in spline interpolation employs a surface interpolation formula, as depicted by the subsequent equation:

$$S_{(x,y)} = T_{(x,y)} + \sum_{j=1}^N \lambda_j R(r_j) \quad (4)$$

Description:

j = 1, 2, 3, ..., n

N = Number of points

λ_j = Coefficients obtained from linear equation system

r_j = Distance between point i and point j

$T(x,y)$ and $R(r)$ defined differently, based on the selection method (regularized spline and tension spline).

2.3.2 Accuracy assessment

The process of evaluating the accuracy of spatial interpolation involves a sequence of steps. Initially, the Pearson correlation (r) is computed for each employed interpolation method. This correlation coefficient (r) assesses the intensity and direction of the linear association between two variables. Subsequently, the level of error for each interpolation method is determined using statistical error metrics like NMSE (Normalized Mean Square Error), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Squared Error), MSE (Mean Squared Error), and MAE (Mean Absolute Error) applied to the processed rainfall data. In the last, the data will be showed in a box plot (a whisker plot) to represent the dataset's distribution that provides a display the spread, tendency, and outliers distribution from three methods.

a. NMSE (Normalized Mean Square Error)

NMSE serves as an evaluation metric for assessing regression performance, measuring the disparity between the predicted and actual values of a regression model. It is computed by dividing the Mean Square Error (MSE) by the variance of the target variable.

$$NMSE = \frac{n \sum_{i=1}^n (y_k - o_k)^2}{\sum_{i=1}^n y_k \sum_{i=1}^n o_k} \quad (5)$$

Where:

- y_k : Observed data at location k
- o_k : Predicted data at location k
- n : The total number of data

Typically expressed in percentage form, a lower NMSE value indicates enhanced performance of the regression model. NMSE proves useful for evaluating and comparing the effectiveness of various regression models using a standardized data scale.

b. MAPE (Mean Absolute Percentage Error)

MAPE is computed by dividing the absolute error for each specific time period by the corresponding observed actual value within that period. Subsequently, the average of these absolute percentage errors is determined. In essence, MAPE represents the mean of absolute percentage errors, categorizing errors according to their origins [11].

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{y_k - o_k}{y_k} \right| \times 100\% \quad (6)$$

Where:

- y_k : Observed data at location k
- o_k : Predicted data at location k
- n : The total number of data

c. MAE (Mean Absolute Error)

MAE (Mean Absolute Error) serves as a metric for assessing accuracy by calculating the average absolute discrepancy between observed values and the estimated results obtained through interpolation.

$$MAE = \frac{1}{n} \sum_{k=1}^n |y_k - o_k| \quad (7)$$

Where:

- y_k : Observed data at location k
- o_k : Predicted data at location k
- n : The total number of data

d. RMSE (Root Mean Squared Error)

RMSE (Root Mean Squared Error) is a technique frequently employed to assess machine learning algorithms, encompassing those that exceed the complexity of linear regression. Broadly speaking, this parameter quantifies the square root of the mean of squared deviations between observed data and predictions. Given that RMSE entails squaring the error in estimation, it proves to be more responsive to significant errors compared to MAE [12].

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - o_k)^2} \quad (8)$$

Where:

- y_k : Observed data at location k
- o_k : Predicted data at location k
- n : The total number of data

e. MSE (Mean Squared Error)

MSE (Mean Squared Error) is the average of squared errors. Since MSE squares the error, it places more weight on larger errors. According to Suryaningrum [13], MSE is another method for evaluating forecasting methods. Each error or residual is squared. This approach allows for large forecasting errors because errors are squared [13].

$$MSE = \frac{1}{n} \sum_{k=1}^n (y_k - o_k)^2 \quad (9)$$

Where:

- y_k : Observed data at location k
- o_k : Predicted data at location k
- n : The total number of data

The best values for NMSE, MAPE, MAE, RMSE, and MSE are typically those that are lower and closer to 0. In simpler terms, when these values are lower, it signifies better performance and accuracy of the model.

3. Results and discussion

3.1. Inverse Weighted Distance (IDW)

The Inverse Distance Weighting (IDW) technique is a method for interpolating surfaces, grounded in the concept that input points can serve as central cells with random or uniform distribution. The IDW approach predicts attribute values at unobserved points using a linear combination of the sample values. This calculation is determined based on the inverse distance function between the points [14]. The interpolation outcomes of the IDW method are depicted in Figure 4.

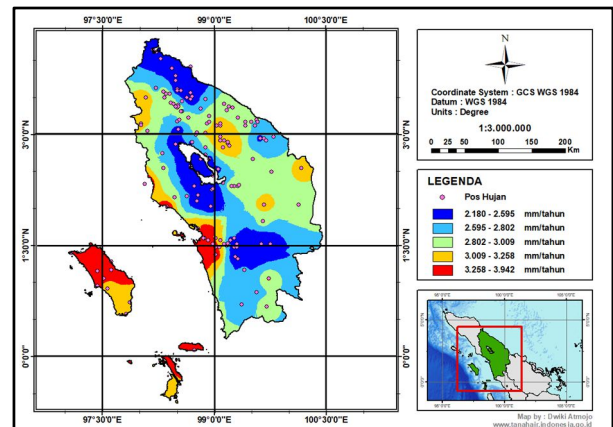


Fig. 4. IDW interpolation.

Based on the data depicted in Figure 4, the results derived from the IDW interpolation reveal distinctive

patterns of monthly rainfall within North Sumatra Province. Areas marked by elevated terrains or mountainous landscapes showcase relatively modest annual precipitation, averaging between 160 and 232 mm monthly. In contrast, the coastal belt, particularly the eastern littoral, encounters more substantial monthly rainfall, typically from 210 to 255 mm. Notably, the western coastal zones, including the surrounding islands, exhibit the most significant precipitation levels compared to all other zones, with values ranging from 255 to 335 mm monthly.

The effectiveness of the IDW interpolation technique is grounded in its ability to yield superior estimates for sample data in proximity compared to those situated at greater distances. This phenomenon arises from the method's utilization of the average of sample data. According to Ashraf et al. (1997), the inverse distance approach proves highly efficient in approximating sample values at specific locations. The effectiveness of the IDW interpolation is amplified when the sample points are abundant and closely spaced, as emphasized by the dense sample points employed [15]. Optimal outcomes are achieved when the sampling data closely aligns with local variations. However, if the samples are scarce and unevenly distributed, the anticipated results may not be realized as intended [16].

3.2. Ordinary Kriging (OK)

Kriging interpolation falls under the category of stochastic interpolation. This type of interpolation offers an estimate of the error associated with the predicted value, considering the presence of random error. Comparable to IDW, the kriging approach is a stochastic estimator employing linear weighting to approximate weights between data points. Initially formulated by D.L. Krige for estimating values in mining materials, this method assumes spatial correlation based on the distance and direction between data points [16]. The results of kriging interpolation are shown in Figure 5.

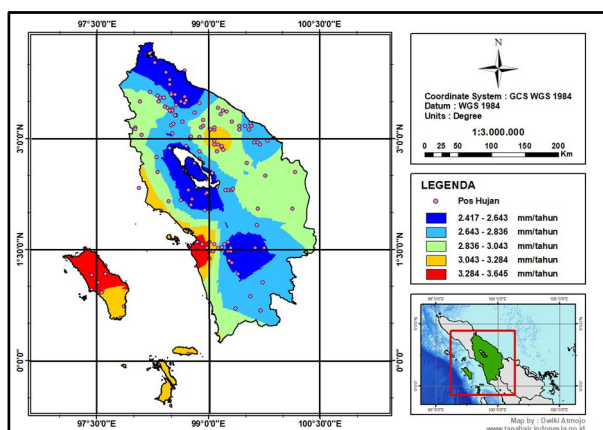


Fig. 5. OK interpolation.

Figure 5 of kriging interpolation results shows a rainfall distribution pattern similar to the IDW interpolation results but not as variable as the IDW interpolation results. The rainfall ranges from 170 to 231 mm/month in the highlands or mountainous areas, while

the eastern coastal area ranges from 210 to 255 mm/month. In the western coastal and island areas, the rainfall ranges from 255-320 mm/month, the highest monthly rainfall in each area. The kriging method has advantages and disadvantages. According to Lagueche (2006), its advantage is the ability to quantify the variance of the estimation value so that the level of estimation accuracy can be known. The kriging method can still be used even if no spatial correlation exists among the data. The disadvantage of kriging is that it assumes the data is usually distributed, while most field data do not meet this requirement. In addition, calculating semi-variances for one data set does not apply to other data sets. Therefore, estimating the semi-variogram will be difficult if the sample points are insufficient [17].

3.3. Spline

Interpolation Spline is a method that estimates values using mathematical functions to minimize the total curvature of the surface [18]. ESRI (1996) states that spline interpolation can calculate cell values based on the average data points from each sample [19]. The results of spline interpolation are shown in Figure 6.

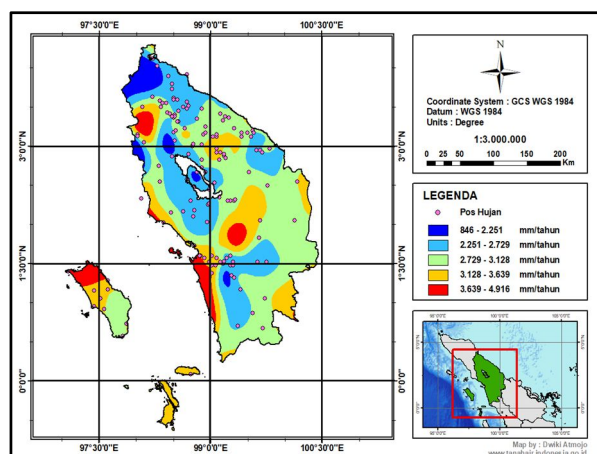


Fig. 6. Spline interpolation.

The rainfall ranges from 140-420 mm/month. The rainfall distribution pattern in the spline interpolation result is very diverse. The rain is relatively lower in the northern mountainous region (140-200 mm/month) than in the central and southern regions (200-232 mm/month). In the eastern coastal area, the rainfall ranges from 232 to 270 mm/month. The rainfall ranges from 270-420 mm/month in the western coastal and inland areas. The spline method passes the resulting surface through the sample points. The advantage of the spline method is its ability to produce a relatively accurate surface even with a small amount of data. According to Pasaribu and Haryana (2012), this method is ineffective when applied in situations where significant values differ at very close distances. The disadvantage of the spline method is that when neighboring sample points have very different values, the spline method may not work well. The inconsistency of the spline method is because the spline method uses slope calculations that change based on distance to estimate the shape of the surface [20].

3.4. Accuracy assessment

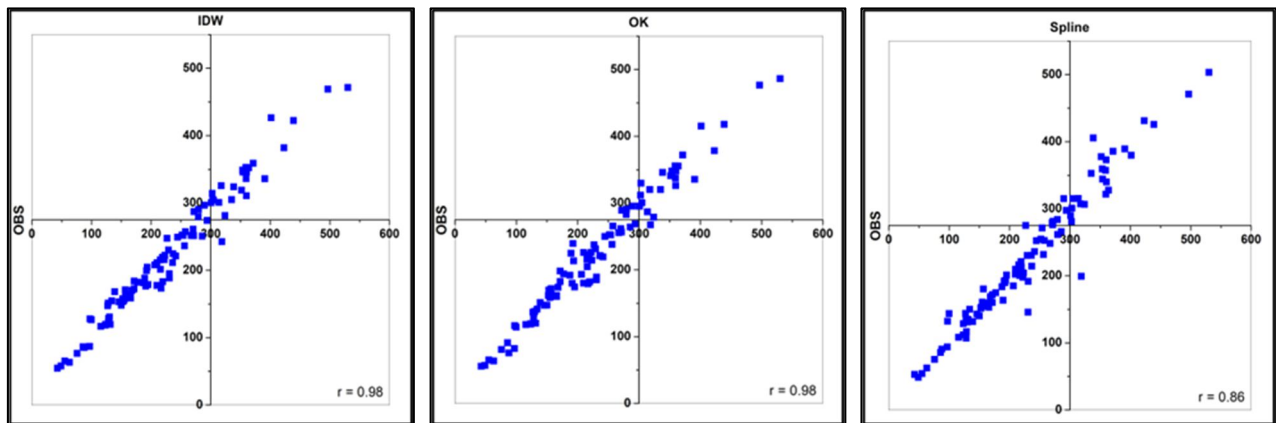


Fig. 7. The Pearson Correlation (r) for each interpolation.

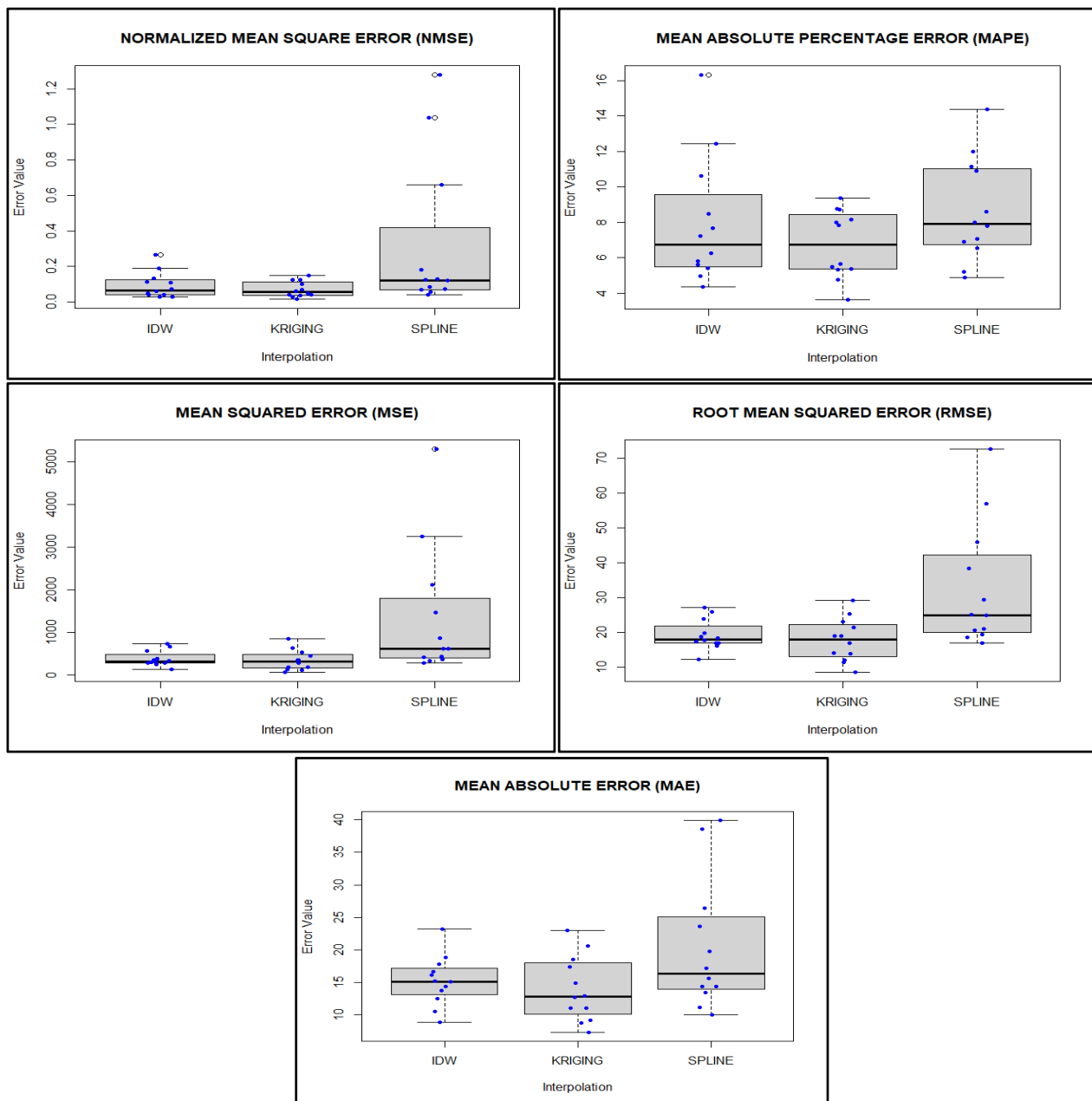


Fig. 8. Boxplot for each error parameter.

By comparing the results of Pearson correlation calculations for each interpolation method used, it can be seen that the correlation coefficient (r) values for the IDW and Kriging interpolation methods are the same, which is 0.98. Meanwhile, the correlation coefficient (r) value is 0.86 for the spline interpolation method (Figure 7). The correlation coefficient (r) shows that the correlation coefficient (r) for the IDW and OK interpolation methods is greater than the Spline interpolation method. The correlation coefficient value can be considered significant if it approaches 1. Therefore, to determine the accuracy and proper validation between the interpolation methods used, two processes with the same correlation coefficient (r) value, further error calculations are needed to validate which interpolation method is appropriate. The results of the error calculations are presented in Table 1. By comparing the values of NMSE, MAPE, RMSE, MSE, and MAE calculations in the three interpolation methods used (IDW, OK, spline), it can be seen that the IDW and OK interpolation method consistently shows smaller errors compared to spline interpolation methods.

Table 1. Calculation of error in each interpolation.

Parameter	Interpolation Method		
	IDW	OK	Spline
r	0,98	0,98	0,86
NMSE	0,01	0,03	0,03
MAPE	3,97	5,67	4,64
MAE	10,19	13,44	11,54
RMSE	13,58	16,15	16,11
MSE	184,60	260,83	259,71

The validity of these observations gains additional substantiation through a comprehensive analysis of the boxplot representations for each error calculation. Depicted in Fig 8, the boxplots unveil a consistent trend wherein the IDW and OK methodologies consistently yield notably diminished error values compared to the kriging and spline interpolation techniques. Such a disparity alludes to the fact that within the IDW and OK interpolation techniques, the range of diversity or distribution among the observation data points remains notably constrained. The utilization of boxplots, a graphical tool entrenched within the realm of descriptive statistics, further enriches our understanding of this phenomenon. These boxplots serve as a visual avenue to present numerical data in an intuitively comprehensive manner. Boxplots effectively unravel the intricate fabric of the data's variation and location by encapsulating pertinent statistical metrics such as medians, quartiles, and potential outliers. Noteworthy is that their efficacy extends to discerning and articulating shifts in variability and centrality across distinct data cluster. In this context, the boxplots serve as a lens through which we can discern the inherent disparities in the variability and distribution of errors among the different interpolation methods, underpinning the consistent trend favouring IDW and OK methods over kriging and spline methods

4 Conclusion

This research employed three distinct spatial interpolation techniques (inverse distance weighting, kriging, and spline) on a dataset of precipitation information derived from the CHIRPS satellite. The primary aim was to illustrate the complexities of the validation process for the outcomes of spatial interpolation. Three distinct validation approaches were employed: a statistical methodology based on Pearson correlation (r), error metrics, and a visualization in the form of a box plot chart. Each validation approach exposed distinct attributes of the spatial interpolation methods. Through quantitative evaluation, it was evident that both inverse distance weighting and kriging exhibited similar performance, outperforming the spline method. The box plot chart demonstrated that the spline technique had a tendency for generating outlier values along the perimeters of the study area.

References

1. HK, B.T., Climatology, 2nd Edition. Bandung–Penerbit ITB, (2004)
2. Prasetyo, B., H. Irwandi, and N. Pusparini, J of Science & Weather Modification Technology. **19** (1), (2018)
3. Enyew, B. and G. Steeneveld, J. Geol. Geosci. **3** (2), (2014)
4. Yi-Hwa, W. and H. Ming-Chih, (2016)
5. Kurniadi, H., et al. GEOTIK National Seminar 2018. (2018)
6. Bhowmik, A.K., (2012)
7. Siregar, S., Statistical Parameters for Quantitative Research. Bumi Aksara. Jakarta, (2013)
8. Sudjana, Statistical Method. 5 ed.Tarsito. (2005)
9. Hadi, B.S., **11** (2). (2013)
10. Alfiana, A.N., Ordinary Kriging methods in Applied Statistic. Universitas Negeri Yogyakarta, (2010)
11. Pakaja, F., A. Naba, and P. Purwanto, J of EECCIS). **6** (1), (2012)
12. Conway, D. and J. White, Machine learning for hackers." O'Reilly Media, Inc." (2012)
13. Suryaningrum, K. and S. Wijaya, **18**. (2015)
14. Hayati, F.D., IPB University, (2012)
15. Ashraf, M., J.C. Loftis, and K. Hubbard, **84** (3-4) (2017)
16. Pramono, G.H., .American Journal of Applied Sciences. **3**(6), (2006)
18. Tran, Q.B. and T.T. Nguyen, Assessment of the influence of interpolation techniques on the accuracy of digital elevation model. (2008)
19. Institute, E.S.R., ArcView Spatial Analyst: Advanced Spatial Analysis Using Raster and Vector Data. 1996: ESRI.
20. Pasaribu, J.M. and N.S. Haryani, J of Remote Sensing and Digital Image Processing. **9**(2). (2012)