Original Article

Check for updates

Available online at
www.heca-analitika.com/ijds

Infolitika Journal of Data Science

Vol. 1, No. 2, 2023

Unraveling Geospatial Determinants: Robust Geographically Weighted Regression Analysis of Maternal Mortality in Indonesia

Latifah Rahayu [1], Elvitra Mutia Ulfa [1], Novi Reandy Sasmita [1,*], Hizir Sofyan [1], Rumaisa Kruba [1], Selvi Mardalena [1] and Arif Saputra [2]

[1] Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; latifah.rahayu@usk.ac.id (L.R.); elvitramutia@gmail.com (E.M.U.); novireandys@usk.ac.id (N.R.S.); hizir@usk.ac.id (H.S.); rumaisakruba@usk.ac.id (R.K.); selvimardalena@usk.ac.id (S.M.)

[2] Department of Epidemiology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Thailand; 6510320009@email.psu.ac.th (A.S.)

* Correspondence: novireandys@usk.ac.id

Article History

Received 31 October 2023
Revised 12 December 2023
Accepted 21 December 2023
Available Online 28 December 2023

Keywords:
Robust Geographically Weighted Regression (RGWR)
Maternal Mortality
Indonesia
Statistics

Abstract

Maternal Mortality Rate (MMR) in Indonesia has experienced a concerning annual increase, reaching 4,627 deaths in 2020 compared to 4,221 in 2019. This upward trajectory underscores the urgency of investigating the factors contributing to MMR. Recognizing the spatial heterogeneity and outliers in the data, our study employs the Robust Geographically Weighted Regression (RGWR) method with the Least Absolute Deviation approach. Using secondary data from the 2020 Indonesian Health Profile publication, the research seeks to establish province-specific models for MMR in 2020 and identify the key influencing factors in each region. Standard regression analyses fall short in addressing the complexities present in the data, making the RGWR approach crucial for understanding the nuanced relationships. The chosen RGWR model utilizes the Least Absolute Deviation method and a fixed kernel exponential weighting function. Notably, this model maintains a consistent bandwidth value across all locations, showcasing its robustness. In evaluating the model variations, the exponential fixed kernel weighting function emerges as the most optimal, boasting the smallest Akaike Information Criterion (AIC) value of 23.990 and the highest coefficient of determination R^2 value of 93.66%. The outcomes of this research yield 24 distinct models, each tailored to the unique characteristics of every province in Indonesia. This nuanced, location-specific approach is vital for developing effective interventions and policies to address the persistently high MMR. By providing insights into the complex interplay of factors influencing maternal mortality in different regions, the study contributes to the groundwork for targeted and impactful public health initiatives across Indonesia.

1. Introduction

In essence, the quality of maternal and child health is inseparable from the progress of a nation because a responsible generation emerges from the good health of mothers [1]. Family development is carried out to establish a healthy family living in a wholesome environment. One criterion for a good family is not only a healthy environment but also the health condition of each family member. By meeting nutritional needs and maintaining the health of family members, a family contributes to optimizing the growth, development, and productivity of all its members [2]. Maternal Mortality Rate (MMR) remains high due to the vulnerability of the health status of mothers and children, especially in the most susceptible groups, such as pregnant women, postpartum mothers, and newborns. Therefore, maternal health initiatives become one of the primary goals of Indonesia's health development. Since mothers are highly vulnerable to the conditions of the family and their environment in general, they deserve priority in health intervention efforts. Hence, it is crucial to evaluate the health situation and the effectiveness of maternal health initiatives [3]. Consequently, in this current study, the focus lies on utilizing population indicators obtained from the Aceh Population Registration Office in 2021, encompassing 23 districts and cities in Aceh Province.

The AKI (Maternal Mortality Ratio) serves as a measure of the developmental status of a region, particularly in the realm of health. Depending on the quality levels achievable by the government, community, and geographical location of a region, the AKI in one area of Indonesia may differ from that in another. To effectively implement policies and minimize the AKI in a specific area, it is essential to understand the factors influencing the AKI. Regression analysis is one statistical technique that can be employed to identify variables affecting the AKI. However, the regression approach does not account for differing geographical factors between locations, necessitating the use of spatial analysis [4].

As the distance between two observations increases, the characteristics of each observation will start to differ from those nearby. On the other hand, spatial non-stationarity describes varying variance values from one region to another. Spatial heterogeneity is another term for non-uniform variance [5]. Classical regression model assumptions are not met due to spatial heterogeneity. Consequently, the variance values produced vary for each observation in each location. Therefore, a regression model is created by allowing error model variance to vary for each region, particularly by introducing local regression coefficients, meaning that each area will have its own separate regression coefficients. Geographically Weighted Regression (GWR) is a process of constructing a regression model that considers local factors by incorporating spatial weighting [6].

The robust geographically weighted regression (RGWR) approach is widely employed in various research analyses. In the study conducted by [7] on modeling diarrhea occurrences in the city of Semarang, a robust geographically weighted regression with the absolute least deviation method was utilized. Due to the presence of outlier values in the diarrhea occurrence data, the traditional geographically weighted regression (GWR) model struggled to accurately capture the relationship between diarrhea issues and related variables. As a result, it can be concluded that the diarrhea cases in Semarang, when modeled using the RGWR approach, yielded more accurate parameter estimates compared to the GWR model. The percentage of households lacking proper sanitation facilities or variable X5 exhibited the highest regression coefficient when estimating diarrhea occurrences using the RGWR model [7].

Addressing outliers in data analysis is crucial as these atypical data points possess characteristics that significantly differ from the rest of the dataset. If left unaddressed, outliers can distort the overall distribution of data, affecting the accuracy of statistical analyses and parameter estimations. In the context of this research on predicting Maternal Mortality Ratio (AKI) cases in Indonesian provinces for 2020, outliers could skew the understanding of influential factors and hinder the creation of effective models. The necessity of the RGWR approach lies in its capability to effectively handle outliers. Unlike traditional methods that are sensitive to outliers and can produce biased estimations, RGWR, utilizing techniques like the Least Absolute Deviation (LAD), offers robustness against outliers' influence. By assigning less weight to outliers and focusing on the median absolute deviation, RGWR ensures a more accurate estimation of parameters, particularly in geospatial analyses where outlier presence is common. Examining spatial patterns in the distribution have been done by Southern India [8]. GWR analysis also applied to assess predictors of home birth hot spots in Ethiopia [9].

In the specific context of predicting Maternal Mortality Ratio in Indonesian provinces for 2020, the relevance of RGWR becomes evident. Maternal health data often encompasses diverse factors influenced by geographic variations. Outliers in this dataset could stem from various sources, such as extreme socio-economic conditions or exceptional health care facilities. These outliers might skew the understanding of factors influencing maternal mortality rates. Thus, employing RGWR in this scenario becomes paramount to create a more accurate and reliable predictive model by robustly addressing the challenges posed by outliers within the geographic context of Indonesia's provinces. By doing so, the RGWR method ensures a more nuanced understanding of the complex relationships between various factors contributing to maternal mortality, thereby enhancing the accuracy of predictions for effective healthcare planning and interventions.

2. Literature Review

2.1. Regression Analysis

Regression analysis is one of the most widely used statistical methods in applied sciences [10–12]. It is a method for establishing a cause-and-effect relationship between a dependent variable and one or more independent variables [13, 14]. In general, the regression equation can be observed in Equation 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_i \qquad (1)$$
$$i = 1,2,\dots,n$$

In the regression equation, where Y is the dependent variable, X_i is the independent variable, β_0 is the intercept, β_1 is the slope, and ε_i is the residual. β_0, β_1, β_2, ..., β_k, are the regression coefficients that need to be estimated. The Ordinary Least Squares (OLS) method is a commonly used estimation method for regression analysis [15].

2.2. Geographically Weighted Regression (GWR)

The Geographically Weighted Regression (GWR) model is an extension of the regression model with parameters calculated at each observation location. The GWR method is employed to estimate spatially non-stationary data or data with spatial heterogeneity. Spatial heterogeneity occurs when the same independent variable elicits different responses at different locations within a study area [16–18]. The GWR model is applied to continuous dependent variables, producing localized parameter estimates for regression models at each location.

GWR generates localized parameter estimates for each location where the data is observed [19]. In the GWR model, the dependent variable y is estimated with independent variables, each with regression coefficients that depend on the observed location. The GWR estimation method utilizes Weighted Least Squares (WLS) or commonly known as weighted least squares. GWR employs weights based on the distance between one observation location and another. The GWR model can be formulated as in Equation 2 [20].

$$y_i = \beta_0(u_i, v_i) + \sum_{i=1}^{n} \beta_k(u_i, v_i) x_{ki} + \varepsilon_i \qquad (2)$$

Where y_i is the dependent variable for location *i*-th, x_k is the *k*-th independent variable at location *i*-th, $(u_i v_i)$ is coordinates (longitude and latitude) at location *i*-th. $\beta_0(u_i, v_i)$ is the slope for the regression model. $\beta_k(u_i, v_i)$ is the regression parameter model. ε_i is residual at location *i*-th.

The parameter estimation procedure for Geographically Weighted Regression (GWR) is similar to the procedure used to obtain values for each location in a classical regression model. However, when predicting parameters at each location, we take into account information from nearby locations in addition to the information from the location itself. The contribution of a location in determining parameter estimates at the observed location increases with the proximity of that location to the observed one. Weighting is applied to the contributions from the surrounding environment. As a result, the Weighted Least Squares (WLS) method is used to estimate parameters in the GWR model. This method involves assigning different weights to each point where data is observed [19].

2.3. Estimating Parameters Using the Least Absolute Deviation (LAD)

Estimating the regression model as in Equation (1) using the Least Absolute Deviation (LAD) method is performed by minimizing the sum of absolute residuals, as expressed by the Equation 3.

$$\sum_{i=1}^{n} w_i(u_i v_i)|\varepsilon_i| = \sum_{i=1}^{n} \left| y_i - \beta_0(u_i\ v_i) - \sum_{k=1}^{K} \beta_k(u_i\ v_i)\, X_{ik} \right| \qquad (3)$$

The solution to obtain regression parameters $(\beta_0, \beta_{1,} \dots, \beta_k)$ cannot be achieved through a differentiation process, as the Weighted Least Squares (WLS) method. This is because the form of Equation 3 is non-differentiable at all points [5].

3. Materials and Methods

3.1. Data

The data was sourced from the 2020 Indonesian Health Profile, published by the Ministry of Health of the Republic of Indonesia. It relates to MMR and maternal health data for 34 provinces in Indonesia in 2020. The definitions of variables are shown in Table 1.

3.2. Data Analysis Procedure

The data analysis in this research involves several steps. The steps undertaken to achieve the research objectives are as follows: 1) conduct descriptive analysis of the dependent and independent variables by mapping the spatial distribution of both variable types. Explore initial characteristics through summary statistics used in the study [21–24]. 2) Examine multicollinearity in the independent variables by assessing the Variance Inflation Factor (VIF). If any independent variable has a VIF ≥ 10, it indicates multicollinearity [25–28]. 3) Perform a spatial heterogeneity test using the Breusch Pagan test. The objective is to identify spatial variation [29–31]. 4) Conduct RGWR analysis due to spatial heterogeneity and the presence of outliers [32]. 5) Perform a significance test on RGWR model parameters to identify which

Table 1. Variable definition

Variables	Variable Definition
Y	Number of maternal deaths during pregnancy, childbirth, and postpartum period
X_1	Percentage of births attended by health workers
X_2	Percentage of deliveries assisted by health facilities
X_3	Percentage of postpartum women receiving vitamin A
X_4	Percentage of services for pregnant women K1 (first quarter period)
X_5	Percentage of maternity services K4 (period close to delivery)
X_6	Percentage of health centers conducting pregnancy classes
X_7	Percentage of 9-month pregnant women receiving blood supplement tablets
X_8	Percentage of pregnant women at risk of chronic energy deficiency
X_9	Number of births waiting for homes

Table 2. Descriptive statistics.

Variables	Min	Max	Mean	Median	Standard Deviation
Y	18	745	136.1	86.5	161.37
X_1	0,00	108.30	81.66	86.80	21.45
X_2	31.40	99.60	78.79	82.85	19.61
X_3	0.00	106.50	80.84	85.00	21.79
X_4	0,00	108.80	86,70	90,35	19,44
X_5	27,50	98,90	77,40	80,95	19.40
X_6	5.61	100.00	77.15	85.30	26,23
X_7	25.30	98.10	80.22	84.05	17.06
X_8	65.70	100.00	92.64	95.90	8.90
X_9	0.00	233.00	36.71	16.00	54.44

variables significantly influence AKI in each province [33]. 6) Create a map illustrating the distribution of AKI based on significant variables. 7) Interpret one of the RGWR models obtained. 8) Conclude and provide recommendations based on the analysis results.

4. Results and Discussion

4.1. Descriptive Statistics

Descriptive statistics can be used to describe all research variables to determine the characteristics of the variables used, the results can be seen in Table 2. The descriptive analysis presented information on the maximum, minimum, mean, median, and standard deviation of each research variable. These results show that the MMR in Indonesia varies, with the highest MMR recorded in West Java at 745 deaths, and the lowest MMR recorded in North Kalimantan at only 18 deaths. The national average MMR is 136. Factors that may influence the MMR, such as the percentage of deliveries assisted by health professionals, the percentage of deliveries assisted by health facilities, the percentage of postpartum mothers receiving vitamin A, the percentage of antenatal care services for K1 and K4, the percentage of health centers conducting antenatal classes, the percentage of pregnant women receiving iron supplementation tablets, the percentage of pregnant women at risk of chronic energy deficiency, and the presence of maternity waiting homes in each province, are also reported. For example, the highest percentage of deliveries assisted by health professionals is 108.3% in Banten province, while the highest percentage of maternal care services for K4 is recorded at 98.90% in DKI Jakarta province. With the average percentage for each variable, this study provides a comprehensive overview of the factors that may contribute to variations in MMR across different provinces in Indonesia.

Figure 1 shows that the MMR in Indonesia in 2020 in each province is divided into three categories, namely the first category there are 3 provinces, namely West Java, East Java and Central Java Provinces with a value range of 243 to 745. Furthermore, the second category has 12 provinces, namely Aceh, North Sumatra, West Sumatra, Riau, South Sumatra, Lampung, DKI Jakarta, Banten, West Nusa Tenggara, East Nusa Tenggara, West Kalimantan and South Sulawesi including with a value range of 98 to 242 and the third category has 19 provinces, namely Jambi, Bengkulu, Bangka Belitung Islands, Riau Islands, DI Yogyakarta, Bali, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku West Papua and Papua with a value range of 18 to 97.

4.2. Spatial Heterogeneity and Spatial Weight Matrix

The assumption of spatial heterogeneity aims to determine whether in the regression model there is inequality in the variance of the residuals between one observation and another [34, 35]. Based on the results of

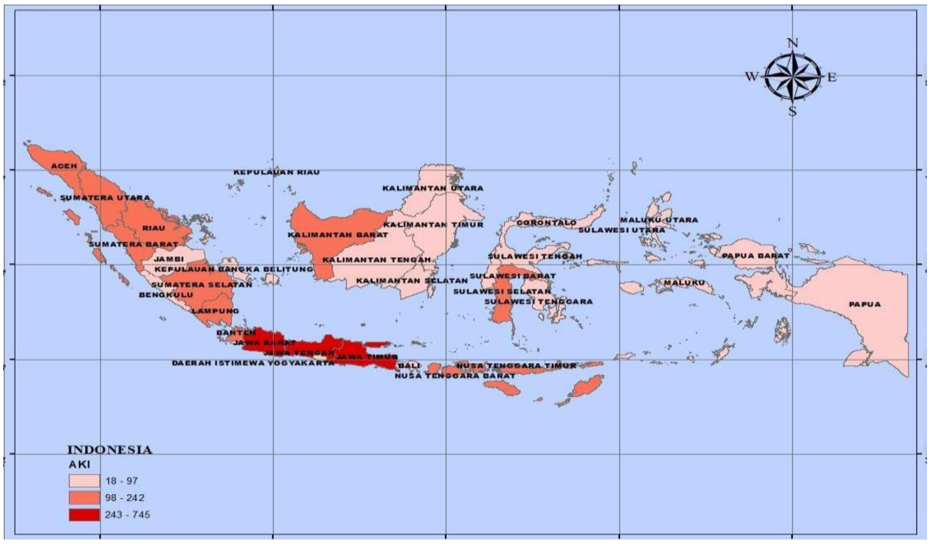

Figure 1. Distribution of MMR in Indonesia in 2020.

Table 3. Kernel weighting functions.

Bandwidth	Kernel Function	R^2	AIC
Fixed	Gaussian	0.905	35.488
	Bisquare	0.753	59.077
	Tricube	0.762	58.691
	Exponential	0.936	23.99
Adaptive	Gaussian	0.650	66.937
	Bisquare	0.859	43.119
	Tricube	0.853	44.043
	Exponential	0.697	63.840

Table 4. VIF values for multicollinearity detection.

Variable	VIF
X_1	17.76
X_2	15.37
X_3	31.45
X_4	30.81
X_5	13.50
X_6	2.05
X_7	4.55
X_8	2.34
X_9	1.18

Table 5. The RGWR model parameters.

Parameter	Min	Q1	Median	Q3	Max
β_0	-0.397	-0.246	0.263	0.547	1.093
B_6	-0.752	-0.436	-0.350	-0.221	-0.050
B_7	0.732	0.023	0.093	0.143	0.467
B_8	-1.140	-1.038	-0.852	-0.362	0.090
B_9	-0.490	-0.029	0.067	0.090	0.121

the Breusch-Pagan analysis, the P-value (0.04254) < α (0.05) is obtained, thus indicating that there is spatial heterogeneity or there are different characteristics in maternal mortality data in Indonesia in 2020 in each province.

There are several ways to determine spatial weights, including using kernel weighting functions consisting of four types of kernel functions: Gaussian, Exponential, Bisquare, and Tricube [36, 37]. This study selected the optimum kernel weighting function based on the minimum Akaike Information Criterion (AIC) and the maximum R^2 value criteria. The following are the AIC and R^2 value obtained from the kernel function.

Based on Table 3, the Fixed Kernel Exponential weighting function has the maximum R^2 value (0.9366) and the minimum AIC value (23.9902), hence it will be used to determine the optimum bandwidth in the RGWR model with the ACV method. The optimum bandwidth result obtained is 3.214 with the minimum ACV value of 18,470

4.3. Robust Geographically Weighted Regression (RGWR)

Before running a regression model, it is important to check for multicollinearity to detect correlations between independent variables [38–40]. Multicollinearity reduces the precision of the estimated coefficients, which weakens the statistical power of the regression model [41–43]. p-values may not be reliable for the identification of independent variables that are statistically significant [44–46]. The value of the variance inflation factor (VIF) can

Table 6. Parameter significance test on the RGWR model for each province.

No	Province	t statistics			
		X6	X7	X8	X9
1	Aceh	-0.685	0.183	-0.708	0.478
2	Sumatera Utara	-1.338	0.280	-2.135*	0.468
3	Sumatera Barat	-1.792	0.408	-4.039*	0.445
4	Riau	-1.787	0.314	-3.792*	0.370
5	Jambi	-1.931	0.221	-5.401*	0.220
6	Sumatera Selatan	-1.790	0.054	-5.767*	0.001
7	Bengkulu	-1.892	0.218	-5.577*	0.059
8	Lampung	-1.531	-0.666	-6.000*	-0.700
9	Kepulauan Bangka Belitung	-2.017	-0.486	-6.210*	-0.556
10	Kepulauan Riau	-2.549*	0.474	-3.712*	-0.453
11	DKI Jakarta	-1.146	-1.514	-6.437*	-1.941
12	Jawa Barat	-1.872	-1.599	-6.616*	-2.072*
13	Jawa Tengah	-3.336*	-1.505	-6.946*	-1.788
14	DI Yogyakarta	-3.164*	-1.334	-6.900*	-1.533
15	Jawa Timur	-3.399*	-0.668	-5.992*	-1.312
16	Banten	-0.901	-1.257	-6.248*	-1.848
17	Bali	-2.130*	0.439	-5.589*	-0.063
18	Nusa Tenggara Barat	-1.792	1.418	-4.691*	0.483
19	Nusa Tenggara Timur	-1.115	1.961	-3.883*	0.39
20	Kalimantan Barat	-2.667*	0.475	-6.433*	-0.118
21	Kalimantan Tengah	-2.272*	0.425	-5.829*	-0.217
22	Kalimantan Selatan	-1.725	0.400	-5.317*	-0.009
23	Kalimantan Timur	-0.95	0.661	-2.802*	0.503
24	Kalimantan Utara	-1.402	0.625	-1.291	0.803
25	Sulawesi Utara	-0.794	0.236	-0.098	0.600
26	Sulawesi Tengah	-1.01	0.511	-0.844	0.474
27	Sulawesi Selatan	-1.424	0.807	-2.515*	0.42
28	Sulawesi Tenggara	-1.108	1.007	-1.411	0.452
29	Gorontalo	-0.824	0.276	-0.275	0.54
30	Sulawesi Barat	-1.441	0.51	-2.672*	0.473
31	Maluku	-0.533	0.165	0.235	0.526
32	Maluku Utara	-0.370	0.213	0.144	0.644
33	Papua Barat	-0.136	0.076	0.158	0.516
34	Papua	-0.144	0.206	-0.046	0.198

be used to detect multicollinearity, where a VIF value greater than 10 indicates multicollinearity.

Based on Table 4, VIF variables X_1, X_2, X_3, X_4, and X_5 are greater than 10, which means there is multicollinearity between these variables. Hence, the independent variables that have multicollinearity are eliminated from the regression model.

The estimation results of the RGWR model parameters with the Least Absolute Deviation method of the Exponential Fixed Kernel function can be seen in Table 5. The values in Table 5 summarize parameter estimates for each observation point and are not applicable globally because each observation has different weights. Furthermore, the parameter significance test of the RGWR model was conducted to find out which independent variables have a significant effect on MMR in each province in Indonesia in 2020, Each province has its own model because the parameter characteristics are different from other regions.

H0: $\beta_j(u_i, v_i) = 0$ (Independent variables do not affect MMR)

H0: $\beta_j(u_i, v_i) \neq 0$ (Independent variables affect MMR)

The significant test of the RGWR model parameters is obtained by comparing the $t_{statistics}$ value and the t_{table} value. The $t_{table} = t_{(0.025;29)} = 2.045$, if the $t_{statistics}$ value at the *i*-th location is greater than the t_{table} value, then the parameter is declared to have a significant effect on MMR at the *i*-th observation location. Based on Table 6, four groups of provinces in Indonesia can be formed based on the similarity of independent variables that have a significant effect on MMR. The grouping can be seen in Figure 2.

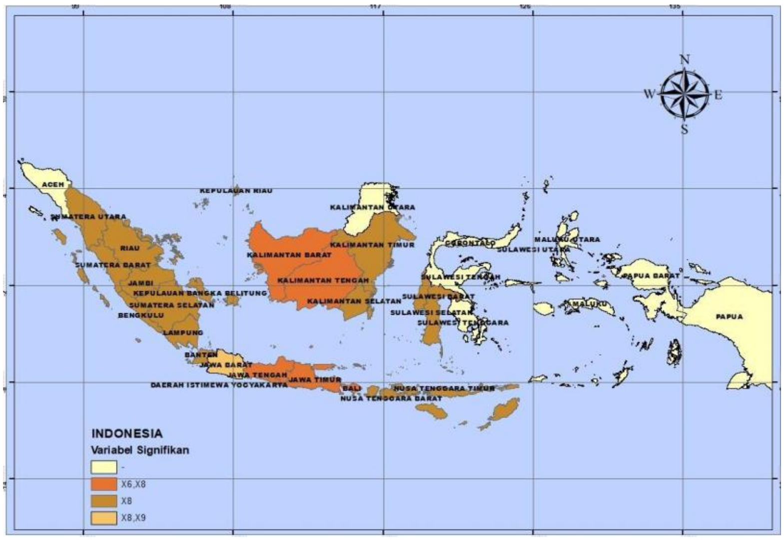

Figure 2. Provincial groups based on significant variables in the RGWR model.

Figure 2 shows that there are four groups of provinces based on variables that significantly affect MMR in Indonesia in 2020. There are 16 provinces in the first group with a significant variable is the percentage of pregnant women at risk of chronic energy deficiency (X_8). there are 7 provinces in the second group with significant variables are the percentage of health centers conducting classes for pregnant women (X_6) and the variable percentage of pregnant women at risk of chronic energy deficiency (X_8). there is only 1 province in the third group with significant variables are the percentage of pregnant women at risk of chronic energy deficiency (X_8) and the variable number of waiting for births (X_9). there are 10 provinces in the last group with no significant variables that affect MMR in Indonesia in 2020.

5. Conclusions

The application of RGWR analysis in identifying factors influencing MMR in Indonesia for the year 2020 has yielded notable insights. The findings reveal that among the various RGWR models tested, the Exponential Fixed Kernel weighting function, utilizing the Least Absolute Deviation method, exhibited superior performance with the smallest AIC value. This led to the derivation of 34 distinct models for each province in Indonesia. There are 4 groups of provinces based on significant variables, where each province in Indonesia that are close together tend to have similarities in variables that significantly affect maternal mortality rates. However, this study is not without limitations. The analysis is confined to data available for the year 2020, potentially limiting the comprehensive understanding of long-term trends or dynamic changes over time. Additionally, while the identified variables are significant, other contextual or socio-economic factors that were not part of this study could also influence MMR.

Moving forward, future research endeavors could encompass a broader temporal scope, capturing trends over multiple years, to offer a more comprehensive understanding of evolving patterns in MMR and its influencing factors across Indonesia. Furthermore, exploring additional variables and incorporating socio-cultural or healthcare infrastructure factors might enrich the depth of analysis, providing a more holistic perspective on maternal healthcare challenges and effective intervention strategies. Moreover, employing advanced modeling techniques or integrating multidisciplinary approaches could enhance the accuracy and applicability of the predictive models, facilitating more targeted and effective policy-making in maternal healthcare

Author Contributions: Conceptualization, N.R.S.; methodology, N.R.S. and H.S.; software, E.M.U.; validation, N.R.S., R.K. and S.M.; formal analysis, H.S.; investigation, N.R.S., R.K. and S.M.; resources E.M.U.; data curation, R.K. and S.M.; writing—original draft preparation, N.R.S., H.S., R.K. and S.M.; writing—review and editing, L.R.; visualization, H.S.; supervision, N.R.S.; project administration; N.R.S. All authors have read and agreed to the published version of the manuscript

Funding: This study does not receive external funding.

Ethical Clearance: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request from the authors.

Acknowledgments: We extend our sincere gratitude to the anonymous reviewers whose thorough and constructive feedback greatly strengthened the quality and clarity of this

article. Their expertise and invaluable insights have been instrumental in shaping this work.

Conflicts of Interest: All the authors declare that there are no conflicts of interest.

References

1. Sasmita, N. R., Ramadeska, S., Utami, R., Adha, Z., Putri, U., Haezah, R., Reskiaddin, L. O., Kamal, S., Yarmaliza, Y., Muliadi, M., and Saputra, A. (2023). Maternal and Child Healthcare Services in Aceh Province , Indonesia : A Correlation and Clustering Analysis in Statistics, *Infolitika Journal of Data Science*, Vol. 1, No. 1, 22–31. doi:10.60084/ijds.v1i1.88.
2. Intan Fazrin, Katarina Kaka Daha, and Kamaru Ilmron Musa. (2022). The Role of Parents in Preparing Balanced Menu with Children's Nutritional Status, *Journal Of Nursing Practice*, Vol. 5, No. 2, 229–238. doi:10.30994/jnp.v5i2.149.
3. Zuber, P. L. F., Moran, A. C., Chou, D., Renaud, F., Halleux, C., Peña-Rosas, J. P., Viswanathan, K., Lackritz, E., Jakob, R., Mason, E., Lamprianou, S., and Guillard-Maure, C. (2018). Mapping the landscape of global programmes to evaluate health interventions in pregnancy: the need for harmonised approaches, standards and tools, *BMJ Global Health*, Vol. 3, No. 5, e001053. doi:10.1136/bmjgh-2018-001053.
4. Zhang, H., Dong, G., Wang, J., Zhang, T.-L., Meng, X., Yang, D., Liu, Y., and Lu, B. (2023). Understanding and extending the geographical detector model under a linear regression framework, *International Journal of Geographical Information Science*, Vol. 37, No. 11, 2437–2453. doi:10.1080/13658816.2023.2266497.
5. 'Afifah, R., Andriyana, Y., and Jaya, I. G. N. M. (2017). Robust geographically weighted regression with least absolute deviation method in case of poverty in Java Island, *Proceedings of the 2nd International Conference on Applied Statistics (ICAS II)*, AIP Publishing, Jawa Barat, 020023. doi:10.1063/1.4979439.
6. Warsito, B., Yasin, H., Ispriyanti, D., and Hoyyi, A. (2018). Robust geographically weighted regression of modeling the Air Polluter Standard Index (APSI), *Journal of Physics: Conference Series*, Vol. 1025, 012096. doi:10.1088/1742-6596/1025/1/012096.
7. Nurhayati, I. C., Rusgiyono, A., and Yasin, H. (2020). Robust Geographically Weighted Regression Dengan Metode Mutlak Simpangan Terkecil Pada Pemodelan Kejadian Diare Di Kota Semarang, *Jurnal Gaussian*, Vol. 7, No. 2, 1–9. doi:https://doi.org/10.14710/j.gauss.7.2.143-152.
8. Francis, M. R., P.S, R., Mohan, V. R., Balraj, V., and George, K. (2012). Examining spatial patterns in the distribution of Low Birth Weight babies in Southern India- the role of maternal, socio-economic and environmental factors, *International Journal of Biological & Medical Research*, Vol. 3, No. 1, 1254–1259.
9. Hailegebreal, S., Haile, F., Haile, Y., Simegn, A. E., and Enyew, E. B. (2023). Using geographically weighted regression analysis to assess predictors of home birth hot spots in Ethiopia, *PLOS ONE*, Vol. 18, No. 6, e0286704. doi:10.1371/journal.pone.0286704.
10. Noviandy, T. R., Maulana, A., Sasmita, N. R., Suhendra, R., Irvanizam, I., Muslem, M., Idroes, G. M., Yusuf, M., Sofyan, H., Abidin, T. F., and Idroes, R. (2022). The Prediction of Kovats Retention Indices of Essential Oils At Gas Chromatography Using Genetic Algorithm-Multiple Linear Regression and Support Vector Regression, *Journal of Engineering Science and Technology*, Vol. 17, No. 1, 306–326.
11. Idroes, R., Noviandy, T. R., Maulana, A., Suhendra, R., Sasmita, N. R., Muslem, M., Idroes, G. M., Kemala, P., and Irvanizam, I. (2021). Application of Genetic Algorithm-Multiple Linear Regression and Artificial Neural Network Determinations for Prediction of Kovats Retention Index, *International Review on Modelling and Simulations (IREMOS)*, Vol. 14, No. 2, 137. doi:10.15866/iremos.v14i2.20460.
12. Idroes, R., Noviandy, T. R., Maulana, A., Suhendra, R., Sasmita, N. R., Muslem, M., Idroes, G. M., and Irvanizam, I. (2019). Retention Index Prediction of Flavor and Fragrance by Multiple Linear Regression and the Genetic Algorithm, *International Review on Modelling and Simulations (IREMOS)*, Vol. 12, No. 6, 373–380. doi:10.15866/iremos.v12i6.18353.
13. Marcinkiewicz, K., and Schnapp, K.-U. (2021). Regression Analysis, *Research Methods in the Social Sciences: An A-Z of Key Concepts*, Oxford University Press, 233–238. doi:10.1093/hepl/9780198850298.003.0054.
14. Agustia, M., Noviandy, T. R., Maulana, A., Suhendra, R., Muslem, M., Sasmita, N. R., Idroes, G. M., Rahimah, S., Afidh, R. P. F., Subianto, M., Irvanizam, I., and Idroes, R. (2022). Application of Fuzzy Support Vector Regression to Predict the Kovats Retention Indices of Flavors and Fragrances, *2022 International Conference on Electrical Engineering and Informatics (ICELTICs)* (Vol. 2022-Septe), IEEE, 13–18. doi:10.1109/ICELTICs56128.2022.9932124.
15. Montgomery, D. C., and Peck, E. A. (1992). *Introduction to Linear Regression Analysis* (Second.), Wiley, New. York.
16. Rahayu, L., Sasmita, N. R., Adila, W. F., Kesuma, Z. M., and Kruba, R. (2023). Spatial Estimation of Relative Risk for Dengue Fever in Aceh Province using Conditional Autoregressive Method, *Journal of Applied Data Sciences*, Vol. 4, No. 4, 466–479. doi:10.47738/jads.v4i4.141.
17. Sasmita, N. R., Khairul, M., Sofyan, H., Kruba, R., Mardalena, S., Dahlawy, A., Apriliansyah, F., Muliadi, M., Saputra, D. C. E., Noviandy, T. R., and Maula, A. W. (2023). A Statistical Clustering Approach: Mapping Population Indicators Through Probabilistic Analysis in Aceh Province, Indonesia, *Infolitika Journal of Data Science*, Vol. 1, No. 2, 1–9. doi:https://doi.org/10.60084/ijds.v1i2.130.
18. Sasmita, N. R., Phonna, R. A., Fikri, M. K., Khairul, M., Apriliansyah, F., Idroes, G. M., Puspitasari, A., and Saputra, F. E. (2023). Statistical Assessment of Human Development Index Variations and Their Correlates : A Case Study of Aceh Province , Indonesia, *Grimsa Journal of Business and Economics Studies*, Vol. 1, No. 1, 1–13.
19. Fadliana, A., Pramoedyo, H., and Fitriani, R. (2019). Parameter Estimation of Locally Compensated Ridge-Geographically Weighted Regression Model, *IOP Conference Series: Materials Science and Engineering*, Vol. 546, No. 5, 052022. doi:10.1088/1757-899X/546/5/052022.
20. Liu, Y., and Goudie, R. J. B. (2023). Generalized Geographically Weighted Regression Model within a Modularized Bayesian Framework, *Bayesian Analysis*, Vol. 1, No. 1. doi:10.1214/22-BA1357.
21. Avval, T. G., Moeini, B., Carver, V., Fairley, N., Smith, E. F., Baltrusaitis, J., Fernandez, V., Tyler, B. J., Gallagher, N., and Linford, M. R. (2021). The Often-Overlooked Power of Summary Statistics in Exploratory Data Analysis: Comparison of Pattern Recognition Entropy (PRE) to Other Summary Statistics and Introduction of Divided Spectrum-PRE (DS-PRE), *Journal of Chemical Information and Modeling*, Vol. 61, No. 9, 4173–4189. doi:10.1021/acs.jcim.1c00244.
22. Idroes, G. M., Hardi, I., Noviandy, T. R., Sasmita, N. R., Hilal, I. S., Kusumo, F., and Idroes, R. (2023). A Deep Dive into Indonesia's CO2 Emissions: The Role of Energy Consumption, Economic Growth and Natural Disasters, *Ekonomikalia Journal of Economics*, Vol. 1, No. 2, 69–81. doi:10.60084/eje.v1i2.115.
23. Iffaty, A., Salsabila, A., Rafiqhi, A. A., Suhendra, R., Yusuf, M., and Sasmita, N. R. (2023). Enhancing Water Quality Assessment in Indonesia Through Digital Image Processing and Machine Learning, *Grimsa Journal of Science, Engineering and Technology*, Vol. 1, No. 1, 1–8.
24. Suhendra, R., Suryadi, S., Husdayanti, N., Maulana, A., Noviandy, T. R., Sasmita, N. R., Subianto, M., Earlia, N., Niode, N. J., and Idroes, R. (2023). Evaluation of Gradient Boosted Classifier in

Atopic Dermatitis Severity Score Classification, *Heca Journal of Applied Sciences*, Vol. 1, No. 2, 54–61. doi:10.60084/hjas.v1i2.85.

25. Senaviratna, N. A. M. R., and A. Cooray, T. M. J. (2019). Diagnosing Multicollinearity of Logistic Regression Model, *Asian Journal of Probability and Statistics*, 1–9. doi:10.9734/ajpas/2019/v5i230132.

26. Rachmawati, R., Novita, R., Erwandi, E., Fitriyanin, E., Idroes, R., and Reandy Sasmita, N. (2020). The Effectiveness of Counseling on Complementary Food for Mothers and Supplementary Feeding for Increasing weight of Breastfed babies (12-24 Months of Age) in the Aceh Besar District of Indonesia, *Pakistan Journal of Nutrition*, Vol. 19, No. 6, 279–284. doi:10.3923/pjn.2020.279.284.

27. Sasmita, N. R., Ikhwan, M., Suyanto, S., and Chongsuvivatwong, V. (2020). Optimal control on a mathematical model to pattern the progression of coronavirus disease 2019 (COVID-19) in Indonesia, *Global Health Research and Policy*, Vol. 5, No. 1, 1–38. doi:10.1186/s41256-020-00163-2.

28. Earlia, N., Bulqiah, M., Muslem, M., Karma, T., Suhendra, R., Maulana, A., Mohamad Amin, M. A., Sasmita, N. R., Idroes, G. M., Prakoeswa, C., Suhartono, E., Khairan, K., and Idroes, R. (2021). Protective Effects of Acehnese Traditionally Fermented Coconut Oil (Pliek U Oil) and its Residue (Pliek U) in Ointment against UV Light Exposure: Studies on Male Wistar Rat Skin (Rattus novergicus), *Sains Malaysiana*, Vol. 50, No. 5, 1285–1295. doi:10.17576/jsm-2021-5005-09.

29. Li, Z., and Yao, J. (2019). Testing for heteroscedasticity in high-dimensional regressions, *Econometrics and Statistics*, Vol. 9, 122–139. doi:10.1016/j.ecosta.2018.01.001.

30. Idroes, R., Husan, I., Muslem, Mahmudi, Rusyana, A., Helwati, Z., Idroes, G. M., Suhendra, R., Yandri, E., Rahimah, S., and Sasmita, N. R. (2019). Analysis of Temperature and Column Variation in Gas Chromatography to Dead Time of Inert Gas and N-Alkane Homologous Series Using Randomized Block Design, *IOP Conference Series: Earth and Environmental Science*, Vol. 364, No. 012020, 1–7. doi:10.1088/1755-1315/364/1/012020.

31. Nadia, Y., Ramli, M., Muslem, Japnur, A. F., Rusyana, A., Idroes, G. M., Suhendra, R., Muhammad, Sasmita, N. R., Tallei, T. E., Idroes, R., Jafnur, A. F., Rusyana, A., Idroes, G. M., Suhendra, R., Muhammad, Sasmita, N. R., Tallei, T. E., and Idroes, R. (2019). Simple Combination Method of FTIR Spectroscopy and Chemometrics for Qualitative Identification of Cattle Bones, *IOP Conference Series: Earth and Environmental Science*, Vol. 364, No. 1, 1–6. doi:10.1088/1755-1315/364/1/012040.

32. Erda, G., Indahwati, and Djuraidah, A. (2019). Outlier handling of Robust Geographically and Temporally Weighted Regression, *Journal of Physics: Conference Series*, Vol. 1175, 012041. doi:10.1088/1742-6596/1175/1/012041.

33. Sasmita, N. R., Geater, A. F., Kammanee, A., and Chongsuvivatwong, V. (2019). Is The Recovery Rate In Latent Tuberculosis Infection Significant In Reducing Tuberculosis Transmission In Indonesia?: A Mathematical Model Study In Epidemiology, *RSU Research Conference 2019*, Rangsit University, Pathum Thani, Thailand, 101–111.

34. Idroes, R., Maulana, A., Noviandy, T. R., Suhendra, R., Sasmita, N. R., Lala, A., and Irvanizam. A Genetic Algorithm to Determine Research Consultation Schedules in Campus Environment, , 796 IOP Conference Series: Materials Science and Engineering 1–8 (2020), 1–8. doi:10.1088/1757-899X/796/1/012033.

35. Sasmita, N. R., Sofyan, H., and Subianto, M. (2012). The Comparison of Fuzzy C-Means (FCM) and Fuzzy C-Shell (FCS) Methods By Using Quickbird Satellite Image (Case Study In Peukan Bada), *SEMIRATA 2012*, State University of Medan, Medan, 211–218.

36. Sasaki, D., Sofyan, H., Sasmita, N. R., Affan, M., and Nizamuddin, N. (2021). Assessing the Intermediate Function of Local Academic Institutions During the Rehabilitation and Reconstruction of Aceh, Indonesia, *Journal of Disaster Research*, Vol. 16, No. 8, 1265–1273. doi:10.20965/jdr.2021.p1265.

37. Idroes, G. M., Syahnur, S., Majid, S. A., Sasmita, N. R., and Idroes, R. (2021). Provincial economic level analysis in Indonesia based on the geothermal energy potential and growth regional domestic products using cluster analysis, *IOP Conference Series: Materials Science and Engineering*, Vol. 1087, No. 1, 012079. doi:10.1088/1757-899X/1087/1/012079.

38. Singh, P., Singh, S., and Paprzycki, M. (2023). Detection and elimination of multicollinearity in regression analysis, *International Journal of Knowledge-Based and Intelligent Engineering Systems*, Vol. 27, No. 1, 105–111. doi:10.3233/KES-221622.

39. Idroes, R., Noviandy, T. R., Maulana, A., Suhendra, R., and Sasmita, N. R. (2023). ANFIS-Based QSRR Modelling for Kovats Retention Index Prediction in Gas Chromatography, *Infolitika Journal of Data Science*, Vol. 1, No. 1, 1–14. doi:10.60084/ijds.v1i1.73.

40. Maulana, A., Noviandy, T. R., Sasmita, N. R., Paristiowati, M., Suhendra, R., Yandri, E., Satrio, J., and Idroes, R. (2023). Optimizing University Admissions: A Machine Learning Perspective, *Journal of Educational Management and Learning*, Vol. 1, No. 1, 1–7. doi:10.60084/jeml.v1i1.46.

41. Ullah, M. I., Aslam, M., Altaf, S., and Ahmed, M. (2019). Some New Diagnostics of Multicollinearity in Linear Regression Model, *Sains Malaysiana*, Vol. 48, No. 9, 2051–2060. doi:10.17576/jsm-2019-4809-26.

42. Sofyan, H., Diba, F., Susanti, S. S., Marthoenis, M., Ichsan, I., Sasmita, N. R., Seuring, T., and Vollmer, S. (2023). The state of diabetes care and obstacles to better care in Aceh, Indonesia: a mixed-methods study, *BMC Health Services Research*, Vol. 23, No. 1, 1–7. doi:10.1186/s12913-023-09288-9.

43. Maulana, A., Noviandy, T. R., Idroes, R., Sasmita, N. R., Suhendra, R., and Irvanizam, I. (2020). Prediction of Kovats Retention Indices for Fragrance and Flavor using Artificial Neural Network, *Proceedings of the International Conference on Electrical Engineering and Informatics*, Vols 2020-Octob, 1–5. doi:10.1109/ICELTICs50595.2020.9315391.

44. Maiti, S. I., and Saikia, S. K. (2019). p<0.05 Is Not Enough, *Advances in Zoology and Botany*, Vol. 7, No. 2, 24–27. doi:10.13189/azb.2019.070202.

45. Azharuddin, Sasmita, N. R., Idroes, G. M., Andid, R., Raihan, Fadlilah, T., Earlia, N., Ridwan, T., Maya, I., Farnida, and Idroes, R. (2023). Patient Satisfaction And Its Socio-Demographic Correlates In Zainoel Abidin Hospital, Indonesia: A Cross-Sectional Study, *Unnes Journal of Public Health*, Vol. 12, No. 2, 57–67. doi:10.15294/ujph.v12i2.69233.

46. Noviandy, T. R., Maulana, A., Sasmita, N. R., Suhendra, R., Muslem, Idroes, G. M., Paristiowati, M., Helwani, Z., Yandri, E., Rahimah, S., Muhammad, Irvanizam, and Idroes, R. (2021). The implementation of K-Means clustering in kovats retention index on gas chromatography, *IOP Conference Series: Materials Science and Engineering*, Vol. 1087, No. 1, 012051. doi:10.1088/1757-899X/1087/1/012051.