

Analisis Faktor-Faktor Yang Berpengaruh Terhadap Pencemaran Air Sungai Mahakam Menggunakan Pemodelan *Geographically Weighted Logistic Regression* Pada Data *Dissolved Oxygen*

Analysis of Factors That Influencing the Mahakam River Water Pollution By Using Geographically Weighted Logistic Regression On Dissolved Oxygen Data

Vivi Dwi Lestari¹, Suyitno² dan Meiliyani Siringoringo³

^{1,2}Laboratorium Statistika Terapan FMIPA Universitas Mulawarman

³Laboratorium Statistika Komputasi FMIPA Universitas Mulawarman

Email: vividwilestari2@gmail.com

Abstract

Geographically Weighted Logistic Regression (GWLR) model is a local model of the logistic regression model applied to spatial data. Parameter estimation is performed at each observation location using spatial weighting. The spatial weighting is calculated by using an adaptive tricube function and bandwidth optimum was obtained based on Generalized Cross Validation (GCV) criteria. The purpose of this study was to obtain a GWLR model on the water pollution indicator Dissolve Oxygen (DO) in Mahakam River in East Kalimantan Province and to find factors affecting the probability of the Mahakam River water was not polluted based on DO indicator. The research data is secondary obtained from Environmental Department of East Kalimantan. The parameter estimation method was Maximum Likelihood Estimation (MLE). The research result showed that the closed form of ML estimator could not be found analytically and it can be approximated by using Newton-Raphson iterative methods. Based on the result of partial hypothesis test, the factors influencing the probability of the Mahakam River water was not polluted is different for every observation location. They were phosphate concentration, total dissolved solid and nitrite concentration. The factor influencing globally was total dissolved solid.

Keywords: *Adaptive Tricube, DO, GWLR, MLE, Newton-Raphson*

Pendahuluan

Analisis regresi adalah analisis yang dapat menjelaskan hubungan antara variabel respon dan variabel prediktor. Analisis regresi dapat diaplikasikan di berbagai bidang seperti bidang ekonomi, sosial, kesehatan dan lingkungan. Data pada bidang tersebut sering ditemukan berupa data spasial. Data spasial adalah data yang mengandung informasi atribut dan lokasi (Fotheringham dkk, 2002).

Data respon lapangan sering ditemukan berupa data respon spasial dan bersifat kualitatif atau kuantitatif yang dapat dikategorikan. Data respon bersifat dikotomis artinya data berskala nominal dengan dua kategori yaitu bernilai 1 atau 0, dimana nilai 1 menyatakan hasil percobaan sukses dan nilai 0 menyatakan hasil percobaan gagal. Salah satu model regresi yang dapat menjelaskan hubungan antara variabel respon dikotomis kategorik dengan variabel prediktor adalah regresi logistik biner . (Hosmer dkk, 2013).

Pemodelan regresi logistik biner pada data respon dikotomis yang memuat heterogenitas spasial adalah *Geographically Weighted Logistic*

Regression (GWLR). GWLR merupakan pengembangan dari model GWR dengan data variabel respon bersifat kategorik. Pemodelan GWLR merupakan model lokal dari regresi logistik dengan memperhatikan faktor lokasi. (Hosmer dkk, 2013). Model GWLR pada penelitian ini akan diaplikasikan pada data *Dissolved Oxygen (DO)* di Daerah Aliran Sungai Mahakam Kalimantan Timur.

Sungai Mahakam merupakan sungai yang penting untuk menunjang kebutuhan manusia sekitarnya yang terletak di Provinsi Kalimantan Timur. Aktivitas-aktivitas di sepanjang DAS Mahakam antara lain aktivitas rumah tangga, rumah makan, peternakan, perikanan, perkebunan, industri, transportasi dan pelabuhan bongkar muatan (Susilowati dkk, 2012). Aktivitas tersebut berpotensi menghasilkan limbah pada aliran air sungai. Limbah tersebut merupakan ancaman air Sungai Mahakam menjadi tercemar, oleh karenanya perlu tindakan pencegahan melalui indikator pencemaran air. Salah satu indikator untuk mendeteksi air sungai tercemar adalah *Dissolve Oxygen (DO)*. DO adalah kadar oksigen terlarut dalam air yang diperlukan untuk

metabolisme semua organisme perairan (Salmin, 2005).

Tujuan penelitian ini adalah mendapatkan model GWLR data DO di DAS Mahakam dan mendapatkan faktor-faktor yang berpengaruh terhadap pencemaran air di DAS Mahakam serta menginterpretasikan hasil yang diperoleh. Pemodelan GWLR lebih baik digunakan untuk memodelkan data DO di DAS Mahakam daripada model regresi logistik biasa karena penaksiran parameter dilakukan di setiap lokasi pengamatan (Suyitno dkk, 2016).

Penelitian sebelumnya antara lain oleh Suyitno, Purnadi, Sutikno dan Irhamah (2016) yaitu *Parameter Estimation of Geographically Weighted Trivariate Weibull Regression Model* menunjukkan bahwa model GWTWR lebih baik daripada model global berdasarkan nilai GCV. Penelitian oleh Pravitasary, Hajarisman dan Sunendiari (2015) yaitu *Pemodelan Faktor-Faktor yang Berpengaruh Terhadap Angka Buta Huruf di Provinsi Jawa Barat dengan Geographically Weighted Logistic Regression* menunjukkan bahwa model GWLR lebih baik daripada model regresi logistik global berdasarkan nilai AIC yang diperoleh.

Distribusi Bernoulli

Distribusi Bernoulli yang dikenalkan oleh James Bernoulli (1654-1705) menggambarkan percobaan probabilistik, dimana percobaan hanya memiliki dua hasil yang mungkin yaitu sukses dan gagal. Fungsi probabilitas peubah acak Y yang berdistribusi Bernoulli dengan parameter π , yaitu

$$P(Y = y) = f(y; \pi) = \pi^y (1 - \pi)^{1-y}; y = 0, 1 \quad (1)$$

dengan $y = 1$ menyatakan hasil percobaan sukses dan $y = 0$ menyatakan hasil percobaan gagal. $P(Y = 1) = \pi$ menyatakan probabilitas sukses dan $P(Y = 0) = 1 - \pi$ menyatakan probabilitas gagal (Nelson dkk, 2003).

Salah satu metode penaksiran parameter distribusi Bernoulli adalah metode *Maximum Likelihood Estimation* (MLE). Metode MLE adalah metode untuk mencari nilai penaksir parameter dengan memaksimalkan fungsi *likelihood* (Hosmer dkk, 2013).

Fungsi *likelihood* didefinisikan oleh

$$L(\pi) = \prod_{i=1}^n f(y_i; \pi) = \pi^{\sum_{i=1}^n y_i} (1 - \pi)^{n - \sum_{i=1}^n y_i} \quad (2)$$

Penaksir distribusi Bernoulli lebih mudah diperoleh dengan memaksimalkan fungsi *log-likelihood*.

Fungsi *log-likelihood* diperoleh dengan menerapkan logaritma natural pada kedua ruas persamaan (5) dan diperoleh

$$\ell(\pi) = \left(\sum_{i=1}^n y_i \right) \ln(\pi) + \left(n - \sum_{i=1}^n y_i \right) \ln(1 - \pi) \quad (3)$$

Penaksir π yang memaksimalkan fungsi *log-likelihood* $\ell(\pi)$ diperoleh dari turunan pertama persamaan (3) terhadap parameter π dan disamakan dengan nol

$$\frac{\partial \ell(\pi)}{\partial(\pi)} = \sum_{i=1}^n y_i - n\pi = 0 \quad (4)$$

Penaksir π diperoleh dengan menyelesaikan persamaan (4) dan didapat

$$\hat{\pi} = \frac{\sum_{i=1}^n y_i}{n} = \bar{Y} \quad (5)$$

(Hosmer dkk, 2013).

Regresi Logistik Biner

Model regresi logistik biner merupakan salah satu model regresi yang digunakan untuk mengetahui hubungan variabel respon yang bersifat dikotomis dengan satu atau lebih variabel prediktor kontinu atau kategorik, atau gabungan dari kontinu dan kategorik (Agresti, 2002).

Model regresi logistik dengan k variabel bebas dapat ditulis sebagai berikut (Hosmer dkk, 2013)

$$\pi_i = \frac{\exp(\beta^T \mathbf{x}_i)}{1 + \exp(\beta^T \mathbf{x}_i)} \quad (6)$$

Penaksiran parameter model regresi logistik biner menggunakan metode MLE. Fungsi *likelihood* berdasarkan n pengamatan $y_i \sim B(1, \pi_i)$, $i = 1, 2, \dots, n$ didefinisikan oleh (Agresti, 2002) :

$$L(\beta) = \left(\prod_{i=1}^n (1 + \exp(\beta^T \mathbf{x}_i))^{-1} \right) \exp \left(\sum_{i=1}^n y_i \beta^T \mathbf{x}_i \right) \quad (7)$$

Penaksir ML model regresi logistik biner adalah nilai vektor $\hat{\beta} = [\hat{\beta}_0 \ \hat{\beta}_1 \ \hat{\beta}_2 \ \dots \ \hat{\beta}_p]^T$ yang memaksimalkan fungsi *likelihood* $L(\beta)$ dan juga fungsi *log-likelihood*, tetapi lebih mudah diperoleh dengan memaksimalkan fungsi *log-likelihood*. Fungsi *log-likelihood* diperoleh dengan menerapkan logaritma natural pada kedua ruas persamaan (7) dan diperoleh

$$\ell(\beta) = \sum_{i=1}^n y_i \beta^T x_i - \ln(1 + \exp(\beta^T x_i)) \quad (8)$$

Penaksir β yang memaksimumkan fungsi *log-likelihood* diperoleh dengan melakukan diferensiasi fungsi $\ell(\beta)$ terhadap parameter β dan disamakan dengan nol, yaitu

$$\frac{\partial \ell(\beta)}{\partial \beta} = \mathbf{0} \quad (9)$$

dimana ruas kiri persamaan (9) adalah vektor gradien berukuran $(p+1)$ dan $\mathbf{0}$ adalah vektor nol berukuran $(p+1)$. Vektor gradien diberikan oleh

$$\mathbf{g}(\beta) = \mathbf{X}^T (\mathbf{y} - \boldsymbol{\pi}) \quad (10)$$

Berdasarkan persamaan (10), bahwa persamaan (9) terdiri dari persamaan-persamaan *nonlinier*, sehingga solusi eksak (*closed form*) untuk mendapatkan penaksir ML eksak tidak dapat diperoleh secara analitik. Metode alternatif untuk menentukan solusi persamaan (9) adalah metode iteratif Newton-Raphson (Agresti, 2002).

Algoritma iterasi Newton-Raphson diberikan oleh

$$\hat{\beta}^{(q+1)} = \hat{\beta}^{(q)} - [\mathbf{H}(\hat{\beta}^{(q)})]^{-1} \mathbf{g}(\hat{\beta}^{(q)}) \quad (11)$$

dengan $\mathbf{g}(\beta)$ diberikan oleh persamaan (10) dan \mathbf{H} adalah Matriks Hessian yang diberikan oleh

$$\mathbf{H}(\hat{\beta}) = -\mathbf{X}^T \mathbf{V} \mathbf{X} \quad (12)$$

Proses iterasi dimulai dari penentuan nilai awal $\hat{\beta}^{(0)} = [\hat{\beta}_0^{(0)} \hat{\beta}_1^{(0)} \dots \hat{\beta}_p^{(0)}]^T$ dan iterasi akan berhenti apabila $\|\hat{\beta}^{(q+1)} - \hat{\beta}^{(q)}\| \leq \varepsilon$ dengan ε merupakan bilangan yang sangat kecil (Agresti, 2002).

Negatif nilai ekspektasi dari matriks Hessian merupakan matriks Informasi Fisher dan invers dari matriks Informasi Fisher adalah penaksir dari matriks varians kovarians (Fathurahman dkk, 2016).

$$\text{cov}(\hat{\beta}) = [\mathbf{I}(\hat{\beta})]^{-1} = -[\mathbf{H}(\hat{\beta})]^{-1} \quad (13)$$

Pengujian Parameter Regresi Logistik Biner Secara Serentak

Hipotesis pengujian serentak adalah sebagai berikut

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{Minimal ada satu } \beta_k \neq 0, k = 1, 2, \dots, p$$

Statistik uji diberikan oleh

$$G = 2 \left(\ell(\hat{\Omega}) - \ell(\hat{\omega}) \right) = 2 \left(\sum_{i=1}^n (y_i \ln(\hat{\pi}_i) + (1 - y_i) \ln(1 - \hat{\pi}_i)) - n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n) \right) \quad (14)$$

dengan $G \sim \chi^2_p$ (Agresti, 2002). Hipotesis nol ditolak pada taraf uji α apabila nilai $G > \chi^2_{(\alpha, p)}$ atau $p_{value} < \alpha$ (Hosmer dan Lameshow, 2000).

Pengujian Parameter Regresi Logistik Biner Secara Parsial

Pengujian signifikansi parameter secara parsial menggunakan uji Wald. Hipotesis pengujian parameter secara parsial untuk parameter β_k dengan nilai k tertentu ($k = 0, 1, 2, \dots, p$) adalah

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0$$

Statistik uji diberikan oleh

$$W = \frac{\hat{\beta}_k}{\sqrt{\text{Var}(\hat{\beta}_k)}} \quad (15)$$

dimana $\text{Var}(\hat{\beta}_k)$ adalah elemen diagonal ke- k dari varian kovarians (13). Statistik uji Wald yang diberikan oleh persamaan (15) mengikuti distribusi $N(0,1)$. Hipotesis nol ditolak pada taraf uji α apabila nilai $|W| > Z_{\alpha/2}$ atau $p_{value} < \alpha$ (Hosmer dan Lameshow, 2000).

Pendeteksian Multikolinieritas

Menurut Gujarati (2003), multikolinieritas adalah terjadinya hubungan linier antar variabel prediktor dalam suatu model regresi linier. Cara mendeteksi multikolinieritas salah satunya dengan melihat nilai VIF (*Varians Inflation Factor*). Nilai VIF lebih besar dari 10 menunjukkan adanya multikolinieritas. Nilai VIF dihitung berdasarkan formula berikut (DeMaris, 2004)

$$VIF_k = \frac{1}{1 - R_k^2} \quad (16)$$

dengan R_k^2 adalah koefisien determinasi antara variabel prediktor X_k dengan variabel prediktor lainnya (Hocking, 1996).

Pengujian Heterogenitas Spasial

Salah satu metode yang dapat digunakan untuk pengujian heterogenitas spasial adalah uji

Glejser. Hipotesis yang digunakan adalah

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, n$$

Statistik uji diberikan oleh

$$F_2 = \frac{(\hat{\alpha} \mathbf{X}^T \mathbf{e} - n \bar{e}^2) / p}{(\mathbf{e}^T \mathbf{e} - \hat{\alpha} \mathbf{X}^T \mathbf{e}) / n - p - 1} \quad (17)$$

dimana \mathbf{e} adalah vektor berisi nilai mutlak residual model regresi logistik biner berdimensi n dengan elemen-elemen $|\hat{e}_1|, |\hat{e}_2|, \dots, |\hat{e}_n|$ dan $\hat{\alpha}$ adalah vektor penaksir parameter dalam pemodelan regresi antara $|\hat{e}_i|$ terhadap variabel prediktor. Hipotesis nol ditolak pada taraf uji α apabila nilai $F_{hitung} > F_{\alpha; p, (n-p-1)}$ atau $P_{value} < \alpha$ (Fotheringham dkk, 2002).

Pembobot Spasial Pada Model Geographically Weighted Regression

Pembobot spasial dihitung menggunakan fungsi pembobot yang bergantung pada jarak antar lokasi pengamatan dan *bandwidth*. Jarak antar lokasi pengamatan dihitung berdasarkan titik koordinat lokasi. Jarak lokasi (u_i, v_i) dengan lokasi (u_j, v_j) disimbolkan dengan d_{ij} dan dihitung menggunakan jarak *Euclidean* (Chasco dkk, 2007)

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (18)$$

dengan u_i menyatakan letak garis lintang (*latitude*) dan v_i menyatakan letak garis bujur (*longitude*) (Chasco dkk, 2007).

Salah satu fungsi pembobot adalah fungsi pembobot *Tricube*. Fungsi pembobot *Adaptive Tricube* diberikan oleh persamaan berikut (Fotheringham dkk, 2002) :

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^3\right)^3, & \text{jika } d_{ij} \leq b_i \\ 0, & \text{jika } d_{ij} > b_i \end{cases} \quad (19)$$

dengan w_{ij} adalah bobot spasial yang diberikan kepada pengamatan ke- j untuk penaksiran model GWLR pada lokasi ke- i dan b_i adalah *bandwidth* untuk penaksiran model GWLR pada lokasi ke- i (Fotheringham dkk, 2002).

Salah satu metode untuk menentukan *bandwidth* optimum adalah metode *Generalized Cross Validation* (GCV). *Bandwidth* optimum

akan diperoleh saat nilai GCV minimum. Nilai GCV diperoleh dengan rumus sebagai berikut (Fotheringham dkk, 2002) :

$$GCV = \frac{n \sum_{i=1}^n \hat{e}_i^2}{(n-v)^2} \quad (20)$$

dengan $\hat{e}_i = \hat{\pi}(u_i, v_i) - \bar{Y}$ dan $v = \text{tr}(\mathbf{S})$, dimana

$$\mathbf{S} = \begin{bmatrix} \mathbf{x}_1^T (\mathbf{X}^T \mathbf{W}(u_1, v_1) \mathbf{V}(u_1, v_1) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_1, v_1) \mathbf{V}(u_1, v_1) \\ \mathbf{x}_2^T (\mathbf{X}^T \mathbf{W}(u_2, v_2) \mathbf{V}(u_2, v_2) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_2, v_2) \mathbf{V}(u_2, v_2) \\ \vdots \\ \mathbf{x}_n^T (\mathbf{X}^T \mathbf{W}(u_n, v_n) \mathbf{V}(u_n, v_n) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_n, v_n) \mathbf{V}(u_n, v_n) \end{bmatrix}$$

Model Geographically Weighted Logistic Regression

Model *Geographically Weighted Logistic Regression* (GWLR) merupakan pengembangan dari model GWR dengan data variabel respon bersifat dikotomus. Pemodelan GWLR merupakan model lokal dari regresi logistik pada persamaan (6) dengan memperhatikan faktor lokasi (Atkinson dkk, 2003).

Misalkan koordinat lokasi setiap pengamatan diketahui, maka model umum GWLR pada lokasi (u_i, v_i) adalah

$$\pi(u_i, v_i) = \frac{\exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i)}{1 + \exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i)} \quad (21)$$

Salah satu metode penaksiran parameter model GWLR adalah MLE. Fungsi *likelihood* dengan pembobot spasial untuk penksiran parameter GWLR pada lokasi ke- i adalah

$$L(\boldsymbol{\beta}(u_i, v_i)) = \prod_{j=1}^n (\pi_j^{y_j} (1 - \pi_j)^{1-y_j})^{w_{ij}} = \left(\exp \left(\sum_{j=1}^n y_j \boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_j \right) \left(\prod_{j=1}^n (1 + \exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_j))^{-1} \right) \right)^{w_{ij}} \quad (22)$$

Penaksir *Maximum Likelihood* model GWLR pada lokasi (u_i, v_i) adalah nilai vektor $\hat{\boldsymbol{\beta}}(u_i, v_i)$ yang memaksimumkan fungsi *likelihood* dan juga fungsi *log-likelihood*. Fungsi *log-likelihood* diperoleh dengan menerapkan logaritma natural pada kedua ruas persamaan (22) dan diperoleh

$$\ell(\boldsymbol{\beta}(u_i, v_i)) = \sum_{j=1}^n w_{ij} \left(y_j \boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_j - \ln(1 + \exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_j)) \right) \quad (23)$$

Penaksir $\boldsymbol{\beta}(u_i, v_i)$ yang memaksimumkan fungsi *log-likelihood* diperoleh dengan melakukan

diferensiasi fungsi $\ell(\boldsymbol{\beta}(u_i, v_i))$ terhadap parameter $\beta_k(u_i, v_i)$ dan disamakan dengan nol, yaitu

$$\frac{\partial \ell(\boldsymbol{\beta}(u_i, v_i))}{\partial (\boldsymbol{\beta}(u_i, v_i))} = \mathbf{0} \quad (24)$$

dimana ruas kiri persamaan (24) adalah vektor gradien berukuran $p+1$ dan $\mathbf{0}$ adalah vektor nol berukuran $p+1$. Vektor gradien diberikan oleh

$$\mathbf{g}(\boldsymbol{\beta}(u_i, v_i)) = \mathbf{X}^T \mathbf{W}(u_i, v_i) (\mathbf{y} - \boldsymbol{\pi}(u_i, v_i)) \quad (25)$$

Berdasarkan persamaan (25), persamaan (24) terdiri dari persamaan-persamaan *nonlinier*, sehingga solusi eksak (*closed form*) untuk mendapatkan penaksir ML eksak tidak dapat diperoleh secara analitik. Metode alternatif untuk menentukan solusi persamaan (2.24) adalah metode iteratif Newton-Raphson (Fathurahman dkk, 2016).

Algoritma iterasi Newton-Raphson diberikan oleh

$$\hat{\boldsymbol{\beta}}^{(q+1)}(u_i, v_i) = \hat{\boldsymbol{\beta}}^{(q)}(u_i, v_i) - [\mathbf{H}(\hat{\boldsymbol{\beta}}^{(q)}(u_i, v_i))]^{-1} \mathbf{g}(\hat{\boldsymbol{\beta}}^{(q)}(u_i, v_i)) \quad (26)$$

dengan $\mathbf{g}(\boldsymbol{\beta}(u_i, v_i))$ diberikan oleh persamaan (25) dan $\mathbf{H}(\boldsymbol{\beta}(u_i, v_i))$ adalah Matriks Hessian yang diberikan oleh

$$\mathbf{H}(\boldsymbol{\beta}(u_i, v_i)) = -\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{V}(u_i, v_i) \mathbf{X} \quad (27)$$

Proses iterasi dimulai dari penentuan nilai awal $\boldsymbol{\beta}^{(0)}(u_i, v_i) = [\beta_0^{(0)}(u_i, v_i) \beta_1^{(0)}(u_i, v_i) \dots \beta_p^{(0)}(u_i, v_i)]^T$ dan akan berhenti apabila $\|\boldsymbol{\beta}^{(q+1)}(u_i, v_i) - \boldsymbol{\beta}^{(q)}(u_i, v_i)\| \leq \varepsilon$ dengan ε merupakan bilangan yang sangat kecil. Prosedur iterasi ini diulang untuk setiap lokasi ke- i , sehingga dapat diperoleh penaksir parameter lokal model GWLR (Fathurahman dkk, 2016).

Matriks Informasi Fisher lokasi (u_i, v_i) dapat dinyatakan (Fotheringham dkk, 2002) :

$$\mathbf{I}(\hat{\boldsymbol{\beta}}(u_i, v_i)) = \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{V}(u_i, v_i) \mathbf{X} \quad (28)$$

Pengujian Kesesuaian Model Regresi Logistik Biner dan Model GWLR

Hipotesis pengujian kesesuaian model regresi logistik dan model GWLR adalah

$$H_0: \beta_k(u_1, v_1) = \beta_k(u_2, v_2) = \dots = \beta_k(u_n, v_n) = \beta_k; \quad k = 1, 2, \dots, p$$

(Model regresi logistik biner dan model GWLR identik)

H_1 : Minimal ada satu

$$\beta_k(u_i, v_i) \neq \beta_k; k = 1, 2, \dots, p; i = 1, 2, \dots, n$$

(Model regresi logistik biner dan model GWLR tidak identik)

Statistik uji adalah

$$F_2 = \frac{D(\hat{\boldsymbol{\beta}}) / v_1}{D(\hat{\boldsymbol{\beta}}^*) / v_2} \quad (29)$$

dengan

$$D(\hat{\boldsymbol{\beta}}) = 2 \left(\sum_{i=1}^n (y_i \ln(\hat{\pi}_i) + (1 - y_i) \ln(1 - \hat{\pi}_i)) - n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n) \right)$$

dan

$$D(\hat{\boldsymbol{\beta}}^*) = 2 \left(\sum_{i=1}^n (y_i \ln \hat{\pi}_i(u_i, v_i) + (1 - y_i) \ln(1 - \hat{\pi}_i(u_i, v_i))) - n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n) \right)$$

Statistik $F_2 \sim F_{(p, np)}$ dimana n adalah banyaknya pengamatan dan p adalah banyaknya variabel prediktor. Hipotesis nol ditolak pada taraf uji α apabila nilai $F_2 > F_{\alpha; p, np}$ atau $p_{value} < \alpha$ (Hosmer dan Lameshow, 2000).

Pengujian Parameter Model GWLR Secara Serentak

Hipotesis pengujian serentak adalah sebagai berikut

$$H_0: \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_p(u_i, v_i) = 0, \quad i = 1, 2, \dots, n$$

H_1 : Minimal ada satu

$$\beta_k(u_i, v_i) \neq \beta_k; k = 1, 2, \dots, p; i = 1, 2, \dots, n$$

Statistik uji diberikan oleh

$$G_2 = 2 \left(\ell(\hat{\boldsymbol{\Omega}}_{GWLR}) - \ell(\hat{\omega}_{GWLR}) \right) = 2 \left(\sum_{i=1}^n (y_i \ln \hat{\pi}_i(u_i, v_i) + (1 - y_i) \ln(1 - \hat{\pi}_i(u_i, v_i))) - \sum_{i=1}^n n_{1i} \ln(n_{1i}) + n_{0i} \ln(n_{0i}) - n \ln(n) \right) \quad (30)$$

dengan $G_2 \sim \chi_{(\alpha, v)}^2$. Hipotesis nol ditolak pada taraf uji α apabila $G_2 > \chi_{(\alpha, v)}^2$ atau $p_{value} < \alpha$ (Hosmer dan Lameshow, 2000).

Pengujian Parameter Model GWLR Secara Parsial

Hipotesis pengujian parameter secara parsial untuk parameter β_k dengan nilai k tertentu ($k = 0, 1, 2, \dots, p$) dan i tertentu ($i = 1, 2, \dots, p$) adalah

$$H_0: \beta_k(u_i, v_i) = 0$$

$$H_1: \beta_k(u_i, v_i) \neq 0$$

Statistik uji diberikan oleh

$$W_h = \frac{\hat{\beta}_k(u_i, v_i)}{\sqrt{\text{Var}(\hat{\beta}_k(u_i, v_i))}} \quad (31)$$

dimana $\text{Var}(\hat{\beta}_k(u_i, v_i))$ adalah elemen diagonal ke- k dari varians kovarians (13). Statistik uji Wald yang diberikan oleh persamaan (31) mengikuti distribusi $N(0,1)$ untuk ukuran sampel (n) yang cukup besar. Hipotesis nol ditolak pada taraf uji α apabila nilai $|W_h| > Z_{\alpha/2}$ atau $p_{value} < \alpha$ (Hosmer dan Lameshow, 2000).

Ketepatan Klasifikasi

Kriteria perbandingan teknik klasifikasi didasarkan pada kesalahan klasifikasi yang dikenal dengan *Apparent Error Rate* (APER) merupakan nilai besar kecilnya jumlah observasi yang salah dalam pengklasifikasian berdasarkan suatu fungsi klasifikasi (Johnson & Wichen, 2007).

Tabel 1. Klasifikasi Aktual dan Prediksi

Kelas Aktual	Kelas Prediksi	
	$y = 0$	$y = 1$
$y = 0$	n_{00}	$n_{01} = n_0 - n_{00}$
$y = 1$	$n_{10} = n_1 - n_{11}$	n_{11}

$$APER = \frac{n_{00} + n_{11}}{n_0 + n_1} \times 100\% \quad (32)$$

Interpretasi Model GWLR

Interpretasi parameter pada model GWLR yaitu menggunakan *odds ratio* (Harlan, 2018).

$$OR(u_i, v_i) = \frac{\frac{\pi(u_i, v_i)(1)}{1 - \pi(u_i, v_i)(1)}}{\frac{\pi(u_i, v_i)(0)}{1 - \pi(u_i, v_i)(0)}} = \exp(\beta_1(u_i, v_i)) \quad (33)$$

Penaksiran parameter model GWLR yang dilakukan pada setiap lokasi pengamatan menyebabkan interpretasi parameter juga dilakukan pada setiap lokasi pengamatan.

Dissolve Oxygen (DO)

Oksigen terlarut atau *Dissolve Oxygen* (DO) adalah jumlah oksigen terlarut dalam air yang dinyatakan dalam mg/l (miligram per liter) atau ppm (*part per million*) yang diperlukan untuk metabolisme semua organisme perairan. DO merupakan indikator kualitas perairan. DO berasal dari proses fotosintesis tumbuhan air dan difusi melalui udara dengan kecepatan yang lambat (Salmin, 2005).

Hasil dan Pembahasan

Data penelitian terdiri dari data variabel respon, variabel prediktor dan data koordinat lokasi pengamatan. Data variabel respon adalah data DO di DAS Mahakam. Data variabel respon dikategorikan menjadi 2 kategori (dikotomus) yaitu $y = 0$ jika $DO < 6$ mg/l atau diindikasikan air Sungai Mahakam tercemar dan $y = 1$ jika $DO \geq 6$ atau diindikasikan tidak tercemar (PP No. 82 Tahun 2001).

Data variabel prediktor terdiri dari data konsentrasi amonia (NH_3), konsentrasi fosfat (PO_4), Daya Hantar Listrik (DHL), *Total Dissolved Solid* (TDS), konsentrasi nitrit (NO_3) dan konsentrasi sulfat (SO_4). Data koordinat 26 lokasi pengamatan adalah letak garis lintang dan bujur yang telah ditetapkan oleh Dinas Lingkungan Hidup Provinsi Kalimantan Timur tahun 2018.

Deskripsi Data Penelitian

Deskripsi data dinyatakan dalam statistik deskriptif yang meliputi rata-rata, nilai maksimum, nilai minimum dan standar deviasi. Statistik deskriptif data DO disajikan pada Tabel 2.

Tabel 2. Statistik Deskriptif Data DO

Mean	Maks	Min	SD	PTT (%)
6,02	8,05	3,63	1,03	57,69

Ket PTT : Persentase Tidak Tercemar

SD : Standar Deviasi

Berdasarkan Tabel 1 bahwa persentase lokasi pengamatan terindikasi tidak tercemar sebesar 57,69% atau 15 dari 26 lokasi pengamatan. Rata-rata DO adalah sebesar 6,02 mg/l yang mengindikasikan air Sungai Mahakam tidak tercemar. DO tertinggi adalah 8,05 mg/l, yaitu di lokasi pengamatan Sungai Mahakam Nyan. DO terendah adalah 3,63 mg/l, yaitu di lokasi pengamatan Karang Mumus.

Pendeteksian Multikolinieritas

Pendeteksian multikolinieritas menggunakan kriteria VIF. Hasil nilai VIF masing-masing variabel prediktor disajikan pada Tabel 3. Berdasarkan Tabel 3, nilai VIF setiap variabel prediktor kurang dari 10 sehingga disimpulkan bahwa tidak terdapat multikolinieritas antar variabel prediktor dan keenam variabel dapat dimasukan ke dalam model GWLR.

Tabel 3. Nilai VIF Setiap Variabel Prediktor

Variabel	VIF
X_1	1,0819
X_2	1,4399
X_3	1,1169
X_4	2,3636
X_5	2,0378
X_6	1,1295

Model Regresi Logistik Biner

Penaksiran parameter model regresi logistik biner global menggunakan metode MLE berdasarkan persamaan (2.11), dan diperoleh model regresi logistik biner global yaitu

$$\hat{\pi} = \frac{\exp\left(\begin{matrix} 2,0405 - 20,1538X_1 - 29,9452X_2 + 0,0222X_3 \\ -0,1262X_4 + 252,9913X_5 - 0,3858X_6 \end{matrix}\right)}{1 + \exp\left(\begin{matrix} 2,0405 - 20,1538X_1 - 29,9452X_2 + 0,0222X_3 \\ -0,1262X_4 + 252,9913X_5 - 0,3858X_6 \end{matrix}\right)}$$

dengan $\hat{\pi}$ menyatakan taksiran probabilitas air Sungai Mahakam tidak tercemar berdasarkan indikator DO. X_1 menyatakan konsentrasi amonia, X_2 menyatakan konsentrasi fosfat, X_3 menyatakan DHL, X_4 menyatakan Total Dissolved Solid (TDS), X_5 menyatakan konsentrasi nitrit dan X_6 menyatakan konsentrasi sulfat. Nilai GCV model regresi logistik biner sebesar 0,2214

Pengujian Parameter Secara Serentak

Hipotesis pengujian parameter secara serentak adalah

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$$

$$H_1 : \text{Minimal ada satu } \beta_k \neq 0 ; k = 1, 2, 3, 4, 5, 6$$

Hasil perhitungan uji parameter secara serentak model regresi logistik biner global diperoleh nilai G_{hitung} sebesar 15,1523 dan p -value sebesar 0,0191 sehingga diputuskan menolak H_0 dan disimpulkan bahwa variabel-variabel prediktor secara bersama-sama berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar.

Pengujian Parameter Secara Parsial

Hipotesis pengujian parameter secara parsial untuk parameter β_k dengan nilai k tertentu ($k = 0, 1, 2, 3, 4, 5, 6$) adalah

$$H_0 : \beta_k = 0$$

(Variabel prediktor X_k tidak berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar)

$$H_1 : \beta_k \neq 0$$

(Variabel prediktor X_k berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar)

Hasil perhitungan uji parameter secara parsial model regresi logistik biner global dapat dilihat pada Tabel 4.

Tabel 4. Hasil Statistik Uji Parameter Regresi Logistik Biner Secara Parsial

Variabel	W_{hitung}	p -value	Keputusan Uji
X_0	0,6541	0,5131	H_0 gagal ditolak
X_1	-0,8980	0,3692	H_0 gagal ditolak
X_2	-1,7130	0,0867	H_0 ditolak
X_3	0,7908	0,4290	H_0 gagal ditolak
X_4	0,0569	0,0268	H_0 ditolak
X_5	1,9655	0,0494	H_0 ditolak
X_6	-0,9816	0,3263	H_0 gagal ditolak

Berdasarkan Tabel 4, diperoleh variabel konsentrasi fosfat (X_2), TDS (X_4) dan konsentrasi nitrit (X_5) berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar. Nilai statistik uji W variabel-variabel tersebut lebih dari 1,64 dan p -value lebih kecil dari 0,1. Variabel-variabel konsentrasi amonia (X_1), DHL (X_3) dan konsentrasi sulfat (X_6) masing-masing secara individual tidak berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar.

Pengujian Heterogenitas Spasial

Pengujian heterogenitas spasial dilakukan dengan metode Glejser. Hipotesis pengujian heterogenitas spasial adalah

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{26}^2 = \sigma^2$$

(Tidak terdapat heterogenitas spasial)

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, 26$$

(Terdapat heterogenitas spasial)

Hasil perhitungan statistik uji persamaan (17) diperoleh nilai F_{hitung} sebesar 3,3830 dan p -value sebesar 0,0193 sehingga diputuskan menolak H_0 dan disimpulkan terdapat heterogenitas spasial pada data respon dikotomus.

Model Geographically Weighted Logistic Regression

Langkah pertama dalam pemodelan GWLR adalah mencari jarak antar lokasi pengamatan menggunakan persamaan (18). Tahapan selanjutnya adalah menghitung pembobot dengan

menggunakan fungsi *adaptive tricube* berdasarkan persamaan (19). *Bandwidth* optimum dihitung menggunakan kriteria GCV yang diberikan oleh persamaan (20).

Penaksiran parameter model GWLR menggunakan metode MLE berdasarkan persamaan (26) dan diperoleh penaksir model GWLR yang disajikan pada Tabel 5.

Tabel 5. Penaksir Parameter Model GWLR

<i>i</i>	β_0	β_1	β_2	β_3	β_4	β_5	β_6
1	1,57	-22,15	-44,56	0,03	-0,13	257,36	-0,35
2	3,61	-17,50	-36,06	0,01	-0,14	245,37	-0,56
3	2,04	-20,15	-29,94	0,02	-0,12	252,99	-0,39
4	4,24	-16,69	-34,35	0,01	-0,14	245,53	-0,63
5	4,05	-16,93	-34,40	0,02	-0,14	244,96	-0,61
6	4,83	-16,92	-34,70	0,01	-0,14	254,15	-0,69
7	0,97	-20,79	-36,36	0,03	-0,12	232,60	-0,30
8	0,99	-20,82	-36,31	0,03	-0,12	233,19	-0,30
9	1,18	-21,40	-38,83	0,03	-0,13	242,12	-0,32
10	1,35	-21,03	-33,36	0,03	-0,13	244,10	-0,33
11	5,01	-18,98	-34,15	0,01	-0,14	262,69	-0,69
12	5,64	-18,98	-35,72	0,01	-0,14	269,69	-0,75
13	5,49	-18,98	-35,25	0,01	-0,14	267,02	-0,74
14	2,04	-20,15	-29,95	0,02	-0,13	252,99	-0,39
15	1,49	-22,06	-41,31	0,03	-0,13	255,65	-0,34
16	2,04	-20,15	-29,94	0,02	-0,12	252,99	-0,39
17	1,84	-21,23	-36,63	0,03	-0,14	254,96	-0,37
18	1,92	-20,72	-32,86	0,03	-0,13	253,79	-0,38
19	2,07	-20,43	-32,76	0,02	-0,13	253,24	-0,39
20	2,29	-20,49	-38,11	0,03	-0,14	254,29	-0,42
21	2,24	-20,35	-35,12	0,03	-0,14	253,27	-0,41
22	2,04	-20,15	-29,95	0,02	-0,13	252,99	-0,39
23	2,04	-20,15	-29,95	0,02	-0,13	252,99	-0,39
24	2,08	-20,93	-38,31	0,03	-0,14	255,36	-0,40
25	2,25	-20,54	-37,81	0,03	-0,14	254,01	-0,42
26	1,16	-20,72	-32,92	0,03	-0,12	238,73	-0,31

Pengujian Kesesuaian Model GWLR

Hipotesis pengujian kesesuaian model adalah

$$H_0: \beta_k(u_1, v_1) = \beta_k(u_2, v_2) = \dots = \beta_k(u_{26}, v_{26}) = \beta_k ; k = 1, 2, \dots, 6$$

(Model regresi logistik biner dan model GWLR identik)

H₁: Minimal ada satu

$$\beta_k(u_i, v_i) \neq \beta_k ; k = 1, 2, \dots, 6 ; i = 1, 2, \dots, 26$$

(Model regresi logistik biner dan model GWLR tidak identik)

Hasil perhitungan uji kesesuaian model GWLR diperoleh nilai F_{hitung} sebesar 20,3813 dan *p-value* sebesar 0,0000 sehingga diputuskan menolak H₀ dan disimpulkan bahwa model regresi logistik biner berbeda dengan model GWLR.

Pengujian Parameter Model GWLR Secara Serentak

Hipotesis pengujian parameter secara serentak adalah

$$H_0: \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_6(u_i, v_i) = 0 ; i = 1,$$

2, ..., 26

$$H_1: \text{Paling sedikit ada satu } \beta_k(u_i, v_i) \neq 0 ; i = 1, 2, \dots, 26 ; k = 1, 2, 3, 4, 5, 6$$

Hasil perhitungan uji parameter secara serentak model GWLR diperoleh nilai G_2 sebesar 16,5031 dan *p-value* sebesar 0,0357 sehingga diputuskan menolak H₀ dan disimpulkan bahwa variabel-variabel prediktor secara bersama-sama berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar di setiap lokasi pengamatan.

Pengujian Parameter Model GWLR Secara Parsial

Pengujian parameter secara parsial bertujuan untuk mengetahui faktor-faktor yang berpengaruh terhadap DO di setiap lokasi pengamatan di Daerah Aliran Sungai Mahakam. Misalkan dilakukan pengujian parsial di lokasi ke-2 yaitu Pulau Kumala. Hipotesis pengujian parameter secara parsial untuk parameter β_k dengan nilai *k* tertentu ($k = 0, 1, 2, \dots, 6$) adalah

$$H_0: \beta_k(u_2, v_2) = 0$$

$$H_1: \beta_k(u_2, v_2) \neq 0$$

Hasil perhitungan uji parsial parameter model GWLR di lokasi pengamatan Pulau Kumala dapat dilihat pada Tabel 6.

Tabel 6. Nilai Statistik Uji Parsial di Pulau Kumala

Variabel	W_{hitung}	<i>p-value</i>
X_0	0,8773	0,3803
X_1	-0,5658	0,5715
X_2	-1,5023	0,1330
X_3	0,6225	0,5336
X_4	-2,0251	0,0429*
X_5	1,6926	0,0905*
X_6	-1,0849	0,2779

*) signifikan pada taraf signifikansi 0,1

Berdasarkan Tabel 6, terlihat bahwa nilai $|W_{hitung}|$ untuk variabel X_4 dan X_5 lebih dari nilai $Z_{0,05}$ sebesar 1,645 dan *p-value* kurang dari 0,1. Hal ini menunjukkan bahwa TDS dan konsentrasi nitrit berpengaruh terhadap DO di lokasi Pulau Kumala dan begitu seterusnya untuk lokasi pengamatan yang lainnya.

Berdasarkan hasil pengujian parameter model GWLR seluruh lokasi pengamatan, model GWLR dapat dikelompokkan menjadi 3 berdasarkan

variabel yang berpengaruh dapat dilihat pada Tabel 7.

Tabel 7. Kelompok Model GWLR Berdasarkan Variabel Prediktor yang Berpengaruh

Kelompok	Variabel	Lokasi
1	X_2, X_4 dan X_5	Kalamur, Samarinda, Karang Mumus Hilir, Muara Pahu, Danau Semayang dan Karang Mumus
2	X_4	Sungai Mahakam Nyan, Sungai Mahakam Boh, Jembayan, Tenggarong dan Sungai Karang Mumus Hulu
3	X_4 dan X_5	Bloro, Pulau Kumala, Anggana, Palaran, Long Bagun, Tering, Melak, Batuq, Muara Muntai, Kota Bangun, Sungai Kedang Kepala Hulu, Sungai Kedang Kepala Hilir, Sungai Belayan Hulu, Sungai Belayan Hilir dan Sungai Boh

Berdasarkan Tabel 7, terlihat bahwa variabel konsentrasi fosfat (X_2) berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Kalamur, Samarinda, Karang Mumus Hilir, Muara Pahu, Danau Semayang dan Karang Mumus. Variabel TDS (X_4) berpengaruh global terhadap probabilitas air Sungai Mahakam tidak tercemar di seluruh lokasi pengamatan. Variabel konsentrasi nitrit (X_5) berpengaruh hampir di seluruh lokasi pengamatan, kecuali di Sungai Mahakam Nyan, Sungai Mahakam Boh, Jembayan, Tenggarong dan Karang Mumus Hulu. Variabel yang berpengaruh bersifat lokal adalah konsentrasi fosfat dan konsentrasi nitrit.

Ketepatan Klasifikasi

Ketepatan prediksi nilai respon berdasarkan model GWLR dan nilai aktual dikotomus dapat dilihat pada Tabel 8. Setelah diketahui ketepatan klasifikasi untuk tiap kategori maka ketepatan hasil klasifikasi secara total dapat diketahui dengan menghitung nilai APER. Berdasarkan persamaan (32) maka diperoleh nilai APER sebagai berikut :

$$APER = \frac{1+3}{26} \times 100\% = 15,3846\%$$

Berdasarkan nilai APER tersebut, maka persentase salah klasifikasi sebesar 15,38% dan

ketepatan klasifikasi (KK) berdasarkan pemodelan GWLR adalah sebesar 84,6154%

Tabel 8. Ketepatan Klasifikasi Model GWLR

Aktual (π)	Prediksi ($\hat{\pi}$)	
	Tercemar	Tidak Tercemar
Tercemar	8	3
Tidak Tercemar	1	14

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, maka kesimpulan yang dapat diambil adalah sebagai berikut:

1. Berdasarkan hasil penaksiran parameter pada Tabel 6, salah satu model GWLR kelompok 1 yang menyatakan probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Kalamur, yaitu

$$\hat{\pi}(u_6, v_6) = \frac{\exp\left(\begin{matrix} 4,8257 - 16,92x_{6,1} - 34,695x_{6,2} + 0,0131x_{6,3} \\ -0,1446x_{6,4} + 254,15x_{6,5} - 0,68649x_{6,6} \end{matrix}\right)}{1 + \exp\left(\begin{matrix} 4,8257 - 16,92x_{6,1} - 34,695x_{6,2} + 0,0131x_{6,3} \\ -0,1446x_{6,4} + 254,15x_{6,5} - 0,68649x_{6,6} \end{matrix}\right)}$$

- Salah satu model GWLR kelompok 2 yang menyatakan probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Tenggarong, yaitu

$$\hat{\pi}(u_{12}, v_{12}) = \frac{\exp\left(\begin{matrix} 5,6380 - 18,9830x_{12,1} - 35,7240x_{12,2} + 0,0089x_{12,3} \\ -0,1437x_{12,4} + 269,6900x_{12,5} - 0,7524x_{12,6} \end{matrix}\right)}{1 + \exp\left(\begin{matrix} 5,6380 - 18,9830x_{12,1} - 35,7240x_{12,2} + 0,0089x_{12,3} \\ -0,1437x_{12,4} + 269,6900x_{12,5} - 0,7524x_{12,6} \end{matrix}\right)}$$

- Salah satu model GWLR kelompok 3 yang menyatakan probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Palaran, yaitu

$$\hat{\pi}(u_5, v_5) = \frac{\exp\left(\begin{matrix} 4,0471 - 16,9250x_{5,1} - 34,3970x_{5,2} + 0,0167x_{5,3} \\ -0,1412x_{5,4} + 244,96x_{5,5} - 0,6088x_{5,6} \end{matrix}\right)}{1 + \exp\left(\begin{matrix} 4,0471 - 16,9250x_{5,1} - 34,3970x_{5,2} + 0,0167x_{5,3} \\ -0,1412x_{5,4} + 244,96x_{5,5} - 0,6088x_{5,6} \end{matrix}\right)}$$

2. Faktor-faktor yang mempengaruhi probabilitas air Sungai Mahakam tidak tercemar secara lokal yaitu konsentrasi fosfat dan konsentrasi nitrit, sedangkan TDS berpengaruh secara global. Konsentrasi fosfat berpengaruh terhadap probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Kalamur, Samarinda, Karang Mumus Hilir, Muara Pahu, Danau Semayang dan Karang Mumus. Konsentrasi nitrit berpengaruh hampir di seluruh lokasi pengamatan, kecuali di Sungai Mahakam Nyan, Sungai Mahakam Boh, Jembayan, Tenggarong dan Karang Mumus Hulu.
3. Interpretasi model GWLR yang menyatakan probabilitas air Sungai Mahakam tidak

tercemar di lokasi pengamatan di Kalamur berdasarkan faktor-faktor yang berpengaruh adalah sebagai berikut.

Setiap kenaikan satu mg/l konsentrasi fosfat dan dianggap nilai variabel lainnya tetap akan menurunkan probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Kalamur menjadi 10^{-13} kali. Setiap kenaikan satu mg/l TDS dan dianggap nilai variabel lainnya tetap akan menurunkan probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Kalamur menjadi 0,8653 kali. Setiap kenaikan satu mg/l nitrit dan dianggap nilai variabel lainnya tetap akan meningkatkan probabilitas air Sungai Mahakam tidak tercemar di lokasi pengamatan Kalamur menjadi $2,37 \times 10^{10}$ kali.

Daftar Pustaka

- Agresti, A. (2002). *Categorical Data Analysis 2nd Edition*. New York: John Wiley & Sons.
- Atkinson, P.M., German, S.E., Sear, D.A., dan Clark, M.J. (2003). Exploring The Relationship Between Riverbank Erosion & Geomorphological Controls Using Geographically Weighted Logistic Regression. *Geographically Analysis*. 35(1), 58-82.
- DeMaris, A. (2004). *Regression With Social Data: Modelling Continuous and Limited Response*. Hoboken: John Wiley & Sons.
- Fathurahman, M., Puhadi., Sutikno., dan Ratnasari, V. (2016). Pemodelan Geographically Weighted Logistic Regression pada Indeks Pembangunan Kesehatan Masyarakat di Provinsi Papua. *Prosiding Seminar Nasional MIPA 2016*: 34-42.
- Fotheringham, A.S., Brundson, C., dan Charlton, M.E. (2002). *Geographically Weighted Logistic Regression: The Analysis of Spatially Varying Relationship*. England: John Wiley & Sons.
- Gujarati, D. (2003). *Basic Econometrics 4th Edition*. New York: McGraw-Hill Inc.
- Harlan, Johan. (2018). *Analisis Regresi Logistik*. Depok: Gunadarma.
- Hocking, R. (1996). *Methods and Application of Linier Model*. New York: John Wiley & Sons.
- Hosmer, D.W., dan Lameshow, S. (2000). *Applied Logistic Regression Second Edition*. New Jersey: John Wiley & Sons.
- Hosmer, D.W., Lameshow, S., dan Sturdivant, R.X. (2013). *Applied Logistic Regression Third Edition*. New Jersey: John Wiley & Sons.
- Johnson, R.A., dan Wichen, D.W. (2007). *Applied Multivariat Statistical Sixth Edition*. New Jersey: Prentice Hall International Inc.
- Nelson, P.R., Coffin, M., dan Copeland, K.A.F. (2003). *Introductory Statistics for Engineering Experimentation*. USA: Elsevier Science.
- Peraturan Pemerintah Republik Indonesia Nomor 82 Tahun 2001 Tentang Pengelolaan Kualitas Air dan Pengendalian Pencemaran Air
- Pravitasary, A.T., Hajarisman, N., dan Sunendiani, S. (2015). Pemodelan Faktor-Faktor yang Berpengaruh Terhadap Angka Buta Huruf di Provinsi Jawa Barat dengan Geographically Weighted Logistic Regression. *Prosiding Penelitian SPeSIA*. 2(1), 21-28.
- Salmin. (2005). Oksigen Terlarut (DO) dan Kebutuhan Oksigen Biologi (BOD) Sebagai Salah Satu Indikator untuk Menentukan Kualitas Perairan. *Osen*. 30(3), 21-26.
- Susilowati, Y., Leksono, B.E., dan Harsono, E. (2012). Pemodelan Kualitas Air Sungai Mahakam Sebagai Dasar Perencanaan Pengelolaan Lahan Wilayah Provinsi Kalimantan Timur. *Prosiding Pemaparan Hasil Penelitian Geoteknologi LIPI*. ISBN, 153-165.
- Suyitno, Puhadi, Sutikno, dan Irhamah. (2016). Parameter Estimation of Geographically Weighted Trivariate Weibull Regression Model. *Applied Mathematical Science*. 10(18), 861-878.