

# Modeling Chemical Oxygen Demand of River Water in East Kalimantan Using Fixed Effects and Geographically Weighted Panel Regression

Memi Nor Hayati <sup>1\*,\*\*\*</sup>, Toha Saifudin <sup>2\*\*</sup>, Suyitno <sup>3\*\*\*</sup>

\* Doctoral Study Program MIPA, Faculty of Science and Technology, Airlangga University, Indonesia

\*\* Department of Mathematics, Faculty of Science and Technology, Airlangga University, Indonesia

\*\*\* Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Mulawarman University, Indonesia

[meminorhayati@fmipa.unmul.ac.id](mailto:meminorhayati@fmipa.unmul.ac.id)<sup>1</sup>, [tohasaifudin@fst.unair.ac.id](mailto:tohasaifudin@fst.unair.ac.id)<sup>2</sup>, [suyitno@fmipa.unmul.ac.id](mailto:suyitno@fmipa.unmul.ac.id)<sup>3</sup>

## Article Info

### Article history:

Received 2026-05-25

Revised 2026-05-31

Accepted 2026-06-11

### Keyword:

Chemical Oxygen Demand,  
Fixed Effect Model,  
Geographically Weighted Panel  
Regression,  
River Water Quality,  
SDGs Goal 6.

## ABSTRACT

The degradation of river water quality in East Kalimantan has become an environmental issue related to the Sustainable Development Goals (SDGs), particularly Goal 6 concerning clean water and sanitation. This study aims to model Chemical Oxygen Demand (COD) as an indicator of river water quality using the Fixed Effect Model (FEM) and Geographically Weighted Panel Regression (GWPR) approaches, as well as to compare the performance of both models in representing spatial and temporal heterogeneity. The data used in this study consist of panel data from 31 observation locations over four semesters, with predictor variables including pH, Total Suspended Solid (TSS), Fecal Coliform, Total Dissolved Solid (TDS), and ammonia. FEM estimation was conducted using the within estimator, while the GWPR model employed an adaptive bisquare kernel weighting function. The results show that TSS and ammonia significantly affect COD in the FEM model. In the GWPR model, the effects of predictor variables vary across observation locations, indicating spatial heterogeneity in the factors influencing COD. TSS was identified as the most dominant variable, being significant at 26 locations, followed by ammonia and TDS, which were significant at 18 and 15 locations, respectively. Furthermore, the GWPR model produced a lower RMSE value (1.3777) than the FEM model (2.182011). These findings indicate that GWPR performs better in capturing spatial heterogeneity and temporal information in river water quality and provides more detailed information for supporting location-specific water quality management in East Kalimantan.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

## I. PENDAHULUAN

Air merupakan sumber daya vital yang berperan penting dalam menunjang kehidupan manusia serta menjaga keberlanjutan ekosistem. Ketersediaan air bersih yang aman menjadi salah satu fokus utama dalam *Sustainable Development Goals* (SDGs), khususnya tujuan ke-6 yang menekankan akses terhadap air bersih dan sanitasi yang layak [1]. Selain berorientasi pada peningkatan akses air bersih, SDGs juga mendorong upaya pengurangan pencemaran serta pengelolaan sumber daya air secara berkelanjutan [2]. Namun, dalam beberapa tahun terakhir, kondisi kualitas air sungai mengalami penurunan akibat meningkatnya aktivitas manusia, seperti pertumbuhan penduduk, industrialisasi,

kegiatan pertanian, dan perubahan penggunaan lahan. Berbagai aktivitas domestik, pertanian, dan industri diketahui berkontribusi terhadap peningkatan beban pencemar di badan air sehingga menurunkan kualitas air sungai [3].

Kalimantan Timur merupakan salah satu wilayah dengan dinamika pembangunan yang sangat tinggi di Indonesia. Kegiatan pertambangan batu bara, ekspansi perkebunan kelapa sawit, serta pembangunan infrastruktur dalam rangka pengembangan Ibu Kota Nusantara (IKN) memberikan tekanan ekologis yang signifikan terhadap ekosistem perairan. Sungai-sungai utama seperti Sungai Mahakam dan Sungai Berau tidak hanya menjadi sumber kehidupan masyarakat, tetapi juga berfungsi sebagai penerima limbah

dari berbagai aktivitas tersebut [4], [5]. Kondisi ini berpotensi menurunkan kualitas air secara signifikan apabila tidak dikelola dengan pendekatan yang tepat dan berbasis data.

Dalam evaluasi kualitas air sungai, kondisi perairan umumnya dianalisis melalui parameter fisik, kimia, dan biologis sebagai indikator tingkat pencemaran. Salah satu parameter kimia yang sering digunakan adalah *Chemical Oxygen Demand* (COD), yaitu ukuran jumlah oksigen yang diperlukan dalam proses oksidasi senyawa pencemar organik maupun anorganik secara kimiawi [6]. Parameter COD digunakan secara luas karena mampu menggambarkan total kandungan bahan pencemar dalam badan air secara cepat dan menyeluruh. Nilai COD yang tinggi menunjukkan besarnya kandungan polutan yang berpotensi menurunkan mutu air, mengganggu keseimbangan ekosistem perairan, serta memberikan dampak terhadap kesehatan manusia. Dibandingkan *Biological Oxygen Demand* (BOD) yang hanya mengukur bahan organik yang mudah terurai, COD memberikan informasi yang lebih luas mengenai tingkat pencemaran kimia dalam perairan [7]. Oleh sebab itu, analisis COD menjadi penting sebagai dasar dalam pemantauan dan pengelolaan kualitas air sungai secara berkelanjutan.

Pemantauan kualitas air sungai umumnya menggunakan data yang diamati pada beberapa lokasi dan periode waktu tertentu sehingga membentuk struktur data panel. Pendekatan regresi dengan data panel memiliki beberapa kelebihan dibandingkan analisis data *cross section* maupun *time series* secara terpisah karena mampu meningkatkan jumlah observasi, mengurangi potensi kolinearitas antarvariabel prediktor, serta mengakomodasi variasi antarwilayah dan antarwaktu secara simultan [8]. Secara umum, *Random Effect Model* (REM), *Fixed Effect Model* (FEM), dan *Common Effect Model* (CEM) dikenal sebagai tiga pendekatan utama dalam model regresi data panel [9], [10]. Pendekatan FEM digunakan ketika terdapat perbedaan karakteristik pada masing-masing individu atau wilayah yang memengaruhi intersep model, sedangkan koefisien *slope* diasumsikan tetap [11]. Model ini banyak diterapkan dalam penelitian ekonomi, sosial, maupun lingkungan karena mampu mengendalikan heterogenitas yang tidak teramati (*unobserved heterogeneity*).

Beberapa penelitian terdahulu [8],[9],[10] menerapkan FEM dan menunjukkan bahwa pendekatan FEM memberikan hasil yang baik dalam analisis data panel. Hasil-hasil penelitian tersebut menunjukkan bahwa pendekatan FEM efektif digunakan untuk menganalisis data yang memiliki karakteristik heterogen antarwilayah dan antarperiode waktu, sehingga relevan diterapkan pada penelitian ini dalam memodelkan COD sebagai indikator kualitas air sungai di Kalimantan Timur.

Meskipun demikian, hubungan antara variabel respon dan variabel prediktor dalam model regresi panel global diasumsikan sama pada seluruh lokasi pengamatan. Dalam kajian kualitas air sungai, asumsi tersebut sering kali kurang sesuai karena setiap wilayah memiliki karakteristik geografis, aktivitas manusia, serta kondisi lingkungan yang berbeda-beda. Perbedaan karakteristik tersebut dapat menyebabkan

hubungan antarvariabel tidak bersifat homogen pada seluruh lokasi pengamatan. Pada wilayah Kalimantan Timur, misalnya, daerah yang didominasi aktivitas pertambangan batu bara, perkebunan kelapa sawit, dan kawasan perkotaan memiliki tekanan lingkungan yang berbeda sehingga respons nilai COD terhadap variabel prediktor juga berpotensi bervariasi secara spasial. Kondisi tersebut menunjukkan adanya heterogenitas spasial, yaitu perbedaan karakteristik hubungan antarvariabel pada setiap lokasi pengamatan.

Untuk mengakomodasi adanya heterogenitas spasial, penelitian ini menggunakan pendekatan *Geographically Weighted Regression* (GWR). Metode ini memungkinkan setiap wilayah pengamatan memiliki parameter regresi yang berbeda sehingga hubungan antarvariabel dapat direpresentasikan secara lebih lokal [12], [13]. Pengembangan GWR pada data panel menghasilkan metode *Geographically Weighted Panel Regression* (GWPR) yang dapat memodelkan pengaruh spasial dan temporal secara simultan [14], [15], [16]. Dengan pendekatan ini, pengaruh faktor-faktor yang memengaruhi kualitas air sungai dapat dianalisis berdasarkan perbedaan lokasi dan waktu pengamatan.

Berdasarkan beberapa penelitian sebelumnya [17], [18], [19], [20] model GWPR dinilai lebih mampu merepresentasikan variasi spasial dibandingkan model global. Meskipun penelitian mengenai regresi data panel maupun model geografis telah banyak dilakukan dalam berbagai bidang penelitian, termasuk studi lingkungan dan kualitas air, kajian yang secara khusus memodelkan indikator kualitas air sungai COD menggunakan pendekatan FEM dan GWPR masih relatif terbatas, khususnya di wilayah Kalimantan Timur. Sebagian besar penelitian sebelumnya terkait COD di sekitar wilayah Kalimantan Timur hanya menggunakan analisis deskriptif kuantitatif [21], atau hanya menggunakan pendekatan geografis pada data *cross section* tanpa mempertimbangkan dimensi panel [22]. Padahal, data kualitas air yang diamati pada banyak lokasi dan periode waktu berpotensi menunjukkan heterogenitas spasial yang menyebabkan hubungan antara COD dan faktor-faktor kualitas air berbeda antarwilayah. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut melalui penerapan dan perbandingan model FEM dan GWPR pada data panel indikator kualitas air sungai COD di wilayah Kalimantan Timur.

Penelitian ini bertujuan memodelkan COD sebagai indikator kualitas perairan sungai di Kalimantan Timur melalui pendekatan FEM dan GWPR. Selain itu, penelitian ini dilakukan untuk mengidentifikasi faktor-faktor yang memengaruhi kualitas perairan sungai serta mengevaluasi kemampuan model global dan model spasial panel dalam menggambarkan heterogenitas spasial maupun temporal pada data kualitas air. Melalui penerapan GWPR pada data panel yang mencakup 31 lokasi pengamatan selama empat semester, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan pemodelan kualitas air berbasis spasial-panel sekaligus menghasilkan informasi

spasial yang lebih rinci untuk mendukung pengelolaan kualitas perairan sungai dan perumusan kebijakan lingkungan yang lebih sesuai dengan karakteristik wilayah di Kalimantan Timur.

## II. METODE

Tahapan dalam penelitian ini meliputi pengumpulan data, penentuan variabel penelitian, serta analisis data menggunakan pendekatan FEM dan GWPR untuk mengidentifikasi pengaruh variabel prediktor secara global maupun spasial/lokal. Dalam penelitian ini, FEM digunakan sebagai model panel global pembandingan karena formulasi GWPR yang digunakan dikembangkan berdasarkan pendekatan *within estimator* pada FEM. Oleh karena itu, FEM digunakan sebagai dasar dalam pembentukan model spasial panel untuk mengevaluasi perbedaan representasi hubungan global dan lokal pada data kualitas air sungai. Selanjutnya, performa FEM dibandingkan dengan GWPR untuk mengidentifikasi kemampuan masing-masing model dalam merepresentasikan variasi COD pada lokasi dan periode pengamatan yang berbeda. Proses pengolahan dan analisis data dilakukan dengan bantuan perangkat lunak RStudio dan GNU Octave guna mendukung kebutuhan komputasi serta visualisasi data.

### A. Data Penelitian

Penelitian dilakukan pada 31 titik pengamatan di daerah aliran sungai wilayah Kalimantan Timur yang diukur oleh Dinas Lingkungan Hidup (DLH) Provinsi Kalimantan Timur selama empat semester pada periode 2024–2025. Nama 31 titik sampel lokasi pengamatan daerah aliran Sungai di Kalimantan Timur dapat dilihat pada Tabel I.

TABEL I  
TITIK SAMPEL LOKASI PENGAMATAN

No	Nama Lokasi	Kabupaten/Kota
1	Anggana Kukar	Kutai Kartanegara
2	Bloro Kukar	Kutai Kartanegara
3	Desa Kedang Kepala Kukar	Kutai Kartanegara
4	Desa Sebelimbingan Kukar	Kutai Kartanegara
5	Desa Siran Kukar	Kutai Kartanegara
6	Hilir Berau Coal	Berau
7	Hulu PT. BBE Berau	Berau
8	Jembatan Desa Pampang	Samarinda
9	Jembatan Gunung Tabur	Berau
10	Jembatan Jalan Poros Bontang-Samarinda Km.8	Kutai Timur
11	Jembatan Jl. Soekarno-Hatta	Bontang
12	Jembatan Marangkayu Santan Ulu Kukar	Kutai Kartanegara
13	Jembatan Santan Tengah Kukar	Kutai Kartanegara
14	Kalamur Samarinda	Samarinda
15	Kantor Bupati Berau	Berau
16	Kantor Gubernur Samarinda	Samarinda
17	Keraton Gunung Tabur Berau	Berau
18	Keraton Sambaliung Berau	Berau

No	Nama Lokasi	Kabupaten/Kota
19	Long Bagun Mahulu	Mahakam Ulu
20	Muara Sungai Belayan Kukar	Kutai Kartanegara
21	Nyan Mahulu	Mahakam Ulu
22	Palaran Samarinda	Samarinda
23	Pampang Dalam Samarinda	Samarinda
24	Pulau Kumala Kukar	Kutai Kartanegara
25	Tering Kubar	Kutai Barat
26	Titik Hulu (Sungai 28)	Balikpapan
27	Hulu Sungai Kasungai	Paser
28	Nursery Suring PPU	Penajam Paser Utara
29	Hilir Sungai (Pos Pon 1)	Balikpapan
30	Titik Hilir Sungai Kasungai	Paser
31	Titik Jembatan Sei Sepaku	Penajam Paser Utara

### B. Variabel Penelitian

*Chemical Oxygen Demand* (COD) merupakan variabel respon dalam penelitian ini, kemudian pH air ( $x_1$ ), *Total Suspended Solid* (TSS) ( $x_2$ ), *Fecal Coliform* ( $x_3$ ), *Total Dissolved Solid* (TDS) ( $x_4$ ), dan Amoniak ( $x_5$ ) merupakan variabel prediktor yang diduga memengaruhi COD. Selain itu terdapat variabel koordinat titik lokasi pengamatan ( $u_i, v_i$ ) dengan  $u_i$  menunjukkan posisi garis lintang dan  $v_i$  menunjukkan posisi garis bujur.

Pertimbangan teoritis dan hasil penelitian terdahulu mengenai faktor-faktor yang berkaitan dengan kualitas air menjadi dasar pemilihan variabel prediktor dalam penelitian ini. Variabel pH digunakan karena tingkat keasaman perairan berpengaruh terhadap berbagai proses kimia dan biologis yang dapat memengaruhi kondisi kualitas air [7]. Variabel TSS dipilih karena tingginya konsentrasi padatan tersuspensi dapat meningkatkan kandungan bahan pencemar dan berkontribusi terhadap peningkatan kebutuhan oksigen dalam perairan [6], [7]. *Fecal Coliform* digunakan sebagai indikator pencemaran biologis yang umumnya berasal dari limbah domestik dan aktivitas manusia di sekitar badan air [6]. Variabel TDS merepresentasikan jumlah zat terlarut yang memengaruhi karakteristik fisik dan kimia perairan, sedangkan amoniak merupakan salah satu senyawa nitrogen yang sering ditemukan pada perairan yang menerima beban pencemar organik [7]. Selain memiliki dasar teoritis yang kuat, kelima variabel tersebut juga telah digunakan dalam penelitian sebelumnya terkait pemodelan COD pada daerah aliran sungai di Kalimantan Timur [22], sehingga dipandang relevan untuk digunakan.

### C. Fixed Effect Model (FEM)

FEM digunakan ketika terdapat karakteristik khusus pada masing-masing entitas atau periode waktu yang bersifat tetap serta berkorelasi dengan variabel prediktor [11], [23]. *Within Transformation (demeaning)* adalah salah satu pendekatan untuk mengestimasi parameter FEM [11]. Pada pendekatan ini, data *cross section* untuk variabel respon maupun variabel prediktor ditransformasikan dengan mengurangkan nilai

pengamatan terhadap rata-rata *time series* yang bersesuaian [24]:

$$\begin{aligned} y_{it}^* &= (y_{it} - \bar{y}_i), \\ x_{kit}^* &= (x_{kit} - \bar{x}_{ki}), \\ \varepsilon_{it}^* &= (\varepsilon_{it} - \bar{\varepsilon}_i) \end{aligned} \tag{1}$$

dimana

$$\begin{aligned} \bar{y}_i &= \frac{1}{\tau} \sum_{t=1}^{\tau} y_{it}, \\ \bar{x}_{ki} &= \frac{1}{\tau} \sum_{t=1}^{\tau} x_{kit}, \\ \bar{\varepsilon}_i &= \frac{1}{\tau} \sum_{t=1}^{\tau} \varepsilon_{it}, k = 1, 2, \dots, p \end{aligned} \tag{2}$$

Data hasil transformasi pada persamaan (1) disebut data *demean*. FEM *within estimator* dapat ditulis

$$y_{it}^* = \beta_1 x_{1it}^* + \beta_2 x_{2it}^* + \dots + \beta_p x_{pit}^* + \varepsilon_{it}^*, \tag{3}$$

$t = 1, 2, \dots, \tau$

atau dalam bentuk matriks yaitu

$$\mathbf{y}^* = \mathbf{X}^* \boldsymbol{\beta} + \boldsymbol{\varepsilon}^*, \tag{4}$$

dengan

$$\mathbf{y}^* = \begin{bmatrix} y_{11}^* \\ y_{21}^* \\ \vdots \\ y_{n1}^* \\ y_{12}^* \\ y_{22}^* \\ \vdots \\ y_{n2}^* \\ \vdots \\ y_{1\tau}^* \\ y_{2\tau}^* \\ \vdots \\ y_{n\tau}^* \end{bmatrix}; \mathbf{X}^* = \begin{bmatrix} x_{111}^* & x_{211}^* & \dots & x_{p11}^* \\ x_{121}^* & x_{221}^* & \dots & x_{p21}^* \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n1}^* & x_{2n1}^* & \dots & x_{pn1}^* \\ x_{112}^* & x_{212}^* & \dots & x_{p12}^* \\ x_{122}^* & x_{222}^* & \dots & x_{p22}^* \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n2}^* & x_{2n2}^* & \dots & x_{pn2}^* \\ \vdots & \vdots & \ddots & \vdots \\ x_{11\tau}^* & x_{21\tau}^* & \dots & x_{p1\tau}^* \\ x_{12\tau}^* & x_{22\tau}^* & \dots & x_{p2\tau}^* \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n\tau}^* & x_{2n\tau}^* & \dots & x_{pn\tau}^* \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}; \boldsymbol{\varepsilon}^* = \begin{bmatrix} \varepsilon_{11}^* \\ \varepsilon_{21}^* \\ \vdots \\ \varepsilon_{n1}^* \\ \varepsilon_{12}^* \\ \varepsilon_{22}^* \\ \vdots \\ \varepsilon_{n2}^* \\ \vdots \\ \varepsilon_{1\tau}^* \\ \varepsilon_{2\tau}^* \\ \vdots \\ \varepsilon_{n\tau}^* \end{bmatrix} \tag{5}$$

Penaksiran parameter FEM yang diberikan oleh persamaan (4) adalah metode *Ordinary Least Squares* (OLS).

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*T} \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{y}^* \tag{6}$$

Pengujian simultan pada model FEM dilakukan untuk mengevaluasi signifikansi variabel prediktor secara bersama-sama terhadap variabel respon. Hipotesis yang digunakan adalah:

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \dots = \beta_p = 0 \\ H_1 &: \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, p \end{aligned}$$

Statistik uji simultan dinyatakan sebagai berikut:

$$F_1 = \frac{\left( \sum_{i=1}^n \sum_{t=1}^{\tau} (\hat{y}_{it}^* - \bar{y}^*)^2 \right) / p}{\left( \sum_{i=1}^n \sum_{t=1}^{\tau} (y_{it}^* - \hat{y}_{it}^*)^2 \right) / (n\tau - n - p)} \tag{7}$$

pada taraf signifikansi  $\alpha$ ,  $H_0$  ditolak apabila  $F_1 > F_{(1-\alpha; p; n\tau - n - p)}$ , atau nilai  $p_1$  lebih kecil dari  $\alpha$  [11], [23], [25].

Pengujian parsial pada model FEM digunakan untuk mengevaluasi pengaruh masing-masing variabel prediktor terhadap variabel respon. Hipotesis pengujian parameter  $\beta_k$  dengan  $k = 1, 2, \dots, p$  dinyatakan sebagai berikut:

$$\begin{aligned} H_0 &: \beta_k = 0 \\ H_1 &: \beta_k \neq 0 \end{aligned}$$

Statistik uji parsial pada model FEM dirumuskan sebagai berikut:

$$T_1 = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)}, \tag{8}$$

pada taraf signifikansi  $\alpha$ ,  $H_0$  ditolak apabila  $|T_1| > t_{1-\frac{\alpha}{2}; (n\tau - n - p)}$ , atau nilai  $p_2$  lebih kecil dari  $\alpha$  [11], [23].

#### D. Geographically Weighted Panel Regression (GWPR)

Pengembangan dari metode GWR dengan menambahkan dimensi waktu dalam struktur data panel dikenal dengan model *Geographically Weighted Panel Regression* (GWPR). Pendekatan ini memungkinkan parameter regresi berbeda pada setiap lokasi pengamatan sekaligus mempertimbangkan variasi temporal pada data [14], [17]. Berdasarkan pendekatan *within estimator* pada model FEM di persamaan (3), model GWPR dapat dituliskan sebagai berikut [17], [18], [26]:

$$\begin{aligned} y_{it}^* &= \beta_1(u_i, v_i) x_{1it}^* + \beta_2(u_i, v_i) x_{2it}^* + \dots \\ &\quad + \beta_p(u_i, v_i) x_{pit}^* + \varepsilon_{it}^* \end{aligned} \tag{9}$$

$i = 1, 2, \dots, n; t = 1, 2, \dots, \tau$

Berdasarkan persamaan (9), model GWPR dapat dituliskan dalam bentuk matriks sebagai berikut:

$$\mathbf{y}^* = \mathbf{X}^* \boldsymbol{\beta}(u_i, v_i) + \boldsymbol{\varepsilon}^* \tag{10}$$

Estimasi parameter pada model GWPR dilakukan menggunakan pendekatan *Weighted Least Squares* (WLS). Bentuk estimator parameter GWPR dinyatakan sebagai berikut [17], [27]:

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = (\mathbf{X}^{*T} \mathbf{W}(u_i, v_i) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_i, v_i) \mathbf{y}^* \tag{11}$$

Selanjutnya,  $\hat{\mathbf{y}}^*$  didefinisikan sebagai vektor estimasi variabel respon, sedangkan  $\hat{\boldsymbol{\varepsilon}}^*$  merupakan vektor galat yang diperoleh sebagai berikut:

$$\hat{\mathbf{y}}^* = \begin{bmatrix} \hat{y}_{11}^* \\ \hat{y}_{21}^* \\ \vdots \\ \hat{y}_{n1}^* \\ \hat{y}_{12}^* \\ \hat{y}_{22}^* \\ \vdots \\ \hat{y}_{n2}^* \\ \vdots \\ \hat{y}_{1\tau}^* \\ \hat{y}_{2\tau}^* \\ \vdots \\ \hat{y}_{n\tau}^* \end{bmatrix} = \mathbf{L}^* \mathbf{y}^* \text{ dan } \hat{\boldsymbol{\varepsilon}}^* = \begin{bmatrix} \hat{\varepsilon}_{11}^* \\ \hat{\varepsilon}_{21}^* \\ \vdots \\ \hat{\varepsilon}_{n1}^* \\ \hat{\varepsilon}_{12}^* \\ \hat{\varepsilon}_{22}^* \\ \vdots \\ \hat{\varepsilon}_{n2}^* \\ \vdots \\ \hat{\varepsilon}_{1\tau}^* \\ \hat{\varepsilon}_{2\tau}^* \\ \vdots \\ \hat{\varepsilon}_{n\tau}^* \end{bmatrix} = (\mathbf{I} - \mathbf{L}^*) \mathbf{y}^*, \quad (12)$$

dengan  $\mathbf{I}_{n\tau}$  sebagai matriks identitas berukuran  $n\tau \times n\tau$ , sedangkan bentuk matriks  $\mathbf{L}^*$  dituliskan sebagai berikut:

$$\mathbf{L}^* = \begin{bmatrix} \mathbf{x}_{11}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_1, v_1) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_1, v_1) \\ \mathbf{x}_{21}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_2, v_2) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_{n1}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_n, v_n) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_n, v_n) \\ \mathbf{x}_{12}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_1, v_1) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_1, v_1) \\ \mathbf{x}_{22}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_2, v_2) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_{n2}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_n, v_n) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_n, v_n) \\ \vdots \\ \mathbf{x}_{1\tau}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_1, v_1) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_1, v_1) \\ \mathbf{x}_{2\tau}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_2, v_2) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_2, v_2) \\ \vdots \\ \mathbf{x}_{n\tau}^{*T} (\mathbf{X}^{*T} \mathbf{W}(u_n, v_n) \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W}(u_n, v_n) \end{bmatrix} \quad (13)$$

Matriks pembobot dalam persamaan (11) berupa blok diagonal sebanyak  $\tau$  blok:

$$\mathbf{W}(u_i, v_i) = \begin{bmatrix} \mathbf{W}_i & 0 & \dots & 0 \\ 0 & \mathbf{W}_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{W}_i \end{bmatrix}_{n\tau \times n\tau}$$

dimana  $\mathbf{W}_i = \text{diag}[w_1(u_i, v_i) \ w_2(u_i, v_i) \ \dots \ w_n(u_i, v_i)]$ , matriks  $\mathbf{W}_i$  berukuran  $n \times n$  dan digunakan sebagai matriks pembobot spasial. Pada penelitian ini, bobot spasial dihitung menggunakan fungsi kernel *adaptive bisquare* [12]:

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^2\right)^2, & \text{untuk } d_{ij} \leq b_i \\ 0, & \text{untuk } d_{ij} > b_i \end{cases} \quad (14)$$

Fungsi kernel *adaptive bisquare* dipilih karena mampu menyesuaikan jumlah tetangga yang digunakan pada setiap lokasi pengamatan. Karakteristik ini sesuai untuk data penelitian yang memiliki sebaran titik pengamatan tidak sepenuhnya seragam antarwilayah. Dibandingkan kernel tetap (*fixed kernel*), *adaptive bisquare* dapat memberikan bobot yang lebih fleksibel sehingga estimasi parameter lokal

menjadi lebih stabil pada wilayah dengan kepadatan titik pengamatan yang berbeda. Selain itu, fungsi kernel ini banyak digunakan dalam penelitian GWPR karena mampu menghasilkan estimasi parameter lokal yang baik pada data yang menunjukkan heterogenitas spasial.

Penentuan *bandwidth* optimal  $b_i$  dilakukan menggunakan kriteria *Akaike Information Criterion* (AIC) [28]:

$$AIC = n \log \left( \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \right) + 2p \quad (15)$$

*Bandwidth* optimal diperoleh melalui proses iteratif dengan memilih nilai *bandwidth* yang menghasilkan nilai AIC minimum. Pada penelitian ini setiap kandidat *bandwidth* dievaluasi menggunakan fungsi kernel *adaptive bisquare* dan nilai *bandwidth* yang memberikan AIC terkecil digunakan dalam pembentukan matriks pembobot spasial pada model GWPR.

Hipotesis yang digunakan dalam pengujian kesesuaian model GWPR dinyatakan sebagai berikut [17], [18], [27]:

- $H_0$  :  $\beta_k(u_i, v_i) = \beta_k, i = 1, 2, \dots, n; k = 1, 2, \dots, p$   
(Tidak terdapat perbedaan yang signifikan antara model regresi panel FEM dengan model GWPR)
- $H_1$  : minimal ada satu  $\beta_k(u_i, v_i) \neq \beta_k, i = 1, 2, \dots, n; k = 1, 2, \dots, p$   
(Terdapat perbedaan yang signifikan antara model regresi panel FEM dengan model GWPR)

Statistik uji untuk pengujian kesesuaian model dirumuskan sebagai berikut:

$$F_2 = \frac{(\mathbf{y}^* - \hat{\mathbf{y}})^T (\mathbf{y}^* - \hat{\mathbf{y}}) / db_1}{\mathbf{y}^{*T} (\mathbf{I} - \mathbf{L}^*)^T (\mathbf{I} - \mathbf{L}^*) \mathbf{y}^* / db_2} \quad (16)$$

dimana  $db_1 = (n\tau - p - 1)$  dan  $db_2 = \frac{\delta_1^2}{\delta_2}$ . Nilai dari  $\delta_1 = \text{tr}((\mathbf{I} - \mathbf{L}^*)^T (\mathbf{I} - \mathbf{L}^*))$  dan  $\delta_2 = \text{tr}((\mathbf{I} - \mathbf{L}^*)^T (\mathbf{I} - \mathbf{L}^*))^2$ . Pada taraf signifikansi  $\alpha$ ,  $H_0$  ditolak apabila  $F_2 > F_{(\alpha; db_1; db_2)}$ , atau nilai  $p_3$  lebih kecil dari  $\alpha$ .

Selanjutnya hipotesis dari uji parsial model GWPR adalah [17], [18], [27]:

- $H_0$  :  $\beta_k(u_i, v_i) = 0$   
(variabel prediktor ke-k tidak berpengaruh signifikan pada lokasi pengamatan ke-i)
- $H_1$  :  $\beta_k(u_i, v_i) \neq 0$   
(variabel prediktor ke-k berpengaruh signifikan pada lokasi pengamatan ke-i)

Statistik uji parsial pada model GWPR dirumuskan sebagai berikut:

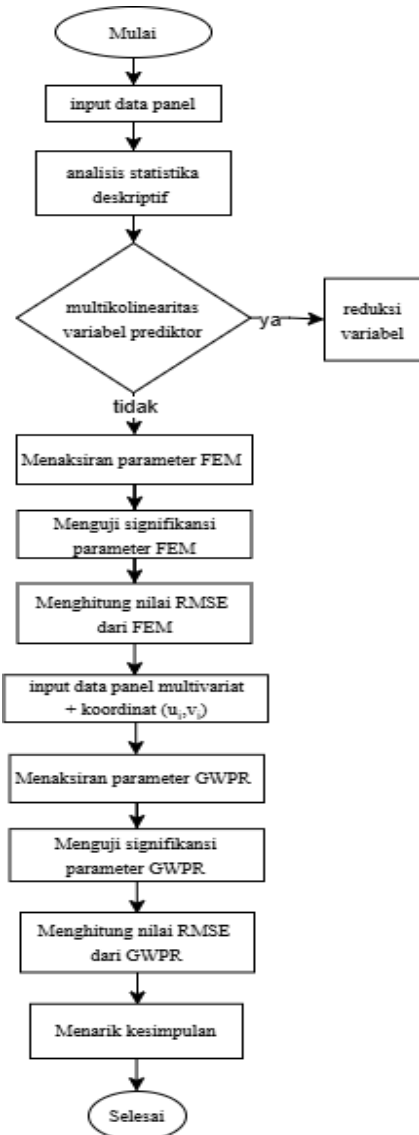
$$T_2 = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma}_{GWPR} \sqrt{c_{kk}}} \quad (17)$$

pada taraf signifikansi  $\alpha$ ,  $H_0$  ditolak apabila  $|T_2| > t_{\left(1 - \frac{\alpha}{2}; \frac{\delta_1^2}{\delta_2}\right)}$ ,

atau nilai  $p_4$  lebih kecil dari  $\alpha$ .

E. Tahapan Analisis Data

Tahapan analisis data pada penelitian ini dilakukan secara bertahap, dimulai dari analisis deskriptif, pembentukan model global menggunakan FEM, hingga pemodelan spasial menggunakan GWPR. Adapun tahapan analisis yang dilakukan adalah sebagai berikut.



Gambar 1. Diagram Alir Penelitian

1. Analisis deskriptif
2. Mengidentifikasi multikolinearitas antarvariabel prediktor menggunakan nilai *Variance Inflation Factor* (VIF). Variabel prediktor ke-k dinyatakan mengalami multikolinearitas apabila nilai VIF lebih dari 10 [23]:

$$VIF_k = \frac{1}{1 - R_k^2} \quad (18)$$

3. Menaksiran parameter FEM berdasarkan persamaan (6)
4. Melakukan pengujian signifikansi parameter FEM secara simultan dan parsial.

5. Menghitung nilai *Root Mean Square Error* (RMSE) pada model FEM berdasarkan persamaan (19) [29], [30]:

$$RMSE = \sqrt{\frac{1}{n\tau} \sum_{i=1}^n \sum_{t=1}^{\tau} (y_{it} - \hat{y}_{it})^2}, \quad (19)$$

6. Menaksiran parameter GWPR berdasarkan persamaan (11)
7. Melakukan pengujian kesesuaian model FEM dan GWPR serta pengujian parsial parameter pada model GWPR.
8. Menghitung nilai RMSE dari GWPR untuk mengevaluasi performa model GWPR dibandingkan model FEM.
9. Membuat kesimpulan.

III. HASIL DAN PEMBAHASAN

A. Analisis Deskriptif Data Penelitian

Analisis deskriptif dilakukan sebagai tahap awal untuk menggambarkan karakteristik umum data penelitian sebelum dilakukan pemodelan lebih lanjut. Variabel yang dianalisis terdiri atas variabel respon COD dengan satuan mg/L dan beberapa variabel prediktor, yaitu pH air, TSS dalam satuan mg/L, *Fecal Coliform* dalam satuan MPN/100 mL, TDS dalam satuan mg/L, serta amoniak dalam satuan mg/L. Ringkasan hasil analisis deskriptif disajikan pada Tabel II.

TABEL II  
STATISTIK DESKRIPTIF VARIABEL PENELITIAN

Variabel	Min	Maks	Range	Rata-Rata	Simpangan Baku
COD (y)	10	26	16	18,0968	3,2269
pH(x <sub>1</sub> )	4,3	8,73	4,43	7,3535	0,6636
TSS(x <sub>2</sub> )	5	185	180	55,0403	37,6381
<i>Fecal Coliform</i> (x <sub>3</sub> )	17	230	213	79,3468	42,1341
TDS(x <sub>4</sub> )	18	304	286	67,7903	58,9065
Amoniak (x <sub>5</sub> )	0,01	0,4	0,39	0,0902	0,0612

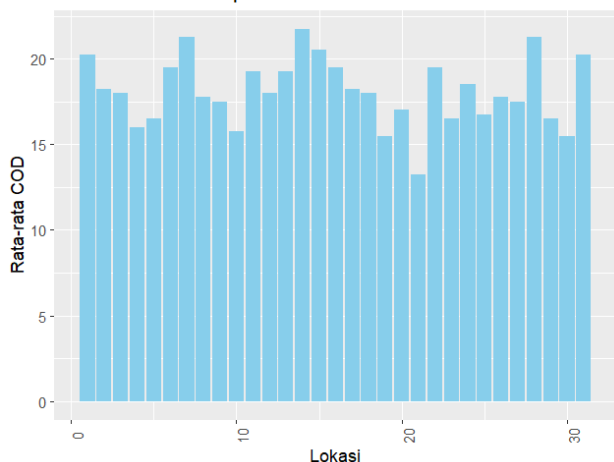
Berdasarkan hasil statistik deskriptif, variabel COD memiliki nilai minimum sebesar 10 mg/L dan maksimum sebesar 26 mg/L dengan rata-rata sebesar 18,1 mg/L. Nilai tersebut menunjukkan tingkat COD pada lokasi pengamatan selama empat semester memiliki variasi yang cukup beragam antarwilayah.

Sementara itu, variabel TSS dan TDS menunjukkan rentang data yang cukup besar dibandingkan variabel lainnya. Kondisi tersebut mengindikasikan adanya variasi karakteristik kualitas air antarwilayah pengamatan. Selain itu, nilai maksimum *Fecal Coliform* yang relatif tinggi menunjukkan adanya beberapa lokasi dengan tingkat kontaminasi biologis yang cukup besar.

Hasil analisis deskriptif secara umum menunjukkan adanya variasi data yang cukup tinggi antarwilayah dan antarperiode pengamatan. Kondisi tersebut mengindikasikan

adanya kemungkinan heterogenitas spasial pada data, sehingga pendekatan model spasial seperti GWPR relevan digunakan untuk mengakomodasi perbedaan karakteristik pada masing-masing lokasi pengamatan.

Rata-rata COD Tiap Lokasi Selama 4 Semester



Gambar 2. Rata-Rata nilai COD pada 31 lokasi

Gambar 2 menunjukkan rata-rata nilai COD pada 31 lokasi pengamatan selama empat semester. Secara umum, rata-rata COD pada setiap lokasi berada pada kisaran 13 mg/L hingga 22 mg/L. Beberapa lokasi memiliki rata-rata COD yang relatif lebih tinggi dibandingkan lokasi lainnya, sehingga menunjukkan adanya perbedaan kondisi kualitas air antarwilayah.

Variasi rata-rata COD antarwilayah tersebut mengindikasikan bahwa karakteristik pencemaran organik pada masing-masing lokasi tidak bersifat homogen. Kondisi ini menjadi salah satu indikasi adanya pengaruh spasial pada data.

**B. Deteksi Multikolinearitas Antar Variabel Prediktor**

Tabel III menunjukkan bahwa seluruh variabel prediktor memiliki nilai VIF di bawah 10. Kondisi tersebut mengindikasikan tidak adanya gejala multikolinearitas antarvariabel prediktor sehingga seluruh variabel dapat digunakan pada tahap pemodelan FEM dan GWPR.

TABEL III  
MULTIKOLINEARITAS MENGGUNAKAN VIF

Variabel	VIF
pH(x <sub>1</sub> )	1,120319
TSS(x <sub>2</sub> )	1,061750
Fecal Coliform(x <sub>3</sub> )	1,164753
TDS(x <sub>4</sub> )	1,233143
Amoniak(x <sub>5</sub> )	1,045763

Selain pemeriksaan multikolinearitas, dilakukan pula pengujian asumsi model FEM yang meliputi normalitas residual, homoskedastisitas, dan autokorelasi. Hasil uji Shapiro–Wilk menunjukkan *p-value* sebesar 0,7763 sehingga residual dapat dikatakan berdistribusi normal pada taraf signifikansi 5%. Selanjutnya, hasil uji Breusch–Pagan

menghasilkan *p-value* sebesar 0,3764 yang menunjukkan tidak terdapat gejala heteroskedastisitas pada model. Sementara itu, hasil uji Breusch–Godfrey/Wooldridge menghasilkan *p-value* sebesar  $1,05 \times 10^{-5}$  sehingga menunjukkan adanya autokorelasi pada residual model. Temuan ini mengindikasikan bahwa model panel global masih menyisakan struktur ketergantungan yang belum sepenuhnya dapat dijelaskan. Meskipun demikian, hasil pengujian tersebut menunjukkan bahwa sebagian besar asumsi model telah terpenuhi sehingga model FEM tetap dapat digunakan sebagai model panel global pembandingan dalam penelitian ini. Selanjutnya, pendekatan GWPR diterapkan untuk mengakomodasi variasi hubungan antarvariabel pada lokasi pengamatan yang berbeda, sebagaimana telah diindikasikan oleh hasil analisis deskriptif yang menunjukkan adanya perbedaan karakteristik kualitas air antarwilayah.

**C. Analisis Data COD Menggunakan Pendekatan FEM**

Pendekatan FEM digunakan untuk menganalisis pengaruh variabel prediktor terhadap COD pada data panel. Estimasi parameter dilakukan menggunakan pendekatan *within estimator* melalui transformasi *demeaning*. Model FEM hasil transformasi *within* yang berdasarkan pada persamaan (3), untuk variabel respon COD dengan lima variabel prediktor pada  $i = 1, 2, \dots, 31$ ;  $t = 1, 2, 3, 4$ :

$$y_{it}^* = \beta_1 x_{1it}^* + \beta_2 x_{2it}^* + \beta_3 x_{3it}^* + \beta_4 x_{4it}^* + \beta_5 x_{5it}^* + \varepsilon_{it} \quad (20)$$

Estimasi parameter model FEM yang diperoleh menggunakan perangkat lunak R disajikan pada Tabel IV.

TABEL IV  
HASIL ESTIMASI PARAMETER FEM

Variabel	$\hat{\beta}_k$	$se(\hat{\beta}_k)$	$T_1$	$p_2$
pH	-0,8396	0,5022	-1,6719	0,0981
TSS	0,0404	0,0094	4,3173	0,000041*
Fecal Coliform	-0,0083	0,0076	-1,1007	0,2740
TDS	-0,0132	0,0105	-1,2610	0,2107
Amoniak	12,3576	4,6639	2,6496	0,0096*

\*Signifikan pada  $\alpha = 0,05$

Model FEM yang terbentuk berdasarkan hasil estimasi parameter pada Tabel IV

$$\hat{y}_{it}^* = -0,8396x_{1it}^* + 0,0404x_{2it}^* - 0,0083x_{3it}^* - 0,0132x_{4it}^* + 12,3576x_{5it}^*, \quad (21)$$

$$i = 1, 2, \dots, 31; t = 1, 2, 3, 4$$

dengan  $\hat{y}_{it}^*$  sebagai taksiran nilai COD, sedangkan  $x_{1it}^*, x_{2it}^*, x_{3it}^*, x_{4it}^*, x_{5it}^*$  masing-masing menyatakan nilai pH, TSS, Fecal Coliform, TDS, dan amoniak pada lokasi ke-*i* dan waktu ke-*t*.

TABEL V  
STATISTIK MODEL FEM

Statistik Model	Nilai
$F_1$	6,8003
$p_1$	$2,0965 \times 10^{-5}$ *
RMSE	2,182011

\*Signifikan pada  $\alpha = 0,05$

Tabel V menunjukkan bahwa nilai statistik uji simultan ( $F_1$ ) 6,8003 dengan nilai  $p$  sebesar  $2,0965 \times 10^{-5}$ . Karena nilai- $p$  berada di bawah taraf signifikansi 5%, variabel prediktor secara bersama-sama memiliki pengaruh signifikan terhadap COD. Dengan demikian, model FEM dinilai sesuai untuk merepresentasikan hubungan antara variabel prediktor dan variabel respon pada data penelitian.

Berdasarkan pada Tabel IV diperoleh hasil statistik uji parsial pada FEM yakni variabel TSS ( $x_2$ ) dan amoniak ( $x_5$ ) berpengaruh signifikan terhadap COD pada taraf signifikansi 5%. Variabel TSS memiliki koefisien positif sebesar 0,0404, yang menunjukkan bahwa peningkatan konsentrasi TSS sebesar 1 mg/L cenderung meningkatkan nilai COD sebesar 0,0404 mg/L dengan asumsi variabel lain konstan. Koefisien variabel amoniak ( $x_5$ ) bernilai positif sebesar 12,3576, yang mengindikasikan bahwa peningkatan kadar amoniak pada perairan cenderung meningkatkan nilai COD.

Sementara itu, variabel pH ( $x_1$ ), *Fecal Coliform* ( $x_3$ ), dan TDS ( $x_4$ ) tidak menunjukkan pengaruh signifikan terhadap COD pada taraf signifikansi 5%. Namun demikian, variabel pH memiliki nilai  $p$ -value mendekati 0,05 sehingga masih menunjukkan kecenderungan pengaruh terhadap COD.

**D. Analisis Data COD Menggunakan Pendekatan GWPR**

Berbeda dengan FEM yang menghasilkan parameter global, GWPR menghasilkan parameter lokal untuk setiap lokasi pengamatan sehingga pengaruh variabel prediktor dapat berbeda antarwilayah. Dengan mempertimbangkan heterogenitas spasial tersebut, model GWPR diharapkan mampu menghasilkan estimasi yang lebih representatif dibandingkan FEM. Dalam penelitian ini, model GWPR dibangun menggunakan data panel yang terdiri atas 31 lokasi pengamatan selama empat semester.

Model GWPR untuk variabel respon COD dengan lima variabel prediktor yang berdasarkan pada persamaan (9), pada  $i = 1, 2, \dots, 31$ ;  $t = 1, 2, 3, 4$ :

$$y_{it}^* = \beta_1(u_i, v_i)x_{1it}^* + \beta_2(u_i, v_i)x_{2it}^* + \beta_3(u_i, v_i)x_{3it}^* + \beta_4(u_i, v_i)x_{4it}^* + \beta_5(u_i, v_i)x_{5it}^* + \varepsilon_{it}^* \quad (22)$$

Pembentukan model GWPR diawali dengan perhitungan jarak *Euclidean* antar lokasi pengamatan, kemudian dilanjutkan dengan penentuan *bandwidth* optimum menggunakan metode AIC pada setiap lokasi. *Bandwidth* yang diperoleh selanjutnya digunakan untuk membentuk matriks pembobot spasial dalam proses estimasi parameter lokal. Pada penelitian ini, fungsi kernel *adaptive bisquare* digunakan dalam perhitungan bobot spasial.

Penggunaan perangkat lunak Octave dalam estimasi parameter GWPR menghasilkan model dengan parameter regresi yang berbeda pada setiap lokasi pengamatan. Dengan demikian, jumlah model GWPR yang diperoleh mengikuti jumlah lokasi pengamatan yang dianalisis, yaitu sebanyak 31 model. Berikut model GWPR untuk lokasi 7 (Hulu PT. BBE Berau), lokasi 16 (Kantor Gubernur Samarinda), lokasi 19 (Long Bagun Mahulu), dan lokasi 25 (Tering Kubar):

$$\hat{y}_{7t}^* = -0,5132x_{17t}^* + 0,0449x_{27t}^* - 0,0241x_{37t}^* + 0,1357x_{47t}^* + 2,839x_{57t}^*, \quad t = 1, 2, 3, 4 \quad (23)$$

$$\hat{y}_{16t}^* = -0,9614x_{116t}^* + 0,0336x_{216t}^* - 0,0098x_{316t}^* - 0,0281x_{416t}^* + 10,4646x_{516t}^*, \quad t = 1, 2, 3, 4 \quad (24)$$

$$\hat{y}_{19t}^* = 0,0043x_{119t}^* - 0,0665x_{219t}^* - 0,0023x_{319t}^* - 0,1521x_{419t}^* - 0,0003x_{519t}^*, \quad t = 1, 2, 3, 4 \quad (25)$$

$$\hat{y}_{25t}^* = 0,0284x_{125t}^* + 0,1663x_{225t}^* - 0,0963x_{325t}^* + 0,2244x_{425t}^* + 0,0016x_{525t}^*, \quad t = 1, 2, 3, 4 \quad (26)$$

TABEL VI  
STATISTIK MODEL GWPR

Statistik Model	Nilai
$F_2$	2,4097
$p_2$	$3,5856 \times 10^{-6}$ *
RMSE	1,3777

\*Signifikan pada  $\alpha = 0,05$

Pengujian kesesuaian model dilakukan untuk membandingkan performa model GWPR terhadap model global FEM. Berdasarkan hasil uji F, diperoleh nilai  $F_2$  sebesar 2,4097 dengan nilai- $p$  sebesar  $3,5856 \times 10^{-6}$ . Karena nilai- $p$  berada di bawah taraf signifikansi 5%, dapat dinyatakan bahwa terdapat perbedaan signifikan antara model FEM dan GWPR.

Hasil kesesuaian model menunjukkan bahwa model GWPR mampu memberikan representasi hubungan yang lebih baik dibandingkan model FEM karena dapat mengakomodasi variasi parameter regresi pada setiap lokasi pengamatan. Dengan demikian, pendekatan GWPR dinilai lebih sesuai untuk memodelkan data COD yang memiliki karakteristik spasial antarwilayah.

Signifikansi variabel prediktor terhadap COD pada setiap lokasi pengamatan dianalisis menggunakan uji parsial pada model GWPR. Perbedaan parameter lokal yang dihasilkan model menyebabkan pengaruh variabel prediktor bervariasi antarwilayah.

TABEL VII  
PENGELOMPOKAN LOKASI BERDASARKAN VARIABEL SIGNIFIKAN MODEL GWPR PADA TARAF SIGNIFIKANSI 5%

Klaster	Variabel Signifikan	Lokasi
1	-	2, 25
2	$x_2$	3, 6, 14, 16, 21, 26
3	$x_5$	4, 5
4	$x_2$ dan $x_5$	1, 8, 11, 12, 13, 24, 28
5	$x_2$ dan $x_4$	7, 9, 15, 17, 18, 29, 31
6	$x_4$	10, 19
7	$x_3$ dan $x_5$	20
8	$x_1$	27
9	$x_1, x_2$ dan $x_5$	22, 23
10	$x_1, x_2, x_4$ dan $x_5$	30

TABEL VIII

PENGELOMPOKAN LOKASI BERDASARKAN VARIABEL SIGNIFIKAN MODEL GWPR PADA TARAF SIGNIFIKANSI 10%

Klaster	Variabel Signifikan	Lokasi
1	$x_5$	4
2	$x_3$	25
3	$x_2$	16
4	$x_4$	19
5	$x_3$ dan $x_5$	5, 20
6	$x_2$ dan $x_5$	3, 6, 8, 11, 12, 13, 14, 21
7	$x_2$ dan $x_4$	7, 9, 15, 17, 18, 29, 31
8	$x_1$ dan $x_2$	2, 26, 27
9	$x_2, x_4$ dan $x_5$	10, 24
10	$x_1, x_2, x_4$ dan $x_5$	1, 22, 23, 28, 30

Berdasarkan hasil pada Tabel VII dan Tabel VIII, pengaruh signifikan variabel prediktor terhadap COD tidak bersifat homogen pada seluruh lokasi pengamatan. Variabel TSS ( $x_2$ ) merupakan variabel yang paling dominan karena signifikan pada 26 lokasi pengamatan. Selain itu, variabel Amoniak ( $x_5$ ) dan TDS ( $x_4$ ) juga menunjukkan pengaruh yang cukup dominan, masing-masing signifikan pada 18 dan 15 lokasi pengamatan.

Beberapa lokasi menunjukkan pengaruh simultan dari lebih dari satu variabel prediktor, terutama kombinasi antara TSS ( $x_2$ ) dan Amoniak ( $x_5$ ). Pada taraf signifikansi 5%, masih terdapat beberapa lokasi yang tidak memiliki variabel signifikan, yaitu lokasi 2 dan 25. Sementara itu, pada tingkat signifikansi 10%, setiap lokasi pengamatan memiliki sedikitnya satu variabel penjelas yang berpengaruh signifikan terhadap COD. Sebaran spasial kelompok lokasi berdasarkan kombinasi variabel signifikan pada model GWPR ditampilkan pada Gambar 3.

Perbedaan pengelompokan lokasi tersebut menunjukkan adanya heterogenitas spasial pada hubungan antara variabel prediktor dan COD. Kondisi ini mengindikasikan bahwa karakteristik kualitas air pada setiap wilayah pengamatan tidak bersifat homogen, sehingga model GWPR mampu

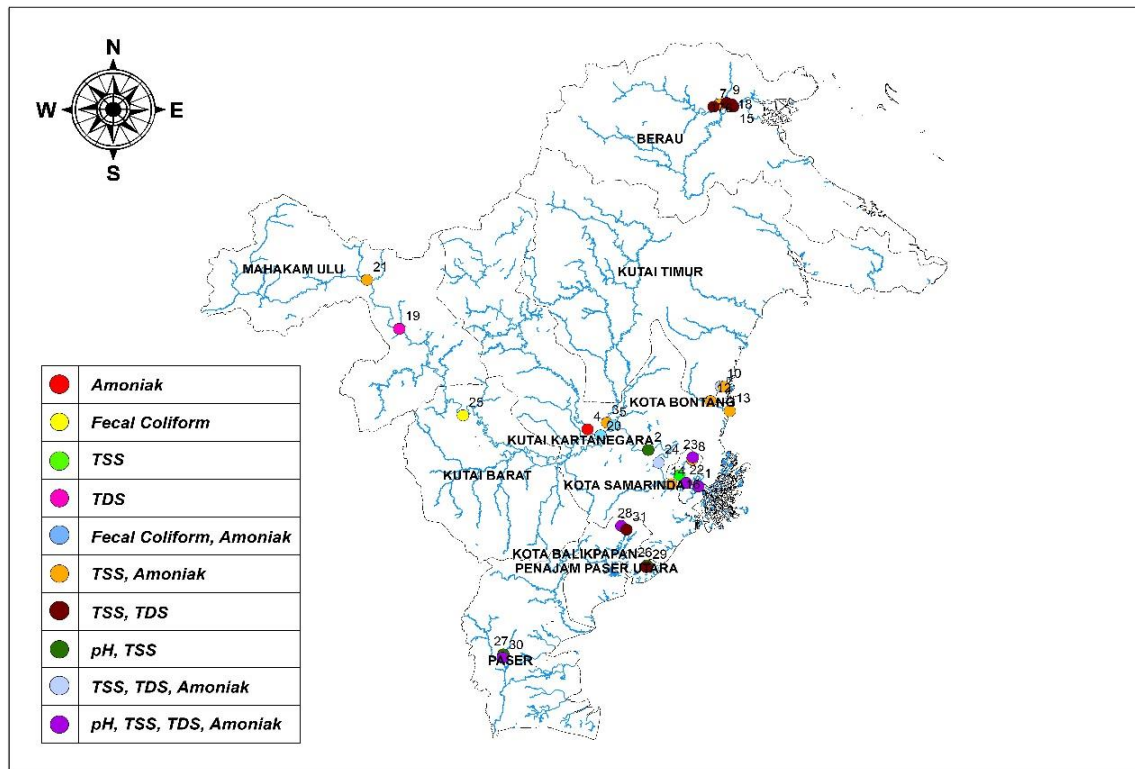
memberikan representasi hubungan lokal yang lebih baik dibandingkan model global FEM.

#### E. Perbandingan Model FEM dan GWPR

Perbandingan performa model dilakukan menggunakan nilai RMSE. Model GWPR menghasilkan nilai RMSE sebesar 1,3777 (Tabel VI), lebih kecil dibandingkan model FEM sebesar 2,182011 (Tabel V). Hasil tersebut mengindikasikan bahwa pendekatan GWPR memberikan performa pemodelan COD yang lebih baik dibandingkan FEM karena dapat mengakomodasi variasi parameter regresi pada setiap lokasi pengamatan. Dengan demikian, model GWPR dinilai lebih sesuai untuk memodelkan data COD yang memiliki karakteristik spasial antarwilayah.

Perbedaan hasil antara FEM dan GWPR menunjukkan bahwa hubungan antara COD dan variabel kualitas air tidak bersifat homogen pada seluruh lokasi pengamatan. FEM menghasilkan koefisien global yang mewakili rata-rata hubungan antarvariabel pada seluruh wilayah, sehingga hanya variabel yang memiliki pengaruh relatif konsisten yang teridentifikasi signifikan secara global. Sebaliknya, GWPR memungkinkan setiap lokasi memiliki koefisien regresi yang berbeda sehingga mampu menangkap variasi pengaruh variabel prediktor yang bersifat lokal.

Hasil GWPR menunjukkan bahwa tingkat signifikansi variabel prediktor berbeda antarwilayah, yang mengindikasikan adanya heterogenitas spasial pada faktor-faktor yang memengaruhi COD. Variabel TSS menjadi faktor yang paling dominan karena signifikan pada 26 lokasi pengamatan, sedangkan amoniak dan TDS signifikan masing-masing pada 18 dan 15 lokasi. Temuan ini menunjukkan bahwa karakteristik kualitas air pada sungai di Kalimantan Timur tidak sepenuhnya seragam, sehingga pendekatan global cenderung menyederhanakan variasi pengaruh faktor-faktor kualitas air terhadap COD yang terjadi pada masing-masing lokasi pengamatan. Kondisi tersebut menjelaskan mengapa model GWPR menghasilkan nilai RMSE yang lebih rendah dibandingkan FEM, karena model ini mampu mengakomodasi perbedaan karakteristik lokal pada setiap wilayah pengamatan.



Gambar 3. distribusi spasial kelompok lokasi berdasarkan variabel signifikan pada model GWPR dengan taraf signifikansi 10%

#### IV. KESIMPULAN

Berdasarkan hasil analisis, model FEM dan GWPR dapat digunakan untuk memodelkan COD sebagai indikator kualitas perairan sungai di Kalimantan Timur. Pada model FEM, variabel TSS dan amoniak berpengaruh signifikan terhadap COD secara global. Namun, FEM menggunakan parameter yang sama pada seluruh lokasi pengamatan sehingga belum mampu merepresentasikan variasi pengaruh antarwilayah secara spesifik.

Hasil pemodelan menggunakan GWPR menunjukkan bahwa pengaruh variabel prediktor terhadap COD berbeda pada setiap lokasi pengamatan. Variabel TSS ( $x_2$ ) menjadi faktor yang paling dominan memengaruhi COD pada sebagian besar lokasi, sedangkan variabel lainnya menunjukkan pengaruh yang bervariasi sesuai karakteristik masing-masing wilayah pengamatan.

Berdasarkan perbandingan performa model, GWPR menghasilkan nilai RMSE yang lebih kecil dibandingkan FEM, yaitu sebesar 1,3777 dibandingkan 2,182011. Hasil tersebut mengindikasikan bahwa pendekatan GWPR memberikan performa pemodelan COD yang lebih baik dibandingkan FEM karena dapat merepresentasikan heterogenitas spasial dan temporal pada data kualitas air sungai di Kalimantan Timur.

Temuan penelitian ini juga memberikan implikasi praktis bagi pengelolaan kualitas air sungai di Kalimantan Timur. Dominannya pengaruh TSS pada sebagian besar lokasi pengamatan menunjukkan bahwa upaya pengendalian sedimen dan padatan tersuspensi perlu menjadi prioritas

dalam pengelolaan kualitas perairan. Selain itu, signifikansi variabel amoniak pada sejumlah lokasi mengindikasikan pentingnya pengawasan terhadap sumber pencemar yang berasal dari limbah domestik, aktivitas peternakan, maupun kegiatan lainnya yang berpotensi meningkatkan kandungan senyawa nitrogen di perairan. Dengan adanya variasi faktor yang berpengaruh antarwilayah, strategi pengelolaan kualitas air sebaiknya disesuaikan dengan karakteristik masing-masing lokasi dan tidak hanya menggunakan pendekatan yang bersifat umum untuk seluruh wilayah pengamatan.

Penelitian selanjutnya disarankan menggunakan jumlah periode pengamatan yang lebih panjang agar pola temporal kualitas air sungai dapat dimodelkan dengan lebih baik. Selain itu, penelitian berikutnya dapat mempertimbangkan penggunaan fungsi kernel pembobot lain untuk membandingkan performa model spasial yang dihasilkan serta memasukkan variabel lingkungan tambahan, seperti curah hujan, penggunaan lahan, atau aktivitas antropogenik yang berpotensi memengaruhi kualitas air sungai.

Di sisi lain, hasil penelitian menunjukkan bahwa variabel TSS cenderung berpengaruh pada sebagian besar lokasi pengamatan. Oleh karena itu, pendekatan *Mixed Geographically Weighted Panel Regression* (MGWPR) dapat dipertimbangkan pada penelitian selanjutnya untuk mengakomodasi variabel yang bersifat global dan lokal secara simultan.

## DAFTAR PUSTAKA

- [1] U. Nations, "Ensure availability and sustainable management of water and sanitation for all," 2015.
- [2] U. N. E. Programme, "Goal 6: Clean Water and Sanitation," 2025.
- [3] M. A. Makarim, "Prevention of decreasing river water quality due to anthropogenic activities: A systematic review of water pollution on Cisadane River," *J. Mar. Probl. Threat.*, vol. 1, no. 1, pp. 44–65, 2024.
- [4] B. P. S. K. Timur, *Kalimantan Timur dalam Angka 2023*. Samarinda, Indonesia: BPS Kalimantan Timur, 2023.
- [5] D. Marganingrum and R. Noviardi, "Pencemaran air dan tanah di kawasan pertambangan batubara di PT. Berau Coal, Kalimantan Timur," *Ris. Geol. dan Pertamb.*, vol. 20, no. 1, pp. 11–20, 2010.
- [6] APHA, *Standard Methods for the Examination of Water and Wastewater*, 23rd ed. Washington DC, USA: American Public Health Association, 2017.
- [7] H. Effendi, *Telaah Kualitas Air: Bagi Pengelolaan Sumberdaya dan Lingkungan Perairan*. Yogyakarta: Kanisius, 2003.
- [8] A. S. F. R. Hufaini, Raupong, and N. Ilyas, "Regresi model data panel efek tetap dengan metode within group pada data indeks pembangunan manusia Provinsi Sulawesi Selatan," *Estimasi J. Stat. Its Appl.*, vol. 1, no. 1, pp. 10–20, 2020.
- [9] A. Indrasetyaningsih and T. K. Wasik, "Model regresi data panel untuk mengetahui faktor yang mempengaruhi tingkat kemiskinan di Pulau Madura," *J. Gaussian*, vol. 9, no. 3, pp. 355–363, 2020.
- [10] A. Dewintha, I. Yahya, and M. Ihwal, "Analisis regresi data panel pada faktor-faktor yang mempengaruhi tingkat kemiskinan di Provinsi Sulawesi Tenggara tahun 2020–2023," *Arus J. Sains dan Teknol.*, vol. 3, no. 1, pp. 83–94, 2025.
- [11] B. H. Baltagi, *Econometric Analysis of Panel Data*, 7th ed. Switzerland: Springer Nature, 2021.
- [12] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: John Wiley & Sons, 2002.
- [13] N. A. Rahman, A. A. Latif, and S. M. Said, "Geographically Weighted Regression (GWR) model for spatial analysis," *Malaysian J. Fundam. Appl. Sci.*, vol. 15, no. 3, pp. 338–343, 2019.
- [14] D. Yu, "Exploring spatiotemporally varying regressed relationships: The geographically weighted panel regression analysis," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 38, 2010.
- [15] R. Cai, D. Yu, and M. Oppenheimer, "Estimating the spatially varying responses of corn yields to weather variations using geographically weighted panel regression," *J. Agric. Resour. Econ.*, vol. 39, no. 2, pp. 230–252, 2014.
- [16] M. R. Sari, A. Hoyyi, and Suparti, "Metode geographically weighted panel regression (GWPR) untuk pemodelan data spasial panel," *J. Gaussian*, vol. 7, no. 3, pp. 239–248, 2018.
- [17] Z. N. Fauziyah, Suyitno, D. A. Nohe, M. N. Hayati, and M. Fauziyah, "Analyzing the Mahakam River water quality using the geographically weighted panel regression model," *MethodsX*, vol. 16, 2025.
- [18] Sifriyani, N. Budiantara, M. Fariz, and F. Mardianto, "Determination of the best geographic weighted function and estimation of spatio temporal model – geographically weighted panel regression using weighted least square," *MethodsX*, vol. 12, 2024.
- [19] N. M. S. Ananda, Suyitno, and M. Siringoringo, "Geographically weighted panel regression modelling of human development index data in East Kalimantan Province in 2017–2020," *J. Mat. Stat. dan Komputasi*, vol. 19, no. 2, pp. 323–341, 2023.
- [20] A. Y. Qur'ani, "Pemodelan geographically weighted regression panel (GWR-panel) sebagai pendekatan model geographically weighted regression (GWR) dengan menggunakan fixed effect model time trend," vol. 2, no. 3, pp. 1–10, 2014.
- [21] D. F. Anwar, T. Alawiyah, and Yulma, "Penentuan status mutu kualitas air Sungai Bidadari di Kelurahan Juata Kerikil Kota Tarakan," *J. Harpodon Borneo*, vol. 15, no. 2, pp. 124–132, 2023, doi: 10.35334/harpodon.v15i2.3008.
- [22] U. S. Primadigna, Suyitno, and M. Siringoringo, "Model geographically weighted Weibull regression pada indikator pencemaran air COD di daerah aliran Sungai Mahakam Kalimantan Timur," *Eksponensial*, vol. 13, no. 2, 2022, doi: 10.30872/eksponensial.v13i2.1050.
- [23] C. Hsiao, *Analysis of Panel Data*, 3rd ed. Cambridge: Cambridge University Press, 2014.
- [24] J. M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. Cambridge, MA, USA: MIT Press, 2010.
- [25] D. N. Gujarati and D. C. Porter, *Basic Econometrics*, 5th ed. New York: McGraw-Hill Education, 2009.
- [26] F. Bruna and D. Yu, "Geographically weighted panel regression and development accounting for European regions," in *Proceedings of the 6th Seminar Jean Paelinck in Spatial Econometrics*, 2016, pp. 1–20.
- [27] S. M. Meutuah, H. Yasin, and I. M. Di Asih, "Pemodelan fixed effect geographically weighted panel regression untuk indeks pembangunan manusia di Jawa Tengah," *J. Gaussian*, vol. 6, no. 2, pp. 241–250, 2017.
- [28] C. S. Purnamasari and Y. Widyaningsih, "Perbandingan performa bandwidth CV, AICc, dan BIC pada model geographically weighted regression," in *Seminar Nasional Statistika XI 2022 BT - Inferensi*, 2023, pp. 71–83.
- [29] A. Dewantoro and T. B. Sasongko, "Comparison of LSTM model performance with classical regression in predicting gaming laptop prices in Indonesia," *J. Appl. Informatics Comput.*, vol. 8, no. 1, pp. 203–212, 2024.
- [30] A. Harismahyanti A., A. Najih, A. I. Yunita, Ratmila, and Nur'eni, "Comparison of FEM-LSDV panel regression with classical panel regression models in analyzing economic growth in Indonesia," *J. Appl. Informatics Comput.*, vol. 9, no. 5, pp. 2450–2460, 2025.