

BAB I PENDAHULUAN

1.1. Latar Belakang

Analisis spasial adalah metode statistik untuk mengevaluasi pola dan hubungan data terkait lokasi geografis dengan asumsi bahwa data pada suatu lokasi dipengaruhi oleh data di lokasi sekitarnya (Wang dkk., 2024). Analisis data spasial memerlukan perhatian khusus karena perbedaan antar lokasi amatan dapat memunculkan heterogenitas spasial (Debataraja dkk., 2021). Heterogenitas spasial adalah kondisi yang terjadi ketika variabel independen yang sama menghasilkan respon berbeda di setiap lokasi dalam suatu wilayah penelitian (Caraka dan Yasin, 2017). Kondisi ini dapat menghasilkan model yang tidak akurat dan menyebabkan kesalahan dalam analisis maupun interpretasi data. Oleh karena itu, dibutuhkan pendekatan analisis yang dapat memperhitungkan pengaruh heterogenitas spasial. *Geographically Weighted Regression* (GWR) merupakan salah satu pendekatan yang mampu memodelkan variasi spasial dengan memberikan bobot berbeda pada setiap lokasi geografis. Hal ini memungkinkan penangkapan heterogenitas spasial dengan lebih akurat.

Seiring berkembangnya analisis spasial, dimensi waktu menjadi unsur penting dalam memahami aktivitas sosial dan lingkungan. Fotheringham dkk. (2015) menjelaskan bahwa selain lokasi, waktu juga merupakan faktor yang perlu dipertimbangkan. Penambahan unsur waktu pada GWR menghasilkan model yang lebih representatif, yaitu *Geographically and Temporally Weighted Regression* (GTWR). Model GTWR memungkinkan pengakomodasian variasi spasial dan temporal secara bersamaan (Oktarina dkk., 2024), serta dapat membentuk model regresi untuk setiap titik di ruang dan waktu (Yu dkk., 2024). Pada analisis GTWR, data dari lokasi dan waktu yang dekat memberikan bobot yang lebih besar terhadap titik yang dianalisis (Wang dkk., 2024). Sebaliknya, data dari lokasi dan waktu yang jauh akan mendapatkan bobot yang lebih kecil. Untuk menentukan seberapa jauh data dari lokasi dan waktu lain mempengaruhi estimasi di suatu titik, GTWR menggunakan parameter *bandwidth* (Ma dkk., 2021).

Bandwidth merupakan parameter penting yang bisa dioptimisasi untuk meningkatkan akurasi model serta efisiensi komputasi (Nishi dan Asami, 2024). *Bandwidth* yang lebih kecil menekankan pada data terdekat dalam ruang dan waktu sehingga efektif untuk menangkap perubahan pada area sekitar meskipun berisiko *overfitting* jika terlalu kecil (Yu dkk., 2020). Sebaliknya, *bandwidth* yang lebih besar mencakup pola yang lebih luas dengan melibatkan data yang lebih jauh, tapi bisa saja mengabaikan informasi spesifik pada area sekitar (Li dkk., 2020). Menurut Wu dkk. (2019), penyesuaian *bandwidth* untuk setiap variabel dapat memaksimalkan kemampuan model dalam menangkap hubungan spesifik antara variabel dan kejadian yang diamati, baik pada skala terbatas maupun lebih luas, sehingga memungkinkan tercapainya *bandwidth* optimum untuk masing-masing variabel. Hal ini tidak tercapai dalam model GTWR, yang hanya menggunakan satu *bandwidth* untuk semua variabel independen, sehingga membatasi kemampuannya dalam

menangkap heterogenitas spasial dan temporal yang kompleks. Oleh karena itu, untuk menghasilkan analisis yang lebih akurat dan komprehensif, dibutuhkan model yang memungkinkan penggunaan *bandwidth* berbeda pada setiap variabel.

Multi-Scale Geographically and Temporally Weighted Regression (MSGTWR) merupakan *Generalized Additive Model* (GAM) yang dapat mengatasi keterbatasan model GTWR karena memungkinkan penggunaan *bandwidth* berbeda untuk setiap variabel (Wu dkk., 2019). Model ini hadir dengan mempertimbangkan efek spasial, temporal, dan *influence scope of variables* atau pengaruh cakupan variabel (Cheng dkk., 2024). *Influence scope* yang dipertimbangkan dalam model MSGTWR memungkinkan representasi yang lebih akurat terhadap heterogenitas antar variabel sehingga menghasilkan analisis yang lebih tepat mengenai dampak faktor-faktor tertentu (Cheng dkk., 2024). Cakupan pengaruh ini disesuaikan menggunakan *bandwidth* berbeda untuk setiap variabel sehingga memungkinkan untuk mengurangi bias dan risiko *overfitting* serta memberikan kinerja model yang lebih baik (Liu dkk., 2021).

Penelitian dengan MSGTWR pertama kali dilakukan oleh Wu dkk. (2019) dan menunjukkan skala fleksibel untuk mendeteksi *non-stasioneritas* spatiotemporal dalam studi *real-estate*. Kemudian Liu dkk. (2021) juga mendapati MSGTWR efektif dalam mengatasi variabilitas spasial dan temporal yang kompleks dalam pemetaan polusi udara berbasis satelit. Penelitian lain juga pernah dilakukan oleh Cheng dkk. (2024) dalam bidang transportasi yang menunjukkan bahwa model MSGTWR memiliki efek regresi yang lebih baik dibandingkan model GTWR klasik. MSGTWR lebih unggul dalam analisis data spatiotemporal yang kompleks karena memungkinkan variabel-variabel berinteraksi dalam ruang dan waktu serta skala yang berbeda-beda. Model ini telah diterapkan dalam bidang *real-estate*, geospasial, dan transportasi. Oleh karena itu, dalam penelitian ini, MSGTWR diterapkan dalam bidang epidemiologi untuk menganalisis data kasus Demam Berdarah Dengue (DBD) di Sulawesi Selatan.

Data kasus DBD di Sulawesi Selatan menunjukkan penurunan signifikan setelah mencapai tingkat kejadian tertinggi pada tahun 2016, yaitu 89 kasus per 100.000 penduduk. Meskipun sempat menurun, angka ini kembali meningkat menjadi 39 kasus per 100.000 penduduk pada tahun 2021 (Kemenkes, 2021). Perubahan ini mencerminkan pola penyebaran yang tidak stabil, sehingga menimbulkan pertanyaan mengenai faktor-faktor yang memengaruhi perubahan tersebut. DBD merupakan penyakit menular yang sangat dipengaruhi oleh faktor spasial dan temporal, seperti kondisi lingkungan, iklim, serta pola aktivitas manusia yang beragam di setiap wilayah. Studi terhadap kasus DBD menjadi penting karena pola penyebarannya bersifat heterogen, baik secara geografis maupun dari waktu ke waktu. Dalam penelitian ini, beberapa variabel digunakan untuk menganalisis faktor-faktor yang memengaruhi penyebaran DBD, di antaranya persentase penduduk miskin, kepadatan penduduk, akses sanitasi layak, jumlah anak di bawah 15 tahun, dan ketersediaan sarana kesehatan. Persentase penduduk miskin mencerminkan kondisi sosial-ekonomi, sedangkan kepadatan penduduk dan akses sanitasi layak menggambarkan karakteristik lingkungan fisik. Jumlah anak di bawah 15 tahun

berhubungan dengan aspek demografi, dan ketersediaan sarana kesehatan mencerminkan akses terhadap layanan kesehatan.

Faktor-faktor ini berkaitan erat dengan konsep determinan sosial kesehatan sebagaimana dijelaskan oleh Marmot dan Wilkinson (2006) mencakup kondisi sosial-ekonomi, lingkungan fisik, jaringan sosial dan demografi, serta akses layanan kesehatan yang tersebar di berbagai tingkat, mulai dari individu hingga komunitas dan kebijakan sistemik. Kombinasi faktor-faktor ini mencerminkan heterogenitas spasial dan temporal yang kompleks, sehingga diperlukan analisis mendalam untuk memahami dinamika penyebaran DBD secara akurat dan terarah. Oleh karena itu, untuk memahami dinamika penyebaran kasus DBD secara lebih mendetail, dilakukan penelitian berjudul **“Pemodelan *Multi-Scale Geographically and Temporally Weighted Regression* pada Kasus Demam Berdarah di Sulawesi Selatan”**. Pendekatan MSGTWR mampu menangkap perbedaan pengaruh variabel pada berbagai skala geografis dan temporal, sehingga diharapkan mampu memberikan wawasan lebih mendalam tentang pola penyebaran DBD di Sulawesi Selatan dan menjadi dasar untuk perencanaan intervensi yang lebih efektif.

1.2. Batasan Penelitian

Batasan dari penelitian ini adalah sebagai berikut:

1. Menggunakan data tingkat kejadian DBD di Sulawesi Selatan tahun 2016-2021.
2. Fungsi pembobot yang digunakan adalah *fixed* kernel gaussian dan eksponensial dengan membandingkan nilai *AICc*.
3. Nilai ambang batas yang digunakan adalah 10^{-5}
4. Penentuan *bandwidth* optimum menggunakan nilai *score of change (SOC)* dari *residual sum of squares (RSS)*.

1.3. Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh model MSGTWR pada data kasus DBD di Sulawesi Selatan tahun 2016-2021
2. Memperoleh faktor-faktor yang secara signifikan memengaruhi DBD di Sulawesi Selatan tahun 2016-2021

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Memberikan informasi kepada pihak berkepentingan mengenai faktor yang memengaruhi kasus DBD di Sulawesi Selatan tahun 2016-2021
2. Menambah dan mengembangkan wawasan terkait pemodelan MSGTWR.

1.4. Teori

1.4.1. Regresi Linear Berganda

Regresi merupakan metode statistika yang digunakan untuk menganalisis serta memodelkan hubungan antara variabel respon dengan variabel prediktor (Efendi dkk., 2020). Tujuan model ini adalah untuk menentukan pengaruh antar variabel atau hubungan fungsional antara dua atau lebih variabel prediktor dengan variabel respon

(Ikhsanudin dan Pasaribu, 2023). Secara umum, model regresi linear dinyatakan dalam Persamaan (1).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (1)$$

Model regresi linier secara umum dapat dinyatakan dalam bentuk matriks pada Persamaan (2).

$$Y = X\beta + \varepsilon \quad (2)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$i = 1, 2, \dots, n$; $k = 1, 2, 3, \dots, p$

dengan

- n : Jumlah pengamatan
- p : Jumlah prediktor
- y_i : Nilai observasi variabel respon pada pengamatan ke- i
- x_{ik} : Nilai observasi variabel prediktor ke- k pada pengamatan ke- i
- β_0 : Nilai intersep model regresi
- β_k : Koefisien regresi variabel prediktor ke- k
- ε_i : Residual pada pengamatan ke- i dengan asumsi independen, identik, dan berdistribusi normal, dengan *mean* nol dan *variansi* konstan σ^2 (Fotheringham dkk., 2002)

1.4.2. Pengujian Multikolinearitas

Multikolinearitas adalah kondisi ketika dua atau lebih variabel independen atau prediktor dalam model memiliki korelasi yang sangat tinggi satu sama lain. Penting untuk menguji keberadaan multikolinearitas di antara variabel-variabel independen sebelum melakukan analisis regresi agar hasil penjelasan dari variabel-variabel tetap akurat (Cheng dkk., 2024). Dalam penelitian ini, multikolinearitas diuji menggunakan *variance inflation factors* (VIFs). Nilai VIF berkisar dari 1 hingga positif tak terhingga. Nilai yang mendekati 1 menunjukkan rendahnya multikolinearitas. Kemudian nilai VIF yang lebih besar dari 10 mengindikasikan adanya multikolinearitas yang serius di antara variabel-variabel dalam model regresi (O'Driscoll dan Ramirez, 2015). VIF untuk variabel independen ke- k dapat dihitung menggunakan Persamaan (1)

$$VIF_k = \frac{1}{1 - R_k^2} \quad (3)$$

dengan R_k^2 koefisien determinasi dari variabel independen x_k terhadap variabel independen lainnya dalam analisis regresi.

1.4.3. Pengujian Heterogenitas Spasial

Heterogenitas spasial adalah kondisi yang terjadi ketika variabel independen yang sama menghasilkan respon berbeda di setiap lokasi dalam suatu wilayah penelitian (Caraka dan Yasin, 2017). Kondisi ini muncul karena perbedaan karakteristik yang ada di setiap wilayah. Heterogenitas spasial dapat diketahui dengan statistik uji

Breusch-Pagan (BP). Suatu data dikatakan memiliki heterogenitas spasial jika nilai $BP > \chi^2_{(\alpha;p)}$ atau $p - value < \alpha$, dengan hipotesis pengujian sebagai berikut:

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (tidak terdapat heterogenitas spasial)

$H_1 : \text{ada } \sigma_i^2 \neq \sigma_j^2 \text{ dengan } i \neq j$ (terdapat heterogenitas spasial)

Nilai dari uji BP dihitung menggunakan Persamaan (4)

$$BP = \left(\frac{1}{2}\right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \quad (4)$$

dengan

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right)$$

e_i^2 adalah residual observasi ke- i berdasarkan hasil estimasi regresi menggunakan *Ordinary Least Squares* (OLS) dan $\mathbf{Z}$ adalah matriks variabel prediktor berukuran $n \times (p + 1)$.

1.4.4. Fungsi Pembobot

Pembobot adalah nilai atau faktor yang digunakan untuk memberikan tingkat kepentingan atau pengaruh yang berbeda pada data atau pengamatan. Pemilihan pembobot sangat penting karena berfungsi untuk menjelaskan pengaruh ketetanggaan antar daerah (Fotheringham dkk., 2009). Pembobot yang dihasilkan berasal dari fungsi kernel. Ada beberapa jenis fungsi kernel yang diantaranya adalah kernel *gaussian*, *bi-square*, *tricube*, dan *exponential* (Anselin dan Rey, 2010). Fungsi kernel yang akan digunakan dalam pembobotan ini ditunjukkan dalam Persamaan (5) dan (6)

1. Fungsi Kernel *Gaussian*

$$w_{ij} = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}^{ST}}{b'} \right)^2 \right] \quad (5)$$

2. Fungsi Kernel *Exponential*

$$w_{ij} = \exp \left(-\frac{d_{ij}^{ST}}{b'} \right) \quad (6)$$

dengan

w_{ij} : Merepresentasikan bobot antara titik pengamatan i dan j

d_{ij}^{ST} : Jarak spasio-temporal antara titik pengamatan i dan j

b' : *Bandwidth* spasio-temporal

Kemudian, jarak spasio-temporal antara titik pengamatan i dan j dihitung berdasarkan Persamaan (7)

$$\frac{(d_{ij}^{ST})^2}{\lambda} = \left[(u_i - u_j)^2 + (v_i - v_j)^2 \right] + \tau (t_i - t_j)^2 \quad (7)$$

dengan menetapkan parameter penyeimbang spasial $\lambda = 1$ dan τ adalah parameter rasio spasio-temporal yang memainkan peran penting dalam membentuk bobot. Pada dasarnya, efek utama dari τ adalah memperbesar atau memperkecil efek jarak temporal agar sesuai dengan jarak spasial (Huang dkk., 2010).

1.4.5. *Bandwidth Optimum*

Bandwidth adalah parameter penting yang mengatur pengaruh spasial dalam model yang bisa ditransformasi serta dioptimisasi untuk meningkatkan akurasi model serta efisiensi komputasi (Nishi dan Asami, 2024). Proses pemilihan *bandwidth* optimum untuk model *Geographically and Temporally Weighted Regression* (GTWR) berbeda dengan model *Multi-Scale Geographically and Temporally Weighted Regression* (MSGTWR). Model GTWR mencari satu *bandwidth* optimum untuk seluruh variabel, sedangkan MSGTWR mencari *bandwidth* optimum untuk setiap variabel.

1.4.5.1 *Bandwidth optimum model GTWR*

Model GTWR merupakan model yang dirancang untuk menangkap pengaruh variabel berbeda di berbagai lokasi geografis dan titik waktu dengan menerapkan bobot yang berbeda pada data terdekat, baik secara spasial maupun temporal. Untuk mencapai performa optimal, GTWR berfokus pada pencarian satu set *bandwidth* optimal yang pada studi ini didapatkan dengan meminimalisasi nilai *AICc* menggunakan metode *grid search*. Metode ini akan secara sistematis menguji berbagai kombinasi parameter dalam rentang yang ditentukan (Shams dkk., 2023). Rentang yang ditentukan dalam konteks ini merujuk pada batas atau cakupan nilai-nilai potensial yang akan diuji untuk parameter *bandwidth* spasial (h) dan rasio spasio-temporal (τ) dengan τ dibatasi hingga 0.9.

1.4.5.2 *Bandwidth optimum model MSGTWR*

Model MSGTWR memungkinkan pemilihan *bandwidth* yang spesifik untuk setiap variabel. Dalam model ini, *bandwidth* dihitung dengan menginisiasi parameter awal menggunakan model GTWR, termasuk penentuan nilai awal untuk koefisien regresi dan *bandwidth*. Setelah itu indeks iterasi (k) diatur menjadi 1 yang menunjukkan bahwa proses dimulai dengan variabel pertama. Langkah selanjutnya memeriksa kondisi $k \leq p$, dengan p adalah banyaknya variabel. Jika k lebih besar dari p , proses akan selesai. Jika tidak, maka lanjut ke langkah berikutnya yaitu menghitung residual awal (ε) berdasarkan Persamaan (8)

$$\varepsilon = y - \sum_{k=1}^p f_k \quad (8)$$

dengan y adalah nilai observasi sebenarnya dan f_k adalah nilai regresi dari variabel ke- k . Selanjutnya residual (ε) ditambahkan dengan nilai regresi variabel ke- k (f_k).

Hasil dari $\varepsilon + f_k$ digunakan sebagai variabel respon untuk melakukan regresi dengan GTWR. Hal ini bertujuan untuk mendapatkan *bandwidth* optimal untuk variabel ke- k dan memperbarui estimasi f_k dengan hasil baru. Setelah regresi, *bandwidth* optimal (b_j) dan estimasi parameter (β_k, b_{ij}) untuk variabel ke- k diperoleh. Selanjutnya nilai k dinaikkan sebesar $1(k+1)$ untuk lanjut ke variabel berikutnya. Setelah nilai k diperbarui, proses kembali memeriksa kondisi $k \leq p$ untuk menentukan ada tidaknya variabel lain yang perlu dihitung. Setelah semua variabel dihitung, model diperiksa untuk melihat konvergensinya, yaitu kondisi yang memperlihatkan estimasi parameter dan *bandwidth* tidak berubah signifikan. Jika

sudah mencapai konvergensi, maka proses selesai atau dihentikan. Namun jika belum, proses akan dimulai kembali dari menghitung residual (ε) sampai tercapai konvergensi (Cheng dkk., 2024; Fotheringham dkk., 2017; Wu dkk., 2019).

Kriteria konvergensi yang umum digunakan adalah skor perubahan (SOC) dari jumlah kuadrat residual (RSS), yang dinyatakan dalam Persamaan (9)

$$SOC_{RSS} = \frac{|RSS_{new} - RSS_{old}|}{RSS_{new}} \quad (9)$$

dengan RSS_{old} menunjukkan RSS dari siklus sebelumnya dan RSS_{new} menunjukkan RSS dari siklus saat ini. Jika nilainya kurang dari parameter yang diberikan, algoritma menentukan bahwa proses iteratif memenuhi kondisi konvergensi; jika tidak, dianggap tidak memenuhi kondisi konvergensi. Nilai ambang batas yang dipilih untuk studi ini adalah 10^{-5} . Nilai ini dipilih berdasarkan strategi efisiensi komputasi MGTWR oleh (Wu dkk., 2019) yang mirip dengan strategi untuk efisiensi komputasi dalam Lu dkk. (2017) dan Fotheringham dkk. (2017). Rumus untuk menghitung RSS ditunjukkan dalam Persamaan (10)

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

dengan $\hat{y}_i$ dan y_i masing-masing menunjukkan nilai pendugaan dan nilai aktual. RSS mengambil nilai dari 0 hingga positif tak terhingga. Semakin besar nilai RSS , semakin buruk kecocokannya.

1.4.6. Geographically and Temporally Weighted Regression

Geographically and Temporally Weighted Regression (GTWR) merupakan model untuk menangani heterogenitas spasial dan temporal secara bersamaan (Yang dkk., 2022). GTWR terbentuk ketika unsur waktu dimasukkan ke dalam model *Geographically Weighted Regression* (GWR). Model ini direpresentasikan oleh Persamaan (11)

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (11)$$

$i = 1, 2, \dots, n$; dan $k = 1, 2, \dots, p$

dengan

- y_i : Variabel dependen lokasi amatan ke- i
- β_0 : Nilai konstanta atau *intercept* model regresi
- (u_i, v_i, t_i) : Koordinat spasio-temporal amatan ke- i
- $\beta_0(u_i, v_i, t_i)$: Intersep pada lokasi amatan u_i, v_i dan waktu t_i
- $\beta_k(u_i, v_i, t_i)$: Koefisien regresi variabel independen ke- k pada lokasi amatan u_i, v_i dan waktu t_i
- x_{ik} : Variabel independen ke- k pada lokasi amatan ke- i
- ε_i : Residual pengamatan pada lokasi ke- i

Koefisien regresi $\hat{\beta}(u_i, v_i, t_i)$ dihitung menggunakan *Weighted Least Squares* (WLS), sesuai dengan Persamaan (12)

$$\hat{\beta}(u_i, v_i, t_i) = [X^T W(u_i, v_i, t_i) X]^{-1} X^T W(u_i, v_i, t_i) Y \quad (12)$$

dengan

X : Matriks observasi variabel independen

Y : Vektor observasi variabel dependen

$W(u_i, v_i, t_i)$ terdiri dari matriks diagonal bobot ruang w_{ij} yang diekspresikan sebagai berikut:

$$W(u_i, v_i, t_i) = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{ij} \end{bmatrix}$$

dengan w_{ij} merepresentasikan bobot antara titik sampel i dan j .

1.4.7. Multi-Scale Geographically and Temporally Weighted Regression

Model *Multi-Scale Geographically and Temporally Weighted Regression* (MSGTWR) dibangun atas dasar model GWR klasik dan menyempurnakan keterbatasan model GTWR. Model ini memungkinkan representasi heterogenitas spasial yang lebih baik di antara variabel. Model MSGTWR dapat dinyatakan dalam Persamaan (13).

$$y_i = \beta_{0,b0}(u_i, v_i, t_i) + \sum_{k=1}^p \beta_{k,bk}(u_i, v_i, t_i)x_{ik} + \varepsilon_i \quad (13)$$

$i = 1, 2, \dots, n$; dan $k = 1, 2, \dots, p$

dengan

- y_i : Variabel dependen lokasi amatan ke- i
- (u_i, v_i, t_i) : Koordinat spasio-temporal amatan ke- i
- $\beta_{0,b0}(u_i, v_i, t_i)$: Intersep dan *bandwidth* intersep ($b0$) pada lokasi amatan u_i, v_i dan waktu t_i
- $\beta_{k,bk}(u_i, v_i, t_i)$: Koefisien regresi variabel independen ke- k dan *bandwidth* variabel independen (bk) pada lokasi amatan u_i, v_i dan waktu t_i
- x_{ik} : Variabel independen ke- k pada lokasi amatan ke- i
- ε_i : Residual pengamatan pada lokasi ke- i

Adapun bentuk vektor dari MSGTWR ditunjukkan pada Persamaan (14) (Wu dkk., 2019).

$$Y = \beta_{bwt0_s0} + \beta_{bwt1_s1} \odot X_1 + \beta_{bwt2_s2} \odot X_2 + \dots + \beta_{bwt p_sp} \odot X_p + \varepsilon \quad (14)$$

dengan $\odot$ menotasikan *element-wise multiplication* dari dua vektor, dan $\beta_{bwt p_sp}$ mewakili koefisien estimasi dari variabel ke- p berdasarkan *bandwidth* spasial ($bwsp$) dan *bandwidth* temporal ($bwtp$) tertentu.

MSGTWR memungkinkan pemilihan *bandwidth* yang bervariasi sesuai jumlah variabel independen. Oleh karena itu, pendekatan dengan WLS tidak berlaku untuk kalibrasi MSGTWR. Pada dasarnya, MSGTWR adalah *generalized additive model* (GAM) sehingga algoritma *back-fitting* menjadi pilihan terbaik untuk kalibrasi (Wu dkk., 2019). Adapun langkah-langkah spesifik dalam proses ini ditunjukkan pada Gambar 1.

Langkah 1. Proses dimulai dengan membuat pendugaan awal untuk semua item adaptif, yaitu $\beta_{bwt0_s0}, \beta_{bwt1_s1} \odot X_1, \beta_{bwt2_s2} \odot X_2, \dots, \beta_{bwt p_sp} \odot X_p$, dan menotasikan setiap item adaptif sebagai $\hat{f}_0^{(0)}, \hat{f}_1^{(0)}, \hat{f}_2^{(0)}, \dots, \hat{f}_p^{(0)}$, sehingga

$$\boldsymbol{\varepsilon}^{(0)} = \mathbf{y} - \hat{\mathbf{f}}_0^{(0)} - \hat{\mathbf{f}}_1^{(0)} - \hat{\mathbf{f}}_2^{(0)} - \dots - \hat{\mathbf{f}}_p^{(0)} \quad (15)$$

Langkah 2. Menetapkan nilai ambang batas sebagai kriteria penghentian algoritma *back-fitting*.

Langkah 3. Setiap *term* kemudian diregresikan terhadap variabel yang sesuai (Intersep menggunakan vektor satuan $\mathbf{I}$, dan term ke- k menggunakan $\mathbf{x}_k$ berdasarkan residual yang diperbarui). *Term* eror $\boldsymbol{\varepsilon}_p^{(l)}$ diperbarui untuk setiap *term* yang sesuai, dengan $\boldsymbol{\varepsilon}_0^{(l+1)} = \boldsymbol{\varepsilon}_p^{(l)}$.

Langkah 4. Membandingkan nilai ambang batas dengan kriteria konvergensi skor perubahan (*SOC*). Koefisien estimasi dari iterasi terakhir adalah koefisien estimasi akhir. Oleh karena itu, nilai *fitting* akhir dari $\mathbf{Y}$ ditunjukkan pada Persamaan (16).

$$\hat{\mathbf{Y}}^{(l+1)} = \hat{\mathbf{f}}_0^{(l+1)} + \hat{\mathbf{f}}_1^{(l+1)} + \hat{\mathbf{f}}_2^{(l+1)} + \dots + \hat{\mathbf{f}}_p^{(l+1)} \quad (16)$$

	$\mathbf{Y}$	=	$\boldsymbol{\beta}_{bwt1_s1} \mathbf{X}_1$	+	$\boldsymbol{\beta}_{bwt2_s2} \mathbf{X}_2$	+	$\dots$	+	$\boldsymbol{\beta}_{bwt p_sp} \mathbf{X}_p$	+	$\boldsymbol{\varepsilon}$
			$f_1(\mathbf{X}_1)$		$f_2(\mathbf{X}_2)$				$f_p(\mathbf{X}_p)$		Model Hasil Iterasi
Perkiraan Awal			$\hat{f}_1^{(0)} \text{ }_{bwt1^{(0)}}$		$\hat{f}_2^{(0)} \text{ }_{bwt2^{(0)}}$				$\hat{f}_p^{(0)} \text{ }_{bwt p^{(0)}}$		$\hat{\mathbf{y}}^{(0)} = \hat{f}_1^{(0)} + \hat{f}_2^{(0)} + \dots + \hat{f}_p^{(0)}$
Iterasi Pertama			$\hat{f}_1^{(1)} = S_1^1[y - \hat{f}_2^{(0)} - \dots - \hat{f}_p^{(0)}] \text{ }_{bwt1^{(1)}}$		$\hat{f}_2^{(1)} = S_2^1[y - \hat{f}_1^{(1)} - \hat{f}_3^{(0)} - \dots - \hat{f}_p^{(0)}] \text{ }_{bwt2^{(1)}}$				$\hat{f}_p^{(1)} = S_p^1[y - \hat{f}_1^{(1)} - \hat{f}_2^{(1)} - \dots - \hat{f}_{p-1}^{(1)}] \text{ }_{bwt p^{(1)}}$		$\hat{\mathbf{y}}^{(1)} = \hat{f}_1^{(1)} + \hat{f}_2^{(1)} + \dots + \hat{f}_p^{(1)}$
			$\vdots$		$\vdots$				$\vdots$		$\vdots$
			$\hat{f}_1^{(l-1)} = S_1^{l-1}[y - \hat{f}_2^{(l-2)} - \dots - \hat{f}_p^{(l-2)}] \text{ }_{bwt1^{(l-1)}}$		$\hat{f}_2^{(l-1)} = S_2^{l-1}[y - \hat{f}_1^{(l-1)} - \hat{f}_3^{(l-2)} - \dots - \hat{f}_p^{(l-2)}] \text{ }_{bwt2^{(l-1)}}$				$\hat{f}_p^{(l-1)} = S_p^{l-1}[y - \hat{f}_1^{(l-1)} - \hat{f}_2^{(l-1)} - \dots - \hat{f}_{p-1}^{(l-1)}] \text{ }_{bwt p^{(l-1)}}$		$\hat{\mathbf{y}}^{(l-1)} = \hat{f}_1^{(l-1)} + \hat{f}_2^{(l-1)} + \dots + \hat{f}_p^{(l-1)}$
Iterasi ke- l			$\hat{f}_1^{(l)} = S_1^l[y - \hat{f}_2^{(l-1)} - \dots - \hat{f}_p^{(l-1)}] \text{ }_{bwt1^{(l)}}$		$\hat{f}_2^{(l)} = S_2^l[y - \hat{f}_1^{(l)} - \hat{f}_3^{(l-1)} - \dots - \hat{f}_p^{(l-1)}] \text{ }_{bwt2^{(l)}}$				$\hat{f}_p^{(l)} = S_p^l[y - \hat{f}_1^{(l)} - \hat{f}_2^{(l)} - \dots - \hat{f}_{p-1}^{(l)}] \text{ }_{bwt p^{(l)}}$		$\hat{\mathbf{y}}^{(l)} = \hat{f}_1^{(l)} + \hat{f}_2^{(l)} + \dots + \hat{f}_p^{(l)}$

Gambar 1. Proses *back-fitting* estimasi parameter MSGTWR

dengan

- $\boldsymbol{\beta}_{bwt p_sp} \mathbf{X}_p$: Item adaptif yang mewaliki koefisien regresi untuk tiap kombinasi *bandwidth* spasial temporal variabel $\mathbf{X}_p$
- $f_p(\mathbf{X}_p)$: Fungsi atau kontribusi yang menghubungkan variabel independen $\mathbf{X}_p$ dan variabel dependen $\mathbf{Y}$
- $\hat{f}_p^{(0)}$: Nilai prediksi awal untuk fungsi $f_p(\mathbf{X}_p)$ pada iterasi ke-0 (pertama)
- $\hat{f}_p^{(l)}$: Nilai prediksi terbaru untuk fungsi $f_p(\mathbf{X}_p)$ pada iterasi ke- l
- S_p^l : Faktor penyesuaian untuk variabel $\mathbf{X}_p$ pada iterasi ke- l
- $\hat{\mathbf{y}}^l$: Nilai prediksi dari model pada iterasi ke- l
- $\mathbf{y}$: Variabel dependen
- $\boldsymbol{\varepsilon}$: Residual pengamatan

1.4.8. Evaluasi Model

Evaluasi model akan dilakukan dengan melihat nilai *Akaike Information Criterion Corrected* (AICc) dan *Root-Mean-Squared Error* (RMSE). Nilai AICc dihitung berdasarkan Persamaan (17) (Brewer dkk., 2016; McMillen, 2004):

$$AICc = n \cdot \ln(RSS) + 2(p + 1) - n \ln(n) + \left[\frac{2p(p + 1)}{n - p - 1} \right] \quad (17)$$

dengan n menunjukkan jumlah sampel dan p menunjukkan jumlah variabel. Nilai AICc berkisar dari 0 hingga tak terhingga secara positif, dan semakin rendah nilai AICc, semakin baik kesesuaian modelnya. Kemudian nilai RMSE dihitung berdasarkan Persamaan (18) (Hodson, 2022; Li dkk., 2023).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (18)$$

dengan $\hat{y}_i$ serta y_i masing-masing menunjukkan nilai yang diprediksi dan nilai aktual. Semakin kecil nilai RMSE yang dihasilkan suatu model maka semakin akurat model tersebut.

1.4.9. Demam Berdarah Dengue

Demam berdarah dengue (DBD) adalah infeksi akut yang disebabkan oleh virus *Flaviviridae* dengan empat serotipe (DEN-1, DEN-2, DEN-3, dan DEN-4). Penyakit ini menyebar melalui gigitan nyamuk *Aedes aegypti* yang membawa virus dengue dalam tubuhnya (Husin dkk., 2024). Gejala DBD sangat bervariasi, mulai dari tanpa gejala hingga bentuk yang lebih parah dengan komplikasi serius yang dapat berakibat fatal, seperti syok dengue dan kegagalan multiorgan (Kularatne dan Dalugama, 2022). Di seluruh dunia, DBD telah menjadi masalah kesehatan endemik di lebih dari seratus negara dengan jumlah kasus tertinggi setiap tahunnya berada di Asia (Nugraheni dkk., 2023; Wang dkk., 2020). Organisasi Kesehatan Dunia (WHO) mengidentifikasi Indonesia sebagai negara dengan tingkat kejadian DBD tertinggi di wilayah Asia Tenggara (Zulkarnain dan Surjadi, 2024). Di Sulawesi Selatan sendiri, tingkat kejadian DBD tertinggi terjadi pada tahun 2016 dengan 89 kasus per 100.000 penduduk (Kemenkes, 2021).

Kasus DBD dipengaruhi oleh beberapa faktor seperti tingkat kepadatan penduduk, kemiskinan, akses sanitasi layak, ketersediaan sarana kesehatan, dan jumlah anak di bawah 15 tahun. Hal ini karena daerah dengan kepadatan penduduk yang tinggi cenderung memiliki lebih banyak tempat yang dapat menjadi sarang bagi nyamuk *Aedes aegypti* (Sobari dkk., 2023). Kemudian akses sanitasi layak juga dapat memengaruhi demam berdarah karena daerah kumuh yang cenderung diperparah oleh faktor kemiskinan sering kali tidak memiliki sanitasi yang baik sehingga berpotensi meningkatkan pertumbuhan dan perkembangbiakan nyamuk *Aedes aegypti* (Safriyan dan Prastowo, 2024). Ketersediaan sarana kesehatan yakni rumah sakit, rumah sakit bersalin, poliklinik, puskesmas, puskesmas pembantu, dan apotek juga dapat memengaruhi tingkat kejadian DBD karena mendukung deteksi dan diagnosis cepat, perawatan yang tepat, serta edukasi pencegahan kepada masyarakat sehingga menjadi salah satu faktor yang signifikan dalam upaya pencegahan penyakit tersebut (Jusuf dkk., 2024). Adapun jumlah anak dengan umur di bawah 15 tahun juga dianggap mempengaruhi tingkat kejadian DBD karena faktor imun yang lemah membuat anak-anak rentan terkena infeksi dengue (Arfan dkk., 2024).

BAB II METODE PENELITIAN

2.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari Dinas Kesehatan Provinsi Sulawesi Selatan dan publikasi dari Badan Pusat Statistik (BPS) periode 2016-2021. Unit pengamatan yang digunakan dalam penelitian ini adalah 24 kabupaten/kota di Sulawesi Selatan.

2.2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas variabel respon kasus DBD dan lima variabel prediktor yang diduga memengaruhi DBD di kabupaten/kota Provinsi Sulawesi Selatan. Variabel yang digunakan disajikan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Satuan	Definisi
Y	Angka kejadian Demam Berdarah (DBD)	Jumlah orang per 100.000 penduduk	Jumlah kasus DBD yang dilaporkan di suatu wilayah dalam periode waktu tertentu dan dinyatakan per 100.000 penduduk
X_1	Kepadatan penduduk	Jiwa/ km^2	Jumlah penduduk yang tinggal di suatu area per satuan luas, kilometer persegi dan digunakan untuk mengukur seberapa ramai atau padatnya suatu wilayah berdasarkan populasi yang ada
X_2	Persentase penduduk miskin	%	Proporsi penduduk di suatu wilayah yang hidup di bawah garis kemiskinan, dinyatakan dalam bentuk persentase dari total penduduk
X_3	Sarana kesehatan	Unit	Fasilitas dan infrastruktur yang disediakan untuk memberikan layanan kesehatan kepada masyarakat, termasuk rumah sakit, puskesmas, klinik, dan posyandu
X_4	Akses sanitasi layak	%	Ketersediaan fasilitas sanitasi yang memenuhi standar tertentu untuk menjaga kesehatan masyarakat, seperti toilet yang aman dan higienis
X_5	Jumlah anak berumur kurang dari 15 tahun	Jiwa	Jumlah individu yang berusia kurang dari 15 tahun di suatu wilayah

2.3. Metode Analisis

Metode analisis yang digunakan dalam penelitian ini adalah *Multi-Scale Geographically and Temporally Weighted Regression*. Pengolahan data dilakukan menggunakan bahasa pemrograman Python dengan bantuan *Visual Studio Code* sebagai *Integrated Development Environment* (IDE). Tahapan analisis data yang dilakukan adalah sebagai berikut:

1. Melakukan eksplorasi data dengan memvisualisasikan peta sebaran angka kejadian DBD di setiap kabupaten/kota.
2. Melakukan perhitungan *Variance Inflation Factor* (VIF) untuk setiap variabel independen dengan tujuan mengidentifikasi dan mengevaluasi tingkat multikolinearitas yang mungkin terjadi di antara variabel-variabel tersebut menggunakan Persamaan (3).
3. Mengidentifikasi efek heterogenitas spasial dengan uji *Breusch-Pagan* menggunakan persamaan (4) dan efek heterogenitas temporal menggunakan *boxplot*.
4. Melakukan pemodelan GTWR
 - 1) Memilih *bandwidth* optimum menggunakan metode *grid search* yang secara sistematis menguji berbagai kombinasi parameter *bandwidth* spasial (h) dan rasio spasio-temporal (τ) dengan τ dibatasi hingga 20
 - 2) Menghitung nilai pembobot dengan fungsi kernel terbaik
 - 3) Melakukan estimasi parameter model GTWR
5. Melakukan pemodelan MSGTWR
 - 1) Memilih *bandwidth* optimum untuk setiap variabel
 - a. Menginisialisasi parameter dengan model GTWR
 - b. Menghitung nilai inisialisasi residual berdasarkan persamaan (8) dengan f_k sebagai nilai regresi variabel ke- k
 - c. Menjumlahkan residual ε dengan nilai regresi variabel ke- k , dimulai dari variabel ke-1 (f_1)
 - d. Menggunakan hasil $\varepsilon + f_1$ sebagai variabel dependen untuk perhitungan GTWR dengan variabel independen dan memperoleh *bandwidth* optimal dari variabel independen ke-1
 - e. Menghitung residual baru untuk variabel selanjutnya sampai memenuhi kondisi $k \leq p$ dengan p adalah banyaknya variabel
 - f. Mengulangi langkah c sampai e hingga *bandwidth* optimal dihitung untuk semua variabel
 - g. Menentukan konvergen tidaknya model dengan kriteria konvergensi skor perubahan (*SOC*) yang dihitung menggunakan persamaan (9)
 - h. Mengulangi langkah c hingga dicapai konvergensi atau nilai ambang batas 10^{-5} dan perhitungan selesai.
 - 2) Menghitung nilai pembobot dengan fungsi kernel terbaik
 - 3) Melakukan estimasi parameter model MSGTWR

- 4) Melakukan uji parsial parameter untuk mengevaluasi signifikan tidaknya setiap variabel independen terhadap variabel dependen berdasarkan model MSGTWR
6. Melakukan uji kebaikan model MSGTWR dengan nilai $AICc$ dan $RMSE$ menggunakan persamaan (17) dan (18)
7. Melakukan Interpretasi Model.