J Statistika Vol. 18, No. 1, (2025), Hal. 840 - 850 jurnal.unipasby.ac.id

---

Implementation of Geographically and Temporally Weighted Regression with Cross Validation and Generalized Cross Validation Methods for Deforestation Modeling in Kalimantan

Gema Khusnul Ma'rifah(1), Mohammad Idhom (2), Trimono (3)

1,2,3Program Studi Sains Data, Fakultas Ilmu Komputer,

Universitas Pembangunan Nasional Veteran Jawa Timur

Jl. Rungkut Madya, Surabaya, Jawa Timur
e-mail: 21083010034@student.upnjatim.ac.id (1), idhom@upnjatim.ac.id(2),
trimono.stat@upnjatim.ac.id 3)

ABSTRAK

Deforestasi di Indonesia mendapatkan sorotan di tingkat nasional hingga internasional yang menimbulkan dampak ekologi, ekonomi, dan sosial. Kalimantan termasuk wilayah dengan tingkat deforestasi yang tinggi dipicu oleh berbagai faktor yang bervariasi antar lokasi serta waktu. Penelitian ini bertujuan untuk memodelkan laju deforestasi di Kalimantan selama periode 2014 hingga 2022 menggunakan metode Geographically and Temporally Weighted Regression (GTWR). Model diuji dengan kombinasi fungsi pembobot Fixed dan Adaptive Gaussian Kernel serta metode penentuan bandwidth Cross Validation (CV) dan Generalized Cross Validation (GCV). Hasil menunjukkan bahwa model terbaik adalah GTWR dengan Fixed Gaussian Kernel dan GCV yang didasarkan pada nilai R^2 dan AIC. Analisis spasial-temporal menunjukkan bahwa tidak ada variabel signifikan pada 2014, kebakaran hutan signifikan pada 2019, dan pada tahun-tahun lainnya kedua variabel tersebut signifikan secara luas. Temuan ini memberikan wawasan terhadap dinamika spasial dan temporal faktor-faktor deforestasi untuk mendukung kebijakan berbasis wilayah.

Kata kunci: GTWR; Cross Validation; Generalized Cross Validation

ABSTRACT

Deforestation in Indonesia has received national and international attention for its ecological, economic and social impacts. Kalimantan is an area with high deforestation rates triggered by various factors that vary between locations and time. This study aims to model the deforestation rate in Kalimantan during the period 2014 to 2022 using the Geographically and Temporally Weighted Regression (GTWR) method. The model was tested with a combination of Fixed and Adaptive Gaussian Kernel weighting functions and Cross Validation (CV) and Generalized Cross Validation (GCV) bandwidth determination methods. The results show that the best model is GTWR with Fixed Gaussian Kernel and GCV based on R^2 *and AIC values. Spatial-temporal analysis shows that neither variable is significant in 2014, forest fires are significant in 2019, and in other years both variables are broadly significant. The findings provide insights into the spatial and temporal dynamics of deforestation factors to support area-based policies.*

Keywords: *GTWR; Cross Validation; Generalized Cross Validation*

INTRODUCTION

The rate of deforestation in Indonesia is one of the polemics that has received attention at the national and international levels. Indonesia ranks fifth in global tree cover loss from 2001 to 2023 and second in humid tropical primary forest loss from 2002 to 2023 [1]. Indonesia experienced the highest deforestation rate globally between 1996 and 2000 at 3.5 million hectares and decreased from 2002 to 2014 by 0.75 million hectares per year with the lowest point in 2022 at 104,000 hectares [2]. This decrease in deforestation rate was successfully suppressed by the implementation of the New Forest Moratorium and Reforestation Program. However, this decline has not succeeded in overcoming the various impacts caused both in terms of ecology, economy, and social such as changes in air temperature, loss of biodiversity, increased flooding, soil erosion, and increased emissions resulting in global warming.

Kalimantan recorded a deforestation rate of 1.12% annually from 1990 to 2014 [3]. The five provinces in Kalimantan are included in the 10 provinces with the worst deforestation in 2023 because almost half of the total deforestation in Indonesia occurred in Kalimantan with an area of 124,611 hectares out of 230,760 hectares [4]. Meanwhile, East Kalimantan is predicted to experience deforestation of up to 2,749.16 hectares in the next 10 years as a result of the move of Indonesia's new capital city [5]. This states that Kalimantan is one of the regions in Indonesia that continues to experience deforestation with the possibility of increasing deforestation rates caused by the relocation of the national capital.

Mining and logging activities as one of the significant economic activities have a major influence on ecological damage and high deforestation in Kalimantan along with the influence of population density [6]. The relocation of the national capital also triggers the potential for deforestation as a result of migration by changing the function of forest land into residential, industrial and office land [7]. The conversion of primary forests into oil palm plantations to fulfill market needs is also one of the causes of deforestation in Kalimantan [8]. Another cause of deforestation in Kalimantan that has a wide impact is forest and land fires [9].

Causal factors of deforestation exhibit both spatial and temporal variations, requiring in-depth analysis that considers both. Geographically Weighted Regression (GWR) and Multiscale Geographically Weighted Regression (MGWR) methods are able to capture spatial variations, but do not account for temporal changes. To overcome this limitation, the Geographically and Temporally Weighted Regression (GTWR) method is used because it is able to capture both spatial and temporal variations. Some previous studies have shown that GTWR is superior to global regression and GWR such as in modeling protected forest deforestation in Indonesia and in the analysis of the construction sector in Java Island using bisquare weight functions [10] [11]. Another study used GTWR Robust to overcome the violation of the normality assumption in the analysis of deforestation in Sumatra [12]. Research also showed that using Gaussian and Bisquare kernels gave almost similar results, while the Great Circle Distance approach with an exponential kernel still produced the best model in the analysis of the Human Development Index in West Java [13] [14].

Based on this, research on the GTWR method has shown development in various fields by discussing different aspects, especially related to the use of weight functions. However, there is no research that explicitly discusses the selection of methods to determine the optimum bandwidth of the weight function. The choice of method between Cross Validation (CV) and Generalized Cross Validation (GCV) in determining the optimum bandwidth affects the resulting estimate. Therefore, this study uses Geographically and Temporally Weighted Regression with CV and GCV against two different weighting functions namely Fixed Gaussian Kernel and Adaptive Gaussian Kernel in the context of deforestation in Kalimantan. This research is expected to provide a deeper understanding of the causal factors and patterns of deforestation that varies between regions in Kalimantan each year.

METHOD

Data Collection

The research used several variables including deforestation factors from 2014 to 2022. All data used in this study is sourced from the Badan Pusat Statistika (BPS). The variables used in this study consist of Deforestation (Y), Population Density (X_1), Area of Forest and Land Fires (X_2), Production of Oil Palm Plantation Crops (X_3), Area of Oil Palm Plantation Crops (X_4), GDP Mining and Quarrying (X_5), and GDP Subcategory of Forestry and Logging (X_6). The total amount of data used in this study is 45 rows of data consisting of 5 provinces in Kalimantan, namely West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, and North Kalimantan.

Preprocessing Data

Data preprocessing is a series of stages carried out in preparing the quality of data before further analysis is carried out. The preprocessing carried out consists of Cleaning data; Data Integration; Handling Missing Values is done using the K-Nearest Neighbor Imputation (KNNI) [15] karena KNNI mampu mempertahankan struktur lokal dan pola kompleks dalam data spasial-temporal. KNNI dipilih daripada imputasi rata-rata yang cenderung terlalu menyederhanakan atau metode Multiple Imputation by Chained Equations (MICE) yang membutuhkan asumsi distribusi dan iterasi kompleks yang tidak diperlukan dalam konteks ini; and Handling Reforestation by changing the negative deforestation value, which means that reforestation is greater than deforestation, so to avoid misunderstanding the analysis, the value is changed to 0.

Linear Regression

A multiple linear regression model is used to explain the relationship between the dependent variable and several independent variables [16]. Simultaneous tests are conducted with the F-test [17], while partial tests use the t-test [18]. Classical assumption testing is conducted to ensure the validity of the model, including normality test using Kolmogorov-Smirnov [19], multicollinearity with VIF < 10, heteroscedasticity using Glejser test, and autocorrelation using Durbin-Watson test [20].

Spatial and Temporal Heterogeneity Testing

The Breusch-Pagan Test can be used in testing spatial heterogeneity [21]. The temporal heterogeneity test can be tested using a boxplot [12].

Geographically and Temporally Weighted Regression (GTWR)

The GTWR model is an extension of the GWR model. GTWR integrates spatial and temporal data in the weight matrix. The GTWR model is formulated in equation (1) [22].

$$Y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k (u_i, v_i, t_i) x_{ik} + \varepsilon_i \qquad (1)$$

Y_i is the dependent variable with k as independent variables at location (u_i, v_i) with time t_i. Parameter estimation $\hat{\beta}(u_i, v_i, t_i)$ with the i -th point in GTWR is shown in equation (2) [22].

$$\hat{\beta}(u_i, v_i, t_i) = \left(X^T W(u_i, v_i, t_i)\right)^{-1} X^T W(u_i, v_i, t_i) Y \tag{2}$$

Let n be the number of observations and $W(u_i, v_i, t_i) = diag(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$. The element $\alpha_{i1}(1 \le j \le n)$ represents the spatial-temporal distance function at (u_i, v_i, t_i). The following equation defines the spatial-temporal distance function that combines the spatial and temporal distance characteristics in equation (3) [22]:

$$(d^{ST})^2 = \lambda (d^S)^2 + \mu (d^T)^2 \tag{3}$$

The spatial and temporal distance functions are denoted by d^S and d^T with scale factors λ and μ that balance the different impacts of spatial and temporal distance measurements. Hence the Euclidean distance in equation (4) [22]:

$$\left(d_{ij}^{ST}\right)^2 = \lambda\left\{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right\} + \mu\left(t_i - t_j\right)^2 \tag{4}$$

With $\lambda \neq 0$ let τ be the ratio parameter of $\frac{\mu}{\lambda}$, then the equation becomes equation (5) [22]:

$$\frac{\left(d_{ij}^{ST}\right)^2}{\lambda} = \left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2 + \tau\left(t_i - t_j\right)^2 \tag{5}$$

With the effect of temporal distance, the spatial distance can be enlarged or minimized by the influence of the parameter τ. The model used for spatial weighting factors in GTWR uses the Fixed Gaussian Kernel (6) and Adaptive Gaussian Kernel (7) kernel function as follows [23]:

$$w_i(u_i, v_i) = exp\left[-\frac{1}{2}\left({d_{ij}}/{h}\right)^2\right] \tag{6}$$

$$w_i(u_i, v_i) = exp\left[-\frac{1}{2}\left({d_{ij}}/{h_{i(q)}}\right)^2\right] \tag{7}$$

The optimal bandwidth selection uses two methods namely Cross Validation (CV) and Generalized Cross Validation formulated in (8)(9) [24].

$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(b)]^2 \tag{8}$$

$$GCV = \sum_{i=1}^{n} [y_i - \hat{y}_{=i}(b)]^2 / (n - v_1)^2 \tag{9}$$

Model Evaluation

Model evaluation is performed using R^2 and Akaike Information Criterion (AIC) evaluation metrics is defined in equation (10)(11) [25].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{10}$$

$$AIC = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n + tr(G) \tag{11}$$

RESULT AND DISCUSSION

Descriptive Analysis

Table 1 contains an overview value of each variable in this research.

Table 1. Descriptive Analysis of Research Variable

	Y	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**
Min	0	8.00	82.2	167.7	50.3	6334	706

1st Qu	11564	17.00	1743.8	1343.0	479.3	14715	1425
Median	27240	28.00	7646.0	2192.6	1078.8	25572	2336
Mean	45206	39.07	47662.4	2950.2	1014.3	77116	3090
3rdQu	69585	37.00	47432.6	4100.9	1434.5	42459	4236
Std.dev	42986.9	35.2	102940	2386.6	668.4	115150	2242.7
max	186642	113.00	583833.4	8600.9	2205.8	490046	7488

Table 1 shows a very scattered distribution of the dependent variable (Y) with possible outliers and some areas with very high values. The independent variable (X_1) has a fairly wide spread indicating a positively skewed distribution. Variable (X_2) indicates extreme outliers and a highly abnormal distribution. The variable (X_3) indicates a very scattered distribution with possible outliers and shows a distribution that tends to be right (positively skewed). Variable (X_4) indicates a slightly negative skewed distribution, although not generally and is generally widely distributed. Variable (X_5) has a very wide and extreme data distribution indicating extreme dispersion and strong outlier potential. The variable (X_6) has a more moderate distribution than X_2 and X_5, but has a wide spread and is slightly positively skewed.

Linear Regression Modeling

The first step is to estimate the parameters using Ordinary Least Square (OLS) to explain the relationship between the independent variables and the dependent variable. The results of multiple linear regression model parameter estimation can be seen in Table 2.

Table 2. Parameter Estimation Value

Variable	Coefficient	Standard Error	t-Value	p-Value
Intercept	106200	32440	3.274	0.00226
X_1	-849.40	289.90	-2.930	0.00570
X_2	0.164	0.0529	3.100	0.00364
X_3	-9.434	6.107	-1.545	0.13069
X_4	6.743	19.63	0.344	0.73310
X_5	0.1553	0.1385	1.122	0.26906
X_6	-8.649	8.502	-1.017	0.31546

Furthermore, the simultaneous test shows that the F test statistical value is 4.8 and the p-value is 0.000985. The decision to reject H_0 means that simultaneously the independent variable has a significant influence on the dependent variable. In the partial test, the results show that variables X_1, and X_2 have a significant influence on the dependent variable with a p-value < 0.05. The variables X_3, X_4, X_5, X_6 are not significant because the p-value > 0.05 so they do not make a significant contribution to the dependent variable.

The first linear regression residual assumption test is the normality test with the test results showing a D value of 0.14315 with a p-value of 0.2869 which has a value greater than the 0.05 significance level. So it can be concluded that the data is normally distributed. The multicollinearity test results show that the variable that has a VIF value above 10 is variable X_6. In the heteroscedasticity test, the Glejser test shows a p-value < 0.05 on variable X_1 with a value of

0.02150 while X_2, X_3, X_4, X_5, and X_6 do not show significant heteroscedasticity in the model. The autocorrelation test results show a DW value of 1.8286 with a p-value of 0.1162. Because the p-value > 0.05, it is concluded that there is significant autocorrelation in the residuals of the regression model.

Based on the previous parameter estimation test results, variables X_3, X_4, X_5, and X_6 were excluded from the model due to their low statistical significance (p-value > 0.05), indicating a minimal contribution to the model. Additionally, variable X_6 showed signs of multicollinearity, which could affect the stability of coefficient estimates. Removing non-significant variables helps simplify the model without significantly reducing its predictive power. Since several assumption tests were not met, re-modeling was necessary, and new parameter estimates are presented in Table 3.

Table 3. Parameter Estimation Value

Variable	Coefficient	Standard Error	t-Value	p-Value
Intercept	57960	8589	6.749	3.33e-08
X_1	-513.50	156.70	-3.278	0.00211
X_2	0.1515	0.05346	2.835	0.00703

In the simultaneous test, it was found that the F test statistical value was 9.361 and the p-value was 0.0004345. The decision rejects H_0, which means that simultaneously the independent variable has a significant effect on the dependent variable. In the partial test, the results show that the variables X_1 and X_2 have a significant effect on the dependent variable with a p-value of 0.00211 and 0.00703. The normality test shows a D value of 0.10947 with a p-value of 0.6143 which has a value greater than the 0.05 significance level. The multicollinearity test results show that the variables have a VIF value below 10. The Glejser test shows a p-value < 0.05 on variable X_1 with a value of 0.0000551, while X_2 does not show significant heteroscedasticity in the model. The autocorrelation test results show a DW value of 0.98756 with a p-value of 0.0001141. Because the p-value < 0.05, it is concluded that there is significant autocorrelation in the residuals of the regression model.

Spatial and Temporal Heterogeneity Testing

Testing spatial heterogeneity using Breusch-Pagan results in a BP value of 8.6302 with a p-value of 0.01337. Because the p-value < 0.05, there is spatial heterogeneity. The temporal heterogeneity test is shown in Figure 1.

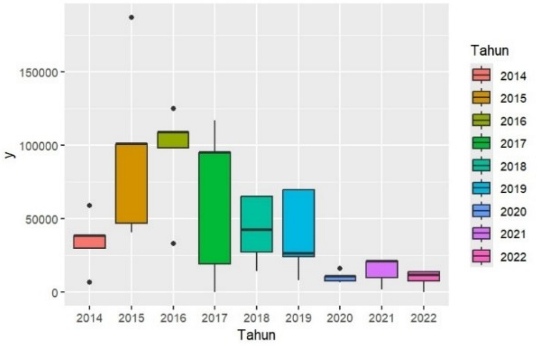

Figure 1. Boxplot of Deforestation in Kalimantan 2014-2022

Figure 1 shows the diversity of the dependent variable (Y) from year to year which illustrates significant fluctuations. Significant differences in the data distribution of some years indicate inhomogeneity in the Deforestation variable which needs to be further analyzed especially the factors that may affect deforestation during those years.

Geographically and Temporally Weighted Regression (GTWR)

Table 4 shows the parameter estimation summary of the Geographically Temporally Weighted Regression model with four kernel configurations: Gaussian Fixed, Gaussian Fixed GCV, Gaussian Adaptive, and Gaussian Adaptive GCV. Coefficient estimates for each variable are presented based on the minimum, maximum, mean, and standard deviation values.

Table 4. Summary of Coefficient Estimates

Models	Variable	Min	Max	Mean	Std. Deviation
Gaussian Fixed	Intercept	52889.34	91798.33	72528.19	13352.6
	X_1	-768.9184	-460.2715	-617.0524	96.53012
	X_2	-0.041245	0.2173	0.1303034	0.069369
Gaussian Fixed GCV	Intercept	55884.45	93006.57	73014.79	13376.2
	X_1	-780.0198	-499.2047	-623.4417	92.64175
	X_2	-0.12756	0.218129	0.1203035	0.094433
Gaussian Adaptive	Intercept	54247.05	65553.6	61295.97	3570.865
	X_1	-574.2845	-478.0567	-539.3515	28.28268
	X_2	0.13508	0.1748805	0.1536025	0.01171776
Gaussian Adaptive GCV	Intercept	57248.89	59723.09	58788.04	823.7398
	X_1	-527.8837	-507.1715	-520.1057	6.571252
	X_2	0.1474668	0.1576054	0.1521811	0.003044776

The GTWR model was compared with the linear regression model using the F test to see whether GTWR provides a significant improvement in explaining the data. Based on the tests conducted, the GTWR model with fixed bandwidth both CV and GCV are superior to the linear regression model because it provides a significant F value (p-value <0.05). In contrast, the GTWR model with adaptive bandwidth did not show a significant advantage over the linear regression model due to the high p-value. Furthermore, to see the spatial and temporal variations produced by the model, a partial test was conducted, the results of which are shown in Figure 2.

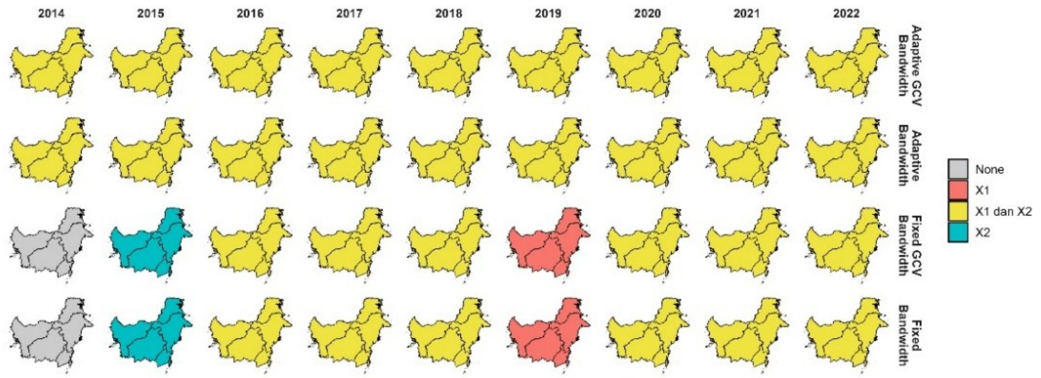

Figure 2. Distribution map based on significant variables

The GTWR model shows that the effect of population density and forest and land fires on deforestation is spatially and temporally variable depending on the year and location. In 2014, the Gaussian Fixed and Gaussian Fixed GCV models did not identify statistically significant predictors. The variable X2 (forest and land fires) started to show a significant effect in 2015. In 2016, both X1 (population density) and X2 became significant in more locations. From 2017-2022, all four models consistently showed that both X1 and X2 variables were statistically significant in most provinces and years, indicating that population density and fires were the main drivers of deforestation. In 2019, the Gaussian Fixed and Gaussian Fixed GCV models only show significance in X1, while the Adaptive model shows that both predictors are significant in all provinces in Kalimantan.

Population density tends to be the dominant predictor in years or areas with development activity, settlement expansion, or urbanization that could potentially increase pressure on forest areas. Forest and land fires are more significant in years with extreme weather conditions such as El Nino or when there is an increase in hotspots. This also reflects uneven fire control policies across provinces. The increased influence of X1 after 2019 can also be attributed to the planned relocation of the capital city to East Kalimantan which triggers population growth and new land activities in the vicinity.

Model Evaluation

Based on the evaluation model based on the R^2 and AIC values of each model, it can be concluded that the Gaussian Fixed GCV model is the model that performs the best modeling of the data. The value of the evaluation model can be seen in Table 5.

Table 5. Model Evaluation

Models	R^2	AIC
Gaussian Fixed	0.4503620	1067.479
Gaussian Fixed GCV	0.4599574	1067.243
Gaussian Adaptive	0.3564924	1070.633
Gaussian Adaptive GCV	0.3203041	1072.631

CONCLUSION

Geographically and Temporally Weighted Regression (GTWR) modeling in the case of deforestation in Kalimantan during the period 2014-2022 was carried out with two types of spatial weighting functions namely Gaussian Fixed and Gaussian Adaptive, each with two bandwidth selection methods Cross Validation (CV) and Generalized Cross Validation (GCV). The evaluation results show that the Gaussian Fixed Kernel model with GCV bandwidth provides the best performance with an R^2 value of 0.4599574 and AIC of 1067.243. This study is limited to the use of Gaussian Fixed and Adaptive kernels as well as CV and GCV methods in bandwidth selection. Future research can develop models with other kernel functions such as Bisquare and consider adding relevant variables to improve model accuracy.

REFERENCE

[1] M. Weisse and E. Goldman, “Top 10 Lists,” World Resources Institute. Accessed: Aug. 20, 2024. [Online]. Available: https://research.wri.org/id/node/77

[2] KLHK, “Pengendalian Deforestasi dan Karhutla di Indonesia,” PPID Kementerian Lingkungan Hidup dan Kehutanan Pejabat Pengelola Informasi dan Dokumentasi. Accessed: Aug. 20, 2024. [Online]. Available: https://ppid.menlhk.go.id/berita/siaran-pers/7594/pengendalian-deforestasi-dan-karhutla-di-indonesia

[3] Yusuf Aguswan, “POLA DEGRADASI DAN DEFORESTASI DI KESATUAN HIDROLOGIS GAMBUT (KHG) PROVINSI KALIMANTAN TENGAH TAHUN 2016 -2017,” *Jurnal Hutan Tropika (Tropical Forest Journal)*, vol. XIV, no. 2, pp. 89–98, Dec. 2019, doi: 10.36873/jht.v14i2.1151.

[4] D. L. Pristiandaru, “10 Provinsi dengan Deforestasi Terparah 2023, Mayoritas di Kalimantan,” Kompas.com. Accessed: Aug. 20, 2024. [Online]. Available: https://lestari.kompas.com/read/2024/03/28/090000586/10-provinsi-dengan-deforestasi-terparah-2023-mayoritas-di-kalimantan

[5] A. H. Nawawi and Evangs Mailoa, “Prediksi Lahan Deforestasi Dan Reforestasi Hutan Kalimantan Timur Dengan Metode Rantai Markov,” *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 1, pp. 251–259, Feb. 2024, doi: 10.51454/decode.v4i1.268.

[6] D. N. Isnaini *et al.*, “Determinants of Deforestation in Kalimantan,” in *Seminar Nasional Official Statistics*, 2021, pp. 978–988. doi: 10.34123/semnasoffstat.v2020i1.570.

[7] T. Berlianty and T. Meiliana, “Potensi Deforestasi di Pulau Kalimantan: Pro dan Kontra Migrasi,” *International Journal of Demos (IJD)*, vol. 5, no. 2, pp. 279–290, Jun. 2023, doi: 10.37950/ijd.v5i2.426.

[8] N. I. Fawzi and M. Y. Iswari, “Analisis Heat Island pada Perkebunan Kelapa Sawit: Studi Kasus di Kabupaten Kayong Utara, Kalimantan Barat,” *Jurnal Wilayah dan Lingkungan*, vol. 8, no. 2, pp. 106–115, Aug. 2020, doi: 10.14710/jwl.8.2.106-115.

[9] F. Ulandari and R. Kurniawan, “PERBANDINGAN ALGORITMA LSDBC DAN DBSCAN PADA PEMETAAN DAERAH RAWAN KEBAKARAN HUTAN: Studi Kasus di Pulau Sumatera, Kalimantan, Sulawesi, dan Papua,” *Jurnal Aplikasi Statistika & Komputasi Statistik*, vol. 12, no. 2, pp. 25–30, Dec. 2020, doi: 10.34123/jurnalasks.v12i2.281.

[10] T. Z. Adiningrum, A. Prahutama, R. Santoso, D. Statistika, F. Sains, and D. Matematika, “PEMODELAN DEFORESTASI HUTAN LINDUNG DI INDONESIA MENGGUNAKAN MODEL GEOGRAPHICALLY AND TEMPORALLY WEIGHTED REGRESSION (GTWR),” *Jurnal Gaussian*, vol. 7, no. 3, pp. 314–325, Aug. 2018, doi: 10.14710/j.gauss.7.3.314-325.

[11] S. Haryanto, M. N. Aidi, and A. Djuraidah, “Analysis of the Geographically and Temporally Weighted Regression (GTWR) of the GRDP the Construction Sector in Java Island,” *Forum Geografi*, vol. 22, no. 1, pp. 130–139, 2019, doi: 10.23917/forgeo.v33i1.7332.

[12] A. D. Putra and S. I. Oktora, “Spatial-Temporal Analysis of Deforestation in Sumatera Island 2011-2019,” in *Proceedings of The International Conference on Data Science and Official Statistics*, 2022, pp. 590–609. doi: 10.34123/icdsos.v2021i1.202.

[13] Harianto, W. H. Nugroho, and E. Sumarminingsih, “Geographically and Temporally Weighted Regression Model with Gaussian Kernel Weighted Function and Bisquare Kernel Weighted Function,” *IOP Conf Ser Mater Sci Eng*, vol. 1115, no. 1, p. 1, Mar. 2021, doi: 10.1088/1757-899x/1115/1/012063.

[14] A. D. Rahmawati, A. H. Wigena, and M. N. Aidi, “Influencing factors for the human development index in West Java using geographically and temporally weighted regression kernel functions,” *Jurnal Pendidikan Geografi: Kajian, Teori, dan Praktek dalam Bidang Pendidikan dan Ilmu Geografi*, vol. 28, no. 2, p. 228, Jun. 2023, doi: 10.17977/um017v28i22023p228-241.

[15] Miraati Laia, “Analisis Kinerja Algoritma K-Nearest Neighbor Imputation (KNNI) Untuk Missing Value Pada Klasifikasi Data Mining,” *Journal of Informatics, Electrical and Electronics Engineering*, vol. 2, no. 3, pp. 92–98, Mar. 2023, doi: 10.47065/jieee.v2i3.891.

[16] T. Maulana Fahrudin, P. Aji Riyantoko, K. Maulida Hindrayani, and I. Gede Susrama Mas Diyasa, “Daily Forecasting for Antam’s Certified Gold Bullion Prices in 2018-2020 using Polynomial Regression and Double Exponential Smoothing,” *Journal of International Conference Proceedings*, vol. 3, no. 4, pp. 45–53, 2020, doi: 10.32535/jicp.v3i4.1009.

[17] T. Arifianto, Y. A. Pangestu, D. S. Oktaria, L. S. Moonlight, and D. I. Pratiwi, “Prediksi Daya Pada Panel Surya Menggunakan Metode Time Series dan Analisis Regresi,” *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, vol. 4, no. 01, pp. 52–63, May 2022, doi: 10.46772/intech.v4i01.674.

[18] Y. Widyaningsih, G. P. Arum, and K. Prawira, “APLIKASI K-FOLD CROSS VALIDATION DALAM PENENTUAN MODEL REGRESI BINOMIAL NEGATIF TERBAIK,” *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 15, no. 2, pp. 315–322, Jun. 2021, doi: 10.30598/barekengvol15iss2pp315-322.

[19] T. Trimono, D. I. Asih Maruddani, and D. Ispriyanti, “PEMODELAN HARGA SAHAM DENGAN GEOMETRIC BROWNIAN MOTION DAN VALUE AT RISK PT CIPUTRA DEVELOPMENT Tbk,” *JURNAL GAUSSIAN*, vol. 6, no. 2, pp. 261–270, 2017, [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/gaussian

[20] G. Mardiatmoko, “PENTINGNYA UJI ASUMSI KLASIK PADA ANALISIS REGRESI LINIER BERGANDA (STUDI KASUS PENYUSUNAN PERSAMAAN ALLOMETRIK KENARI MUDA [CANARIUM INDICUM L.]),” *BAREKENG: Jurnal Ilmu Matematika*

dan Terapan, vol. 14, no. 3, pp. 333–342, Oct. 2020, doi: 10.30598/barekengvol14iss3pp333-342.

[21] Deviani, B. Y. Wulandari, Askariyah, I. J. Fitri, and S. Hariati, "ANALISIS SPASIAL BERBASIS PEMETAAN MENGGUNAKAN GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) PADA SEBARAN JUMLAH WISMA DI KABUPATEN LOMBOK BARAT," *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, vol. 5, no. 2, pp. 660–670, Aug. 2024, doi: 10.46306/lb.v5i2.545.

[22] B. Huang, B. Wu, and M. Barry, "Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices," *International Journal of Geographical Information Science*, vol. 24, no. 3, pp. 383–401, Mar. 2010, doi: 10.1080/13658810802672469.

[23] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression the analysis of spatially varying relationships*. England: John Wiley & Sons Ltd, 2002.

[24] A. T. Damaliana, I. Nyoman Budiantara, and V. Ratnasari, "Comparing between mgcv and agcv methods to choose the optimal knot points in semiparametric regression with spline truncated using longitudinal data," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jul. 2019. doi: 10.1088/1757-899X/546/3/032003.

[25] A. Muhaimin, H. Prabowo, and Suhartono, "Model Selection for Forecasting Rainfall Dataset," *IJDASEA (International Journal of Data Science, Engineering, and Analytics)*, vol. 1, pp. 1–10, 2021, doi: 10.3390/xxxxx.