BAREKENG: Journal of Mathematics and Its Applications
March 2024 Volume 18 Issue 1 Page 0627–0636
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss1pp0627-0636

RAINFALL MODELING USING THE GEOGRAPHICALLY WEIGHTED POISSON REGRESSION METHOD

Atiek Iriany[1*], Wigbertus Ngabu[2], Danang Ariyanto[3]

[1,3]Department Statistics, Faculty of Mathematics and Science, Brawijaya University
Veteran Street, Ketawanggede, Lowokwaru, Malang City, 65145, Indonesia
[2]Statistics Study Program, Faculty of Mathematics and Science, San Pedro University
Kupang City, 85142, Indonesia

Corresponding author e-mail: [1] **atiekiriany@ub.ac.id***

Article History:
Received: 1st October 2023
Revised: 20th January 2024
Accepted: 10th February 2024

Keywords:
GWR;
GWPR;
Rainfall.

ABSTRACT

Rainfall is an important parameter in understanding the climate and environment in the Malang Regency area. This research aims to model the distribution of rainfall in this region using the Geographically Weighted Poisson Regression (GWPR) method. GWPR is a spatial statistical approach that allows us to understand changes in inhomogeneous rainfall patterns throughout the Malang Regency area. Rainfall data collected from weather stations over several years was used in this study. We use GWR to study the relationship between various environmental factors, such as topography, vegetation, and land use, and rainfall distribution in Malang Regency. The results of the GWR analysis provide a deeper understanding of the spatial differences in the influence of these factors on rainfall. By applying GWR, we can find out how certain factors contribute to different rainfall patterns in certain regions. Rainfall modeling using the Geographically Weighted Poisson Regression (GWPR) method combines the power of Poisson regression in analyzing calculated data with the advantages of GWR in modeling spatial variability. GWPR allows us to identify and map rainfall distribution patterns that vary in geographic space. The main advantage of GWPR is its ability to provide local adjustments and capture the spatial variability associated with rainfall distribution. The results of the modeling analysis show that the GWPR is better, marked by the smallest AIC value, namely 336.84, compared to the generalized Poisson regression model, namely 337.76.

How to cite this article:

A. Iriany, W. Ngabu and D. Ariyanto., "RAINFALL MODELING USING THE GEOGRAPHICALLY WEIGHTED POISSON REGRESSION METHOD," *BAREKENG: J. Math. & App.*, vol. 18, iss. 1, pp. 0627-0636, March, 2024.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*
Research Article • Open Access

1. INTRODUCTION

Rainfall is an important hydrological phenomenon in the water cycle on Earth [1]. The variability and spatial distribution of rainfall have a major impact on various aspects of human life, including agriculture, water resources management, and flood risk mitigation. Therefore, rainfall modeling is essential in efforts to understand and anticipate changes in rainfall patterns in the future.

One main factor that influences water availability in an area. Rainfall distribution patterns vary greatly, depending on factors such as regional climate, topography, and land use [2]. Changing rainfall patterns can cause serious impacts, such as droughts that cause water crises or floods that damage infrastructure and the environment. With good rainfall modeling, we can understand the distribution patterns and variability of rainfall in a region, helping sustainable land use planning, water resource management, and developing disaster risk mitigation strategies [3].

In recent decades, advances in technology and data analysis have enabled the development of more sophisticated rainfall modeling methods. One interesting approach is the use of Geographically Weighted Poisson Regression (GWPR), a method that combines Poisson regression with spatial elements [4].

Poisson regression is a statistical method that is commonly used to analyze calculated data, such as the number of events or the frequency of events that occur in a certain time interval [5]. In the context of rainfall modeling, Poisson regression can be used to relate the dependent variable in the form of rainfall frequency with independent variables that influence rainfall, such as temperature, wind, and other factors. However, the Poisson regression model has limitations when applied to spatial data because it ignores spatial variations that can influence rainfall in various locations [6].

Geographically Weighted Regression (GWR) is a development of traditional regression models that takes into account spatial elements. GWR treats each location as a unique entity and allows regression coefficients to vary in geographic space [7]. In other words, GWR allows more detailed and in-depth analysis at the local level. In rainfall modeling, GWR can be used to overcome spatial variability problems related to rainfall distribution. By applying GWR, we can find out how certain factors contribute to different rainfall patterns in certain regions [8].

Rainfall modeling using the Geographically Weighted Poisson Regression (GWPR) method combines the power of Poisson regression in analyzing calculated data with the advantages of GWR in modeling spatial variability. GWPR allows us to identify and map rainfall distribution patterns that vary in geographic space [9]. The main advantage of GWPR is its ability to provide local adjustments and capture the spatial variability associated with rainfall distribution. The results of the GWPR analysis can provide more in-depth information about the factors that influence rainfall in various locations.

Rainfall modeling using the Geographically Weighted Poisson Regression (GWPR) method can make a significant contribution to our understanding of spatial rainfall distribution patterns. By applying this approach, it is hoped that we can make more informed decisions in water resource management and sustainable policy development.

However, it should be remembered that to produce an accurate and reliable model, high-quality rainfall data and appropriate variable selection are required. Rainfall modeling using the Geographically Weighted Poisson Regression (GWPR) method is a promising approach in hydrological and environmental analysis. GWPR combines Poisson regression techniques with more in-depth spatial analysis, allowing us to understand rainfall distribution patterns that vary in geographic space. With more accurate and spatially relevant rainfall modeling, it is hoped that we can take more effective steps in water resource management, flood risk mitigation, and environmental protection [10].

GWPR offers several important advantages in rainfall modeling: Improved Spatial Adjustment GWPR takes into account spatial variability in the relationship between independent variables and rainfall. This produces a model that is more accurate and appropriate to local conditions at each location. Identification of Spatial Patterns: GWPR allows the identification of rainfall distribution patterns that vary within a particular area. This helps in understanding hydrological dynamics and can provide new insights into water resource management.

Benefits for Planning and Decision Making: The results of GWPR can provide more relevant information for policymakers and water resource managers. More precise rainfall modeling can be used to optimize flood risk mitigation strategies, water resource management, and spatial planning.

2. RESEARCH METHODS

2.1 Data

The data collection technique for this research is to use secondary data taken from the digital publication of the East Java Meteorology, Climatology, and Geophysics Agency in May 2023. This research also uses astronomical location data, which includes the latitude and longitude of each sub-district in Malang Regency. The Malang Regency area consists of 33 sub-districts, with the variables used in this research being one dependent variable (Y), namely rainfall, and independent variables (X), namely altitude (X_1), temperature (X_2), and humidity (X_3).

2.2 Generalized Poisson Regression (GPR)

The Generalized Poisson Regression (GPR) model is a model that is suitable for counting data if there is over- or under-dispersion, namely, if the variance is greater or smaller than the mean [11]. So apart from the parameter μ, in GPR there is θ as a dispersion parameter. The GPR model is similar to the Poisson Regression model, but the GPR model assumes that the random component has a generalized Poisson (GP) distribution [12].
The generalized Poisson distribution can be seen in Equation (1) below [13]:

$$f(y,\mu,\theta)=\left(\frac{\mu}{1+\theta\mu}\right)^{y}\frac{(1+\theta y)^{y-1}}{y!}\exp\left(\frac{-\mu(1+\theta y)}{1+\theta\mu}\right), y=0,1,2...,\text{n} \quad (1)$$

The mean and variance of the GPR model are as follows:

$$E(y)=\mu, \text{var}(y)=\mu(1+\theta\mu)^{2} \quad (2)$$

If θ is equal to 0, then the GPR model will be an ordinary Poisson regression model. If θ is greater than 0, then the GPR model represents dispersion count data, and if θ represents count data, then it is under dispersion. The form of the GPR model is the same as the Poisson Regression model, namely as follows:

$$\mu_1=\exp(x_i^T\beta^*) \quad (3)$$

Where i is the unit of observation, namely i = 1, 2, ..., n

2.3 Geographically Weighted Poisson Regression (GWPR)

This GWPR model is a local linear regression model that produces local model parameter estimates for each point or location where the data is collected [14]. The GWPR model was developed from the GWR method, which is a technique that brings the framework from a simple regression model to a weighted regression model. In the GWPR model, the response variable y is predicted by predictor variables whose respective regression coefficients depend on the location where the data is observed. By notating the latitude and longitude coordinate vectors (u_i, v_i), the GWPR model can be written as follows [14].

$$y_i \sim Poisson\left[N_i\exp(\beta_0(u_i,v_i)+\sum_k\beta_k(u_i,v_i)x_{k,i})\right]$$
$$\mu_i=\exp(\beta_0(u_i,v_i)+\beta_1(u_i,v_i)x_{1i}+\beta_2(u_i,v_i)x_{2i}+...+\beta_k(u_k,v_k)x_{ik}+\varepsilon_i \quad (4)$$

The GWPR model parameter estimation was carried out using the Maximum Likelihood Estimation (MLE) method, namely by maximizing the likelihood function [15]. The MLE method is usually used if the distribution of the data being modeled is known. The likelihood function of GWPR is:

$$L(\beta)=\prod_{i=1}^{n}f(y_i)=\prod_{i=1}^{n}\frac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!} \quad (5)$$

Model parameter testing is carried out by testing the parameters partially. This test is to find out which parameters have a significant effect on the response variable for each location.

2.4 Testing Overdispersion

Overdispersion is a condition that occurs in Poisson regression analysis when the variance value of the response variable is greater than the average value [16]. The presence of cases of overdispersion will result in the deviation value of the regression model being large. Apart from that, cases of overdispersion can also result in the standard error of the resulting regression parameter estimates tending to be lower than it should be, so if the Poisson regression model is still used in overdispersion conditions, then the estimated parameters, which should not necessarily be significant, will be considered significant [17].

Overdispersion analysis of Poisson regression data can be seen from the Devian value divided by the degrees of freedom. If the quotient of both is greater than 1, then the data is said to have overdispersion. Statistical value for overdispersion testing in **Equation (6) [18]**

$$\phi = \frac{D}{db} = \frac{2\sum_{i=1}^{n}\left\{ y_i \ln\left(\frac{y_i}{\mu_i}\right) - (y_i - \mu_i) \right\}}{n-k-1} \qquad (6)$$

Where n is the number of observations and k is the number of parameters.

2.5 Moran's I Dependency Test

Moran's I coefficient is a development of Pearson correlation on univariate series data [19]. Moran's I coefficient is used to test spatial dependency or autocorrelation between observations or locations [20]. With the test statistics used in **Equation (7) [21].**

$$I = \frac{n\sum_{i=1}^{n}\sum_{i=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{i=1}^{n} w_{ij}(x_i - \bar{x})^2} \qquad (7)$$

The value of the I index is between -1 and 1. If I > Io, the data has positive autocorrelation. If I < Io, the data has negative autocorrelation.

2.6 Bandwidth and Weighting

Observations that are located close to location i will have more influence in forming model parameters at location i. Points that are within the radius's locations influence the model, which is weighted according to the function used. Selecting the optimum bandwidth will affect the accuracy of the model for the data by adjusting the variance and bias of the model [22]. The method used to select the optimum bandwidth is cross-validation (CV).

$$CV(s) = \sum_{i=1}^{n}(y_1 - \hat{y}_{\neq i}(G))^2 \qquad (8)$$

The weighting used in this research is the fixed bisquare kernel function. The fixed kernel method allows the optimal bandwidth value for each location to be the same or constant. If the data points are distributed regularly in the research area, then using the fixed method will be suitable for modeling.

Weighting is used to give different emphasis to different observations in generating parameters. Before the weighting is determined, $\boldsymbol{d_{ij}}$ must first be calculated, which is the distance between location $(\boldsymbol{u_i}, \boldsymbol{v_i})$ and location $(\boldsymbol{u_j}, \boldsymbol{v_j})$ using Euclidean distance, namely [23].

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (\mathrm{v}_i - v_j)^2} \qquad (9)$$

The Bisquare weighting function is used because it involves the element of distance between observed locations whose value is continuous in building the weighting matrix so that each location will receive a weight according to the distance between that location and the observed location. With the fixed kernel Bisquare weighting formula in the equation below [24],

$$W_j(i) = \begin{cases} \left[1-(\frac{d_{ij}}{b})^2\right]; d_{ij} < b \\ 0; d_{ij} \geq b \end{cases} \quad (10)$$

Information:

d_{ij} = Euclidean distance from location - i to location - j

b = Optimum bandwidth.

2.7 Selection of the Best Model

To get the best model between the Poisson Regression, Generalized Poisson Regression, and GWPR models, the best model was selected. The selection of the best model uses the AIC (Akaike's Information Criterion) criteria [25]. The AIC value is in line with the deviation value of the model. The smaller the deviation value, the smaller the error rate produced by the model, so that the model obtained becomes more precise. Therefore, the best model is the one with the smallest AIC and the smallest deviation [26]. The AIC value is formulated as follows:

$$AIC = D(\hat{\beta}) + 2K \quad (11)$$

Where $D(\hat{\beta})$ is the deviance of each model calculated, including the deviance of the Poisson Regression and GWPR models while K is the number of parameters in the model.

3. RESULTS AND DISCUSSION

3.1 Overdispersion Check

The presence of cases of overdispersion or underdispersion can be detected by checking the deviance value, or Chi-Square value, divided by the degrees of freedom in the goodness of the fit table. If the value is more than one, then there is overdispersion, and conversely, if it is less than one, then there is underdispersion. Overdispersion tests are presented in **Table 1**.

Table 1. Examination of Overdispersion

	Value	Df	Value/df
Deviance	13959.99	33	423.03

It can be seen in **Table 1** that the deviation value from the Poisson regression model is 13959.99, If this value is divided by the degree of freedom value, namely 33, we get 423.03. The result is that the deviance value is greater than 1. So, it can be concluded that there is overdispersion in the data. To overcome this, the generalized Poisson regression model is used.

3.2 Poisson Regression Model

After checking the overdispersion between the predictor variables, the results showed that the data was overdispersed, so the parameters of the Poisson Regression model were estimated for these three variables. By using the Maximum Likelihood Estimation method, parameter estimators for each variable and their constants are obtained, as shown in **Table 2** below.

Table 2. Poisson Regression Parameter Values

Variable	Estimate	Standard error (SE)	P-value
intercept	4.763978	0.502201	2e-16
X1	4.763978	0.000041	2.41e-13
X2	0.003571	0.008646	0.68
X3	0.018194	0.003933	3.73e-06

From Table 2 above, significant variables are obtained, namely altitude and humidity, which influence rainfall in Malang district. The results above do not include spatial elements, so it needs to be investigated further by adding spatial elements.

3.3 Spatial Effect Testing

In the overdispersion test, it was found that there were cases of overdispersion, so it was continued with analysis using the Geographically Weighted Poisson Regression method. The spatial dependency test via Moran's I test aims to see the spatial effect on each variable by looking at the p-value and comparing the value with α; if the p-value < α, then there is a spatial effect on that variable. The Moran's I test value can be seen in Table 3.

Table 3. Spatial dependency test

Variable	Moran's I	P-value
Y	1.425	0.0045
X1	0.985	0.00072
X2	1.4562	0.0011
X3	1.1142	0.003211

From Table 3 above, it can be seen that all variables have a p-value < α, namely the variables rainfall, altitude, temperature, and humidity. So, the four variables above have a spastic effect. And it can be continued with the GWPR model.

3.4 Geographically Weighted Poisson Regression Model

To create a GWPR model, the steps are to select the optimum bandwidth (*b*) for all locations in the Malang district. Where the selection of optimum bandwidth uses CV criteria. Furthermore, the resulting optimum bandwidth can be used to find the weighting matrix in each sub-district in Malang Regency, where this research uses fixed bisquare kernel weighting. From calculations in GWR4 software, the optimum bandwidth using a fixed bisquare kernel is 2,669. The weight used for each observation location is a fixed bisquare kernel function. The fixed kernel method allows the optimal bandwidth value for each location to be the same or constant. If the data points are distributed regularly in the research area, then using the fixed method will be suitable for modeling.

The parameter estimation results for each sub-district are presented in the table below. For example, the model form for the Ampelgading sub-district is:

$$\mu_{Ampelgading} = \exp\,(4.77473 + 0.00004X1 + 0.003504X2 + 0.018067X3)$$

Of the three variables examined, the rainfall variable in Ampelgading sub-district shows a positive relationship; this shows that if the three variables increase, the rainfall increases by one unit. The same applies to other locations in Malang district. Parameters for each location can be presented in Table 4 below:

Table 4. Parameters for Each GWPR Model Location

Subdistrict	β_0	β_1	β_2	β_3
Ampelgading	4.77473	0.000303	0.003504	0.018067
Kasembon	4.787973	0.0003	0.003052	0.018064

Subdistrict	β_0	β_1	β_2	β_3
Tirtoyudo	4.766714	0.000299	0.003522	0.018191
Kepanjen	4.754225	0.0003	0.003577	0.018329
Dampit	4.76853	0.000303	0.003639	0.018106
Lawang	4.744524	0.000305	0.004063	0.018257
Wonosari	4.787618	0.000302	0.003252	0.017991
Kalipare	4.730048	0.000306	0.00431	0.018357
Pakis	4.785186	0.000297	0.002982	0.018139
Pagelaran	4.737717	0.000306	0.004198	0.018297
Wajak	4.782007	0.000296	0.003112	0.018144
Tajinan	4.759631	0.000304	0.003741	0.01818
Karangploso	4.761787	0.000298	0.003549	0.018252
Turen	4.777085	0.000298	0.003275	0.018147
Pakisaji	4.723737	0.000307	0.004432	0.018395
Wagir	4.78465	0.000299	0.002998	0.018131
Pagak	4.776633	0.000303	0.003485	0.01805
Bululawang	4.788322	0.0003	0.003056	0.018058
Ngantang	4.77803	0.000302	0.003439	0.018048
Pujon	4.762566	0.000304	0.003729	0.018144
Sumberpucung	4.767058	0.0003	0.003379	0.018225
Dau	4.783326	0.000302	0.003324	0.018022
Sumbermanjing	4.763087	0.000297	0.003507	0.018253
Tumpang	4.751623	0.000305	0.003915	0.01822
Bantur	4.784941	0.000298	0.003073	0.018111
Gondanglegi	4.724322	0.000306	0.004367	0.018414
Singosari	4.77925	0.000302	0.003419	0.01804
Gedangan	4.774757	0.000303	0.003505	0.018067
Donomulyo	4.770105	0.000303	0.003617	0.018086
Jabung	4.78697	0.000301	0.003255	0.017999
Kromengan	4.746572	0.000304	0.004019	0.018254
Poncokusumo	4.737279	0.000305	0.004186	0.018312

The results of the GWPR parameter estimates for each location show that they are different. The results of the model above obtained the predicted values which are presented in **Figure 1** below.

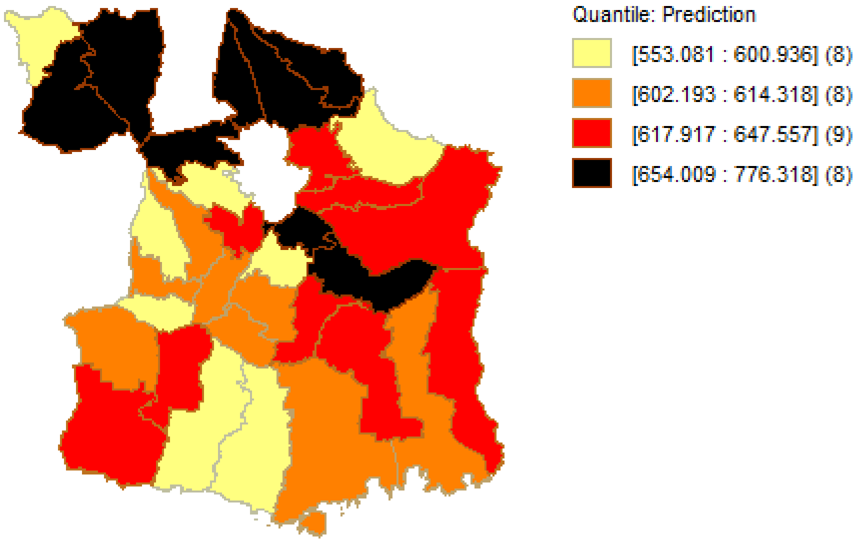

Figure 1. Rainfall Prediction Map in Malang Regency

Based on **Figure 1** above shows the results of rainfall predictions in Malang district, with rainfall between 553 and 600 mm/month in 8 locations, 602 and 614 mm/month in 8 locations, 617 and 648 mm/month in 9 locations, and 654 and 776 mm/month in 8 locations.

3.5 Selection of the Best Model

A comparison of the generalized linear Poisson model, the generalized Poisson regression model, and the GWPR model was carried out to find out which model is better to use in modeling rainfall in Malang Regency. The criteria for selecting the best model used are the AIC (Akaike's Information Criterion) values of the three models. The AIC values are compared, and the best model is selected, namely the model with the smallest AIC value. The AIC values are presented in **Table 5** below:

Table 5. Best Model Testing

Model	AIC Value
Generalized Poisson Regression	337.760
GWPR	336.845

Based on the AIC value in **Table 5** above, it can be shown that the smallest AIC value is in the GWPR model. This means that the GWPR model is more suitable for analyzing rainfall data in Malang Regency.

4. CONCLUSIONS

The results of Geographically Weighted Poisson Regression Modeling with Fixed Kernel Bisquare Weighting show that the AIC value is smaller than the Gen model. These results show that the GWPR model is better than generalized Poisson regression.

ACKNOWLEDGMENT

We would like to thank Brawijaya University for its support for this research. This research is a doctoral grant at the faculty level at Brawijaya University.

REFERENCES

[1] S. Laimeheriwa, E. L. Madubun, and E. D. Rarsina, "Analisis Tren Perubahan Curah Hujan dan Pemetaan Klasifikasi Iklim Schmidt-Ferguson untuk Penentuan Kesesuaian Iklim Tanaman Pala (Myristica fragrans) di Pulau Seram," *Agrologia*, vol. 8, no. 2, pp. 71–81, 2020.

[2] E. Sofia and M. Amalia, "Analisis Karakteristik Curah Hujan di Kota Banjarbaru Berdasarkan Data Stasiun Klimatologi Banjarbaru," *J. Teknol. Berkelanjutan*, vol. 10, no. 01, pp. 36–41, 2021.

[3] F. Dwirani, "Menentukan stasiun hujan dan curah hujan dengan metode polygon thiessen daerah kabupaten lebak," *J. Lingkung. Dan Sumberd. Alam*, vol. 2, no. 2, pp. 139–146, 2019.

[4] C. Xu, Y. Wang, W. Ding, and P. Liu, "Modeling the spatial effects of land-use patterns on traffic safety using geographically weighted Poisson regression," *Networks Spat. Econ.*, vol. 20, pp. 1015–1028, 2020.

[5] S. Ji, Y. Wang, and Y. Wang, "Geographically weighted poisson regression under linear model of coregionalization assistance: Application to a bicycle crash study," *Accid. Anal. Prev.*, vol. 159, p. 106230, 2021.

[6] R. L. Wilby, R. J. Abrahart, and C. W. Dawson, "Detection of conceptual model rainfall—runoff processes inside an artificial neural network," *Hydrol. Sci. J.*, vol. 48, no. 2, pp. 163–181, 2003.

[7] M. Kumari, C. K. Singh, O. Bakimchandra, and A. Basistha, "Geographically weighted regression based quantification of rainfall–topography relationship and rainfall gradient in Central Himalayas," *Int. J. Climatol.*, vol. 37, no. 3, pp. 1299–1309, 2017.

[8] M. A. Gebremedhin, M. W. Lubczynski, B. H. P. Maathuis, and D. Teka, "Novel approach to integrate daily satellite rainfall with in-situ rainfall, Upper Tekeze Basin, Ethiopia," *Atmos. Res.*, vol. 248, p. 105135, 2021.

[9] M. Sachdeva, A. S. Fotheringham, Z. Li, and H. Yu, "On the local modeling of count data: multiscale geographically weighted Poisson regression," *Int. J. Geogr. Inf. Sci.*, pp. 1–24, 2023.

[10] N. Peleg, F. Marra, S. Fatichi, A. Paschalis, P. Molnar, and P. Burlando, "Spatial variability of extreme rainfall at radar subpixel scale," *J. Hydrol.*, vol. 556, pp. 922–933, 2018.

[11] A. Mahama, J. A. Awuni, F. N. Mabe, and S. B. Azumah, "Modelling adoption intensity of improved soybean production technologies in Ghana-a Generalized Poisson approach," *Heliyon*, vol. 6, no. 3, 2020.

[12] G. Gao, H. Wang, and M. V Wüthrich, "Boosting Poisson regression models with telematics car driving data," *Mach. Learn.*, pp. 1–30, 2022.

[13] Y. Asar and A. Genç, "A new two-parameter estimator for the Poisson regression model," *Iran. J. Sci. Technol. Trans. A Sci.*, vol. 42, pp. 793–803, 2018.

[14] D. Murakami, N. Tsutsumida, T. Yoshida, T. Nakaya, B. Lu, and P. Harris, "Stable geographically weighted poisson regression for count data," 2021.

[15] D. N. Sari and Q. Aini, "Geographically weighted bivariate zero inflated generalized Poisson regression model and its application," *Heliyon*, vol. 7, no. 7, 2021.

[16] D. R. S. Saputro, A. Susanti, and N. B. I. Pratiwi, "The handling of overdispersion on Poisson regression model with the generalized Poisson regression model," in *AIP Conference Proceedings*, 2021, vol. 2326, no. 1.

[17] P. G. Hartono, G. M. Tinungki, J. Jakaria, A. B. Hartono, P. G. Hartono, and R. Wijaya, "Overcoming overdispersion on direct mathematics learning model using the quasi poisson regression," in *1st International Conference on Mathematics and Mathematics Education (ICMMEd 2020)*, 2021, pp. 442–449.

[18] E. H. Payne, M. Gebregziabher, J. W. Hardin, V. Ramakrishnan, and L. E. Egede, "An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data," *Commun. Stat. Comput.*, vol. 47, no. 6, pp. 1722–1738, 2018.

[19] W. Ngabu, H. Pramoedyo, R. Fitriani, and A. B. Astuti, "Spatial Modeling of Fixed Effect and Random Effect with Fast Double Bootstrap Approach," *ComTech Comput. Math. Eng. Appl.*, vol. 14, no. 1, pp. 1–9, 2023.

[20] W. Ngabu, R. Fitriani, H. Pramoedyo, and A. B. Astuti, "CLUSTER FAST DOUBLE BOOTSTRAP APPROACH WITH RANDOM EFFECT SPATIAL MODELING," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 945–954, 2023.

[21] D. R. S. Saputro, P. Widyaningsih, N. A. Kurdi, and A. Susanti, "Proporsionalitas Autokorelasi Spasial dengan Indeks Global (Indeks Moran) dan Indeks Lokal (Local Indicator of Spatial Association (LISA))," 2018.

[22] S. Chakraborty and X. Zhang, "A new framework for distance and kernel-based metrics in high dimensions," *Electron. J. Stat.*, vol. 15, no. 2, pp. 5455–5522, 2021.

[23] M. Hoffmann and F. Noé, "Generating valid Euclidean distance matrices," *arXiv Prepr. arXiv1910.03131*, 2019.

[24] A. Iriany, W. Ngabu, D. Arianto, and A. Putra, "CLASSIFICATION OF STUNTING USING GEOGRAPHICALLY WEIGHTED REGRESSION-KRIGING CASE STUDY: STUNTING IN EAST JAVA," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 1, pp. 495–504, 2023.

[25] J. G. Liao, J. E. Cavanaugh, and T. L. McMurry, "Extending AIC to best subset regression," *Comput. Stat.*, vol. 33, pp. 787–806, 2018.

[26] H. Pham, "A new criterion for model selection," *Mathematics*, vol. 7, no. 12, p. 1215, 2019.