Journal of Computer Science, Information Technology and Telecommunication Engineering (JCoSITTE)
Vol. 5, No. 2, September 2024, pp. 679~685
ISSN: 2721-3838, DOI: 10.30596/jcositte.v5i2.20930

Application of Geographically Weighted Panel Regression Based on Seasonal Pattern in Data for the Development of Dengue Hemorrhagic Fever (DHF) in North Sumatra Province

Dinda Kusuma Putri[1], Fibri Rakhmawati[2]
[1,2]Department of Mathematic, State Islamic University of North Sumatera, Indonesia

ABSTRACT

Dengue Hemorrhagic Fever (DHF) is a disease caused by the dengue virus which is transmitted to humans through the bites of Aedes aegypti and Aedes albopictus mosquitoes which were previously infected with the dengue virus from other dengue fever sufferers. This research aims to apply the Geographically Weighted Panel Regression (GWPR) method. in identifying what factors influence the development of Dengue Hemorrhagic Fever in 33 districts in North Sumatra Province starting from September 2022-August 2023 with a rainy and dry season pattern. The results of this research show that the appropriate GWPR model is the fixed effect model with the pepmbobot kerne bisquare function. Based on this model, the districts/cities of North Sumatra are divided into ten groups based on variables with a high significant influence on the percentage of development of Dengue Hemorrhagic Fever. The factors that influence this percentage are gender, namely male and female, then Age, namely ≤ 14 Years and Age ≥ 14 Years.

Keyword : Dengue Hemorrhagic Fever(DHF),GWPR,Fixed Effect Model

Corresponding Author:
Dinda Kusuma Putri
Department of Mathematic
State Islamic University Of North Sumatera
Jl. Lapangan Golf No1. 120,Kp Tengah,Kec.PancurBatu,Sumatera Utara,Indonesia
Email : dinda0703202008@uinsu.ac.id

Article history:
Received Aug 14, 2024
Revised Aug 18, 2024
Accepted Aug 20, 2024

1. INTRODUCTION

Dengue virus reason disease fever Dengue fever (DHF) is contagious to man through bite mosquito Aedes aegypti and Aedes albopictus already infected with the virus from patient fever bloody other. Mosquito Female Aedes aegypti, also known as mosquito fever bleeding, is vector main disease this is in Indonesia. Characteristics typically is own scales with pattern white silvery all over body and legs(Gamayanti et al., 2023). Mosquito bites in the morning and evening, usually on the surface of stagnant water or in a damp place(Al'afi et al., 2020). In Indonesia, fever bloody is disease endemic with amount case high annual rate. Often, disease This result death for those who suffer from it(Pratama et al., 2021)

Tropical and subtropical areas is major global center disease fever dengue bleeding. Disease This is disease with fastest growing, with improvement global frequency as much as 30 times for 50 years final(Umi et al., 2023). Disease This caused by a virus carried by mosquitoes. Even if it's an adult sometimes Can caught disease Here, the children are below age fifteen year are the main victims disease This. Fever bloody moment This is transmitted disease through the most common and growing vector rapidly in the world(Indrasetianingsih & Wasik, 2020). Worldwide, 1.3 billion people live in regions endemic fever bleeding, and 2.5 billion people live in endemic countries fever bloody Where fever bloody Can attack (Rusgiyono & Prahutama, 2021)

In North Sumatra, the number case fever bloody tend increase every the year. There were 8,541 cases fever bleeding in North Sumatra in 2022, increasing sharp of 2,918 cases in year previously. (Ministry of Health, North Sumatra, 2023) . With so We can see that dengue fever is still there ongoing problems increasing in North Sumatra.

The Geographically Weighted Regression (GWR) model and the Panel Data Regression model are combined in GWPR approach. With consider factor geographically, the GWPR model is version local from the panel regression model, according to Ni Made, Suyitno, and Meiliyani (2023) citing Cao K &

Wu B. As far as we know, the global regression model may ignore variant local certain and only reflect global trends.

2. RESEARCH METHOD

Study This carried out at the North Sumatra Provincial Health Service.Time Study This Done from January 2024 to finished. Data used is the affected data Fever Dengue bleeding in September 2022-August 2023 with use patt ern existing season that is Rain and Dry(Wati & Utami, 2020).

Variable (Y) in study This is Number of people affected Fever Dengue blood in districts in North Sumatra. While X is used for the predictor variable used (X_1) Labor Medical, (X_2) It's dense Amount population, (X_3) Type Sex that is Male (X_{3a}) and Female (X_{3b}), (X_4) Health Facilities, namely Level 1(X_{4a}) and Level 2(X_{4b}) ,(X_5) Age that is ≤ 14 Years (X_{5b}) and Age ≥ 14 Years (X_{5b}).

Steps taken in study This is as following:

1. Collect data from source study
2. Do analysis descriptive to variable study.
3. Estimating the parameters of a regression model on panel data
4. Selection of the best model with using the Chow test, Hausman test, Lagrange multiplier test
5. Test assumptions from panel data regression
6. Heterogeneity test spatial with the Breusch-pagan test
7. Input Latitude and longitude data
8. Count distance Euclidean between location (u_i, v_i) to location (u_j, v_j) with formula
   $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$
9. Determine optimum bandwidth for determine bandwidth can seen from minimum CV (Cross validation) value with equation (2.30)
10. Parameter estimates for get the final model fixed effects GWPR.
11. Simultaneous test and partial test of the GWPR model

Panel data regression can be estimated using three different methods: common effect model, fixed effect model, and random effect model techniques.

2.1 Geographically Weighted Panel Regression (GWPR)

1. Common Effects Model (CEM)

The common effect model (CEM) is a panel data estimation technique that combines all data without taking into account differences in time or units(Raihani et al., 2023). This approach operates with the assumption that the intercept value is constant, meaning that the intercept value remains the same for all different units and times(Martha et al., 2021). The following is a general effects model:

$$y_{i,t} = \beta_0 + \beta_1 x_1 + \cdots + \beta_k X_{k,it} + \varepsilon_{i,t} \quad (1)$$

2. Fixed Effect Model (FEM)

Allowing the intercept value to fluctuate for each unit of cross section while maintaining the assumption of stability, the slope coefficient is one way to pay attention to the unit of cross section, [10]. Fixed Effects in the transformation model is one method that can be used to estimate fixed effects models, according to Wooldridge (2002). Because the slope coefficient of the model is constant, the following model is produced. (Woolridge, 2012)

$$y_{it} = \beta_1 x_{1,it} + \beta_2 x_{2,it} + \cdots + \beta_k x_{k,it} + \alpha_1 + \varepsilon_{it} \quad (2)$$

3. Random Effect Model (REM)

Errors are used in random effects models to account for variations in unobserved effects. Error is considered a random variable that fluctuates between individual units. Additionally, unobserved impacts are assumed to be uncorrelated with any predictor variables (Woolridge, 2012). Random effects models, in particular

$$y_{it} = \beta_1 x_{1,it} + \beta_2 x_{2,it} \cdots + \beta_k X_{k,it} + v_{it} \quad (3)$$

2.2 Panel Data Regression Model Selection Test

1. Test Chow

The best model to use, namely between CEM and FEM, is determined using the Chow test. Following is Chow's test hypothesis:

$H_0 = \alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_{10} = 0$(slope and intercept are the same (following the common effects model))

H $_1$ = There is at least one intercept (α_i) that is not the same (Same slope and different intercept (following the fixed effects model), where i= 1,2,...,10; t=1,2,3[11].

The GWPR model operates on the premise that spatiotemporal smooth processes are realized in a time series of observations at a particular place. Nearby observations are more related to distant observations in the distribution of these processes. Therefore, GWPR model parameters will vary by location[12]. The GWPR model was created by combining the GWR model with the Fixed Effect Model (FEM) panel data regression model with an inside estimator[13]. The GWPR model has the following expression in equation (2.3). (Rahayu, 2017)

$$\ddot{y}_{it} = \beta_0(u_{it,}V_{it}) + \sum_{k=1}^{p} \beta_k(u_{it,}v_{it})\ddot{x}_{k,,it} + \ddot{\varepsilon}_{it} \quad (4)$$

$$\beta_0\left(u_{it,}v_{it}\right)\text{And } t = 1, 2, \ldots, T \quad (5)$$

Where:

$\ddot{y}_{it}$:(demeaned) average response value at the *i*th observation and *t* th time.

$\ddot{x}_{k,it}$: corrected (demeaned) average predictor variable value at the *i*th observation and *t* time.

$\beta_0\left(u_{it,}v_{it}\right)$: Constant/intercept of the equation formed at the *i* observation and time t

$\beta_k\left(u_{it,}v_{it}\right)$: Regression coefficients on the kth and kth observations for the average of the kth corrected (reduced) predictor variables.

$\left(u_{it,}v_{it}\right)$: The coordinate point of the observation location at the *i* th observation and the *t* time.

K : Number of predictor variables

$\ddot{\varepsilon}_{it}$: It is assumed that random errors are independent, identical, and normally distributed with zero mean and constant variance.

2.3 Geographically Weighted Panel Regression Model Testing

1. Model Fit Test

The purpose of this test is to find out whether the GWPR fixed effect and the fixed effect panel data regression model (global regression model) are different from each other. The following theory is used[14] . (Leung Mafvitul, 2021)

H_0: β_k = β_k $(u_{it,}v_{it})$There is no real difference between the GWPR fixed effect model and the global regression model for each variable *k* = 1, 2,..., p and *i* = 1, 2,..., n.

H_1 : there is at least one β_k $(u_{it,}v_{it})$ ≠ β_k for *k* = 1, 2, ..., p and *i* = 1, 2, ..., n (there is a significant difference between the global regression model and the GWPR fixed effect model)[15]

Test Statistics:

$$F_1 = \frac{RSS\left(H_1\right)/df_1}{RSS\left(H_0\right)/df_z} \quad (5)$$

With

$RSS(H_0): Y^T(I-H)Y$, Where $H = X(X^TX)^{-1}X^T$

$RSS(H_1): Y^T(I-L)^T(I-L)^Y$

$df_1 = \frac{\delta_1^2}{\delta_2}$ Where $\delta_1 = tr\left(\left[\left(I-L\right)^T\left(I-L\right)^i\right]\right), i = 1,2; df_2 = n-p-1$

I is the identity matrix measuring $nt \times nt$ and ***L*** is the projection matrix from the GWPR model.

Rejection area: Reject H_0 if $T_{hitung} < F_{1-\alpha, df1, df2}$ or The global regression model (fixed effect panel data regression model) and the GWPR fixed effect model are significantly different as indicated by the p-value >α. .(Journal, 2020)

2. Parameter Significance Test

If the data can be described well by the GWPR model, then a parameter significance test is carried out. The purpose of parameter significance testing is to identify parameters that significantly influence the response variable. The premise guiding this examination is (Wati & Utami, 2020)

$$H_0 : \beta_k\left(u_{1,}u_2\right) = 0$$
$$H_1 : \beta_k\left(u_{1,}u_2\right) \neq 0 \quad (6)$$

Test Statistics:

$$T_{hitung} = \frac{\widehat{\beta}_{k(uij,vij)}}{\widehat{\sigma}\sqrt{C_{kk}}} \quad (7)$$

Where C kk is the kth diagonal element of the matrix C it C T it where C it =

$$\ddot{X}^T W\left(u_{1,}u_2\right)\ddot{X})^{-1}\ddot{X}^T W\left(u_{1,}u_2\right) \quad (8)$$

Rejection area: Reject H_0 if | T_{hitung} | > .$T_{(\alpha/2,df)}$) with $df = \frac{\delta_1^2}{\delta_2}$ or p value< α This shows that the response variable is significantly influenced by the predictor variable parameters.(Journal, 2020)

3. RESULTS AND DISCUSSION

Chow test and Hausman test are two possible testing used for choose an estimation model best from panel data regression.

Table 1.1 Chow Test and Hausman Test

Test Chow		Hausman test	
Results	Conclusion	Results	Conclusion
P -Value= 0.001537 F=3.2421	H0, Rejected means the best model is FEM	P-Value=0.005086 F=16.709	H0, Rejected means the best model is FEM

Based on Chow test table and Hausman test above, yes known in study this is the selected FEM model as the best panel data regression model.

3.1 Test Assumptions Classic Panel Data Regression

shows multicollinearity test results using VIF as following.

Table 2 Multicollinearity Tests

Variable	VIF
X3a	4.89
X3b	2.41
X5a	1.16
X5b	1.45

Results of tests carried out can be seen in Table 4.7 which shows that all over variable independent have VIF value is less from or the same with 10. In conclusion, no There is correlation between variable independent, and problematic multicollinearity No There is related at them.

Autocorrelation Test

Autocorrelation test results with the p value is 0.342 or more p value of 0.05 gives conclusion that final model No contain remainder autocorrelation

Heteroscedasticity Test

The data distribution shows proven heteroscedasticity with heteroscedasticity test results; p − value 0.04788 indicates that p − value < 0.05, rejecting H0. This matter show that assumptions homoscedasticity panel data regression does not fulfilled. One of possibility reason happen heteroscedasticity is the existence of different observation zones. Because of the problem heteroscedasticity in study this, then analysis that focuses on effects location / spatial specifically GWPR method will done.

Geographically Weighted Panel Regression (GWPR)

Election ideal weight for entered to in GWPR modeling was carried out after heteroscedasticity test show exists variation spatial data. Determine the ideal bandwidth and distance between districts and cities in North Sumatra Province before build matrix weighting. With use Euclidean formula, distance between districts and cities calculated.

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}$$

$$= \sqrt{\left(99,3673 - 97,7417\right)^2 + \left(0,74324 - 1,08694\right)^2}$$
$$= \sqrt{\left(2,642575\right)^2 + (0,11813)^2}$$
$$= \sqrt{2,760705}$$
$$= 1,661536954$$

North Sumatra Province has three tens three districts and cities. With so distance North Sumatra Province between districts and cities will produce matrix measuring 33x33. Furthermore use AIC for looking for the ideal bandwidth. Function Adaptive Bisquare weighting produce lowest AIC value among tested kernel functions in study this, so make it choice best for choice weighting. Table 4.10 shows results selection of ideal bandwidth

Table 3. Weighting Kernel Functions

Function Kernel Weighter	AIC	R^2
Adaptive Bissquare	649.0368	0.8285833
Adaptive Gaussian	683.6639	0.6313874
Adaptive Exponential	677.1459	0.6797589

Because of function adaptive kernel weighting Bisquare have mark Akaike Information Criterion (AIC) smallest and value coefficient highest determination (R2). so used as weighting. Furthermore different models are produced For every location based on step parameter and time estimation observation. Table 3 below This serve analysis findings estimation and testing significance of parameters at location observations in the District Mandailing Natal and Padang Lawas.

Table 4. Parameter Significance Tests

Variable	Mandailing Christmas		Padang Lawas	
	Parameter	P-Value	Parameter	P-Value
X3a	-0.22101627	0.293	0.1442	0.635
X3b	1.7038611		0.71301	0.222

		0.031		
X7a	7.510793	0.666	-22,995	0.213
X7b	-7.3895757	0.068	24.948	0.265

Results of estimation and significance testing of parameters in districts Mandailing Christmas is different with Padang Lawas , from fourth variable freely in Mandaling Natal variables X3b, X7a and X7b have an effect positive, meanwhile variable X3a has an effect negative, Pg This show that the percentage affected by dengue fever will the more reduce If type sex man , Age ≥ 14 years, and age ≤ 14 Years the more tall whereas type female gender the more low and in Padang Lawas district fourth variable the influential negative.Pg This means percentage affected by dengue fever in Padang Lawas will the more reduce If four variable free it's in the field old the more tall .

By Overall, we got it identify grouping districts that have factor significant to proportion North Sumatra Province is infected dengue fever, based on significant parameter test results using alpha 0.05.

Table 5. Significance of GWPR Model Parameters

Variable Study	Regency
X3a, X3b, X5a, X5b	Dairi, Nias Utara, West Nias
X3a, X3b, X5a,	Simalungun, Mt Sitoli
X3a, X3b, X5b	Tapanuli Selatan, Humbang Hasundutan, Samosir
X3a, X5a, X5b	Labuhanbatu
X3a, X3b	Tapanuli Tengah, Tapanuli Utara, Toba, West Pakpak
X3a, X5a	Asahan, Labura
X3b, X5b	Mandailing Christmas, Karo, Deli Serdang, Langkat
X3b	Cliff High, Rock Bara, Medan, Binjai, Padang Sidempuan, Serdang Different
X5a	Labuhanbatu Selatan
No Influential	Padang Lawas, and field old North

Based on Application of patterned GWPR models seasons in the table has 10 variations spatial or group on development fever bloody dengu in North Sumatra province, namely gender (X3) influences distribution Fever Dengue blood in the district in North Sumatra province except Regency South Labuhan Batu, Padang Lawas, and North Padang Lawas. Besides that, Regency Dairi, North Nias, West Nias, Simalungun, Mount Sitoli, South Tapanuli, Humbang Hasundutan, Samosir, Labuhan Batu, Asahan, North Labuhan Batu, Mandailing Natal, Karo, Deli Serdang, Langkat, and South Labuhan Batu were affected impact. gender variable. Second factor such, age and type gender, no affected in Padang Lawas and North Padang Lawas Regencies.

4. CONCLUSION

As for the factors that significantly influenced the development of Dengue Hemorrhagic Fever is gender, namely male and female, then age, namely ≤ 14 years and age ≥ 14 yearsof Dairi, North Nias, West Nias, Simalungun, Gunung Sitoli, South Tapanuli, Humbang Hasundutan, Samosir, Labuhan Batu, Asahan, North Labuhan Batu, Mandailing Natal, Karo, Deli Serdang, Langkat, South Labuhan Batu. The two variables, namely gender and age, have no effect in the district. Padang Lawas and North Padang Lawas.

REFERENCES

Al'afi, A. M., Widiart, W., Kurniasari, D., & Usman, M. (2020). Peramalan Data Time Series Seasonal Menggunakan Metode Analisis Spektral. Jurnal Siger Matematika, 1(1), 10–15. https://doi.org/10.23960/jsm.v1i1.2484

Febriyanti, L., Pahlevi, M. R., & Rohaini, E. (2022). Perancangan Sistem Informasi Penjualan Sparepart Dan Jasa Service Motor Pada Bengkel RPM Jakarta. 1(April), 110–119.

Gamayanti, N. F., Junaidi, J., Fadjryani, F., & Nur'eni, N. (2023). Analysis of Spatial Effects on Factors Affecting Rice Production in Central Sulawesi Using Geographically Weighted Panel Regression. BAREKENG: Jurnal Ilmu Matematika Dan Terapan, 17(1), 0361–0370. https://doi.org/10.30598/barekengvol17iss1pp0361-0370

Indrasetianingsih, A., & Wasik, T. K. (2020). Model Regresi Data Panel Untuk Mengetahui Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Pulau Madura. Jurnal Gaussian, 9(3), 355–363. https://doi.org/10.14710/j.gauss.v9i3.28925

Journal, C. (2020). 肠道菌群失调参与 2 型糖尿病患者认知损害的研究进展 1 2. 36, 50–56.

Kusumawati, N., & Sukendra, D. M. (2020). Spasiotemporal Demam Berdarah Dengue berdasarkan House Index, Kepadatan Penduduk dan Kepadatan Rumah. HIGEIA (Journal of Public Health Research and Development), 4(2), 168–177.

Martha, S., Yundari, Y., Rizki, S. W., & Tamtama, R. (2021). Penerapan Metode Geographically Weighted Panel Regression (Gwpr) Pada Kasus Kemiskinan Di Indonesia. BAREKENG: Jurnal Ilmu Matematika Dan Terapan, 15(2), 241–248. https://doi.org/10.30598/barekengvol15iss2pp241-248

Meutuah, M. S., Yasin, H., & Maruddani, D. A. I. (2017). Pemodelan Fixed Effect Geographically Weighted Panel Regression untuk IPM di Jawa Tengah. Jurnal Gaussian, 6(2), 241–250.

Mondiana, Y. Q., Pramoedyo, H., Iriany, A., & Marjono. (2024). Applied fixed effect of Geographically Weighted Panel Regression (GWPR) with M- Estimator approach to estimate sugarcane yield data in East Java. Journal of Applied and Natural Science, 16(2), 646–652. https://doi.org/10.31018/jans.v16i2.5443

Pratama, A., Suyitno, S., & Purnamasari, I. (2021). Pemodelan Persentase Penduduk Miskin Di Provinsi Kalimantan Timur Menggunakan Model Geographically Weighted Panel Regression. Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya), 9(2). https://doi.org/10.24252/msa.v9i2.21021

Raihani, R., Sifriyani, S., & Prangga, S. (2023). Geographically Weighted Panel Regression Modelling of Dengue Hemorrhagic Fever Data Using Exponential Kernel Function. JTAM (Jurnal Teori Dan Aplikasi Matematika), 7(4), 961. https://doi.org/10.31764/jtam.v7i4.16235

Rusgiyono, A., & Prahutama, A. (2021). Geographically Weighted Panel Regression With Fixed Effect for Modeling the Number of Infant Mortality in Central Java, Indonesia. Media Statistika, 14(1), 10–20. https://doi.org/10.14710/medstat.14.1.10-20

Tabrani, M., & Aghniya, I. R. (2019). Jurnal Interkom Vol. 14 No. 1 - April 2019. Jurnal Interkom, 14(1), 44–53.

Umi, K., Junaidi, & Fadjryani. (2023). Gender Inequality Index Modeling in Indonesia Using Geographically Weighted Panel Regression Method. Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-228-6_34

Vebiriyana, M., Darsyah, Y. M., & Nur, I. . (2015). Pemodelan Geographically Weighted Regression Dengan Fungsi Kernel Bisquare Terhadap Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Kabupaten Demak. Jurnal Statistika, 3(1), 34–39.

Wati, D. C., & Utami, H. (2020). Model Geographically Weighted Panel Regression (Gwpr) Dengan Fungsi Kernel Fixed Gaussian Pada Indeks Pembangunan Manusia Di Jawa Timur. Jurnal Matematika Thales, 2(1), 78–97. https://doi.org/10.22146/jmt.49230