

## Determining Factors that Influence Unmet Need For Family Planning Using Geographically Weighted Logistic Regression With LASSO

### Penentuan Faktor yang Mempengaruhi *Unmet Need* Keluarga Berencana Menggunakan *Geographically Weighted Logistic Regression* Dengan LASSO

Dian Ayu Permata Sari Rusdy<sup>1</sup>, Sri Astuti Thamrin<sup>\*2</sup>, Anna Islamiyati<sup>3</sup>

<sup>1,2,3</sup>Departemen Statistika, Universitas Hasanuddin, Indonesia

Email: <sup>1</sup>dianrusdy77@gmail.com, <sup>2</sup>tuti@unhas.ac.id, <sup>3</sup>annaislamiyati701@gmail.com

<sup>\*</sup>Corresponding author

Received: 25 May 2024, revised: 14 December 2024, accepted: 25 December 2024

#### Abstract

Binary logistic regression is a regression used for categorical response variables with two possibilities: success or failure. This regression is a global model, making it inappropriate for spatial data. Binary logistic regression was then developed into geographically weighted logistic regression (GWLR). GWLR considers location factors into the model through a weight function. Nevertheless, GWLR is unable to overcome multicollinearity issue. Multicollinearity can cause the estimated parameters to be insignificant, thus it needs to be solved. A method to deal with multicollinearity is least absolute shrinkage and selection operator (LASSO). LASSO is applicable to various areas, including health, namely in the case of unmet need for family planning (FP). Unmet need for FP refers to productive-age women who do not wish to have more children or wish to postpone having children without using contraceptive methods. This study aims to obtain GWLR model with LASSO and influential factors, and acquire the performance of GWLR model with LASSO on unmet need for FP in South Sulawesi. The AIC value of the GWLR with LASSO model, which is 31,918, is less than the AIC value of the GWLR without LASSO, which is 38,879. This implies that GWLR with LASSO method is able to model unmet need for FP better than GWLR model. In addition, it was obtained that the status of unmet need for FP in 22 districts/cities was affected by the percentage of women with junior high school education or equivalent or lower, number of high-fertility women, percentage of husbands/families who refuse family planning, and number of KB staffs, while there were 2 districts/cities where the status of unmet need for KB was determined by the number of high-fertility women, percentage of husbands/families who refuse family planning, and number of FP staffs.



**Keywords:** GWLR, Family Planning, LASSO, Multicollinearity, Binary Logistic Regression, Unmet Need.

### **Abstrak**

Regresi logistik biner merupakan regresi yang digunakan apabila variabel responnya bertipe kategorik dengan dua kemungkinan yaitu sukses atau gagal. Regresi ini bersifat global, sehingga kurang tepat diterapkan pada data spasial. Regresi logistik biner kemudian dikembangkan menjadi *geographically weighted logistic regression* (GWLR). GWLR mempertimbangkan faktor lokasi yang dimasukkan ke dalam model melalui fungsi pembobot. Meski demikian, GWLR tidak dapat mengatasi masalah multikolinearitas. Multikolinearitas dapat menyebabkan hasil estimasi parameter model menjadi tidak signifikan, sehingga perlu diatasi. Salah satu metode untuk mengatasi multikolinearitas adalah *least absolute shrinkage and selection operator* (LASSO). LASSO dapat diterapkan di berbagai bidang, salah satunya dalam bidang kesehatan, yaitu pada kasus faktor yang memengaruhi *unmet need* Keluarga Berencana (KB). *Unmet need* KB merupakan wanita usia produktif dan tidak ingin memiliki anak lagi ataupun ingin menunda memiliki anak tetapi tidak menggunakan metode kontrasepsi. Penelitian ini bertujuan untuk mengidentifikasi faktor yang berpengaruh terhadap *unmet need* KB di Sulawesi Selatan melalui pemodelan GWLR dengan LASSO. Selain itu, dalam penelitian ini juga dianalisis performa pemodelan GWLR dengan LASSO dalam mengidentifikasi faktor yang memengaruhi *unmet need* KB di Sulawesi Selatan. Hasil yang diperoleh menunjukkan pemodelan GWLR dengan metode LASSO (31,918) dapat mengidentifikasi faktor yang memengaruhi *unmet need* KB di Sulawesi Selatan lebih baik dibandingkan model GWLR tanpa LASSO (38,879) berdasarkan nilai *Akaike Information Criterion* (AIC). Selain itu, diperoleh bahwa *unmet need* KB di Sulawesi Selatan, khususnya 22 Kabupaten/Kota dipengaruhi oleh persentase wanita dengan pendidikan SMP/setara atau lebih rendah, banyak wanita usia subur, persentase suami/keluarga yang menolak terhadap KB, dan banyak petugas lapangan KB. Sementara 2 Kabupaten/Kota lainnya, *unmet need* KB dipengaruhi oleh banyak wanita usia subur, persentase suami/keluarga yang menolak terhadap KB, dan banyak petugas lapangan KB.

**Kata kunci:** GWLR, Keluarga Berencana, LASSO, Multikolinearitas, Regresi Logistik Biner, *Unmet need*.

## **1. PENDAHULUAN**

Analisis regresi merupakan salah satu analisis dalam statistika yang digunakan untuk menaksir pola hubungan antara variabel respon dan variabel prediktor. Variabel respon merupakan variabel yang nilainya dipengaruhi oleh variabel prediktor. Sedangkan variabel prediktor merupakan variabel yang dianggap memiliki pengaruh terhadap variabel respon. Salah satu analisis regresi yang umum digunakan adalah regresi logistik biner [2].

Regresi logistik biner merupakan metode yang digunakan untuk menggambarkan hubungan antara variabel respon dengan satu atau lebih variabel prediktor. Variabel respon dari regresi logistik biner merupakan variabel bertipe kategorik dengan dua kemungkinan, yaitu sukses atau gagal [16]. Analisis regresi logistik biner merupakan model yang bersifat global atau tidak melibatkan unsur lokasi geografis, sehingga kurang tepat apabila diterapkan pada data yang dipengaruhi oleh lokasi secara geografis atau yang biasa disebut dengan data spasial [11]. Untuk mengatasi hal tersebut, metode regresi logistik biner kemudian dikembangkan menjadi *geographically weighted logistic regression* (GWLR).

GWLR merupakan metode statistik yang dapat digunakan untuk memodelkan hubungan antara variabel respon dengan variabel prediktor, serta mempertimbangkan faktor lokasi dengan asumsi bahwa data berdistribusi bernoulli [36]. Pada model GWLR, faktor lokasi dimasukkan ke dalam model melalui fungsi pembobot. Fungsi pembobot ini akan mewakili letak data observasi

satu terhadap data observasi lainnya. Model GWLR bersifat lokal sehingga memiliki koefisien regresi yang masing-masing bergantung pada lokasi dimana data tersebut diamati [18]. Namun, model GWLR memiliki kekurangan, yaitu tidak dapat mengatasi masalah multikolinearitas.

Multikolinearitas terjadi karena adanya korelasi yang cukup tinggi di antara variabel prediktor. Adanya multikolinearitas dapat diketahui dengan menggunakan nilai *variance inflation factor* (VIF). Nilai VIF akan menjadi semakin besar jika terdapat korelasi yang semakin besar di antara variabel prediktor [15]. Jika VIF memiliki nilai lebih dari 10, maka multikolinearitas akan memberikan pengaruh yang serius pada estimasi model [25]. Adanya multikolinearitas akan menyebabkan hasil taksiran parameter model menjadi tidak signifikan, sehingga multikolinearitas ini perlu diatasi [41]. Metode yang dapat digunakan untuk mengatasi multikolinearitas adalah *least absolute shrinkage and selection operator* (LASSO).

LASSO merupakan metode yang dapat digunakan untuk mengatasi multikolinearitas yaitu dengan melakukan penyeleksian variabel pada model. Penyeleksian ini dilakukan dengan cara menyusutkan nilai koefisien regresi. Penyusutan nilai koefisien regresi oleh LASSO dapat dilakukan hingga mendekati nol atau tepat nol [8]. Penyusutan nilai ini menjadi kelebihan dari metode LASSO karena dapat menyeleksi variabel prediktor pada model dan hanya variabel yang berpengaruh yang dimasukkan ke dalam model sehingga memudahkan dalam menginterpretasikan model [27]. LASSO dapat diterapkan di berbagai bidang, salah satunya dalam bidang kesehatan, yaitu pada kasus *unmet need* Keluarga Berencana (KB).

*Unmet need* KB didefinisikan sebagai wanita yang memiliki usia produktif dan tidak ingin memiliki anak lagi ataupun ingin menunda anak yang berikutnya tetapi tidak menggunakan metode kontrasepsi [49]. Tingginya angka *unmet need* KB di Indonesia menjadi salah satu permasalahan yang sampai saat ini belum terselesaikan dengan baik. Data statistik rutin Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) menunjukkan bahwa pada tahun 2020 terdapat 12,13 persen dari total Pasangan Usia Subur (PUS) merupakan *unmet need* KB. Kemudian angka ini naik menjadi 12,21 persen pada tahun 2021 [7]. Hal ini menunjukkan bahwa angka *unmet need* KB mengalami peningkatan secara nasional. Peningkatan ini disebabkan tingginya *unmet need* KB di beberapa Provinsi, salah satunya Sulawesi Selatan.

Berdasarkan data statistik rutin BKKBN, angka *unmet need* KB di Sulawesi Selatan mencapai 13,12 persen pada tahun 2020. Angka ini kemudian meningkat menjadi 13,70 persen pada tahun berikutnya [7]. Peningkatan angka ini menjadikan Sulawesi Selatan sebagai salah satu provinsi dengan angka *unmet need* KB yang cukup tinggi di Indonesia. Angka *unmet need* KB yang tinggi akan meningkatkan risiko kehamilan yang tidak diinginkan. Hal ini juga akan berpengaruh terhadap peningkatan risiko kematian ibu dan bayi [31].

Penelitian sebelumnya telah dilakukan oleh Mayfield dkk [23] dan diperoleh kesimpulan bahwa metode GWLR lebih baik dalam memodelkan data spasial dibandingkan regresi logistik. Penelitian juga dilakukan oleh Lestari dkk [19] dengan kesimpulan pemodelan dengan metode LASSO pada *geographically weighted regression* (GWR) memberikan hasil yang lebih baik dibandingkan metode GWR tanpa LASSO. Penelitian lain juga dilakukan oleh Amalah [3] yang melakukan estimasi parameter model *geographically weighted logistic regression* (GWLR) pada data yang mengandung multikolinearitas dengan metode *ridge*. Pada penelitian-penelitian sebelumnya belum terdapat penelitian lebih lanjut terkait penggunaan LASSO pada GWLR. Oleh karena itu menjadi penting untuk melakukan penelitian terkait penggunaan LASSO pada GWLR tersebut khususnya pada kasus identifikasi faktor yang memengaruhi *unmet need* KB di Provinsi Sulawesi Selatan tahun 2021. Adapun tujuan dari penelitian ini adalah untuk mengidentifikasi faktor yang memengaruhi *unmet need* KB di Sulawesi Selatan melalui pemodelan GWLR dengan LASSO. Selain itu juga dianalisis performa pemodelan GWLR dengan LASSO dalam mengidentifikasi faktor yang memengaruhi *unmet need* KB di Sulawesi Selatan tahun 2021.

## 2. METODE PENELITIAN

### 2.1 Data dan Variabel

Data yang digunakan dalam penelitian ini adalah data sekunder mengenai *unmet need* KB, tingkat pendidikan, pekerjaan, dan pengeluaran masyarakat di Provinsi Sulawesi Selatan tahun 2021. Data *unmet need* KB diperoleh dari BKKBN Sulawesi Selatan. Sedangkan data tingkat pendidikan, pekerjaan, dan pengeluaran masyarakat diperoleh dari publikasi Badan Pusat Statistik (BPS) Sulawesi Selatan pada situs <https://sulsel.bps.go.id>. Pada penelitian ini juga digunakan data jarak antar Kabupaten/Kota di Sulawesi Selatan sebagai faktor pembobot lokasi, dengan Kantor Bupati/Walikota sebagai titik acuan pengukuran jarak. Jarak antar Kantor Bupati/Walikota diperoleh melalui aplikasi *Google Maps* dengan mengambil jarak dengan rute terpendek.

Variabel yang digunakan dalam penelitian ini terdiri dari satu variabel respon berbentuk biner (Y) dan tujuh variabel prediktor (X) seperti berikut ini.

#### 1) Variabel Respon

Variabel respon yang digunakan bersifat kategorik, yaitu dengan mengelompokkan setiap kabupaten/kota menjadi *unmet need* KB dan *met need* KB. Pengelompokan ini berdasarkan nilai persentase *unmet need* KB Provinsi Sulawesi Selatan sebesar 13,70 persen dengan pengelompokan sebagai berikut.

0 = *met need* KB (persentase *unmet need* KB kabupaten/kota kurang dari 13,70 persen)

1 = *unmet need* KB (persentase *unmet need* KB kabupaten/kota lebih atau sama dengan dari 13,70 persen)

#### 2) Variabel Prediktor

Variabel prediktor yang digunakan yaitu banyak tempat pelayanan KB ( $X_1$ ), persentase wanita dengan pendidikan SMP/setara atau lebih rendah ( $X_2$ ), pengeluaran rata-rata perkapita masyarakat ( $X_3$ ), banyak wanita bekerja ( $X_4$ ), banyak wanita usia subur (WUS) ( $X_5$ ), persentase suami/keluarga yang menolak terhadap KB ( $X_6$ ), banyak petugas lapangan keluarga berencana (PLKB) ( $X_7$ ).

### 2.2 Regresi Logistik Biner

Regresi logistik biner merupakan metode analisis data yang digunakan untuk mendeskripsikan hubungan antara variabel respon yang bersifat biner atau dikotomus dengan satu atau lebih variabel prediktor. Variabel respon regresi logistik biner terdiri dari dua kategori, yaitu sukses yang dinotasikan dengan 1 atau gagal yang dinotasikan dengan 0 [17]. Variabel respon pada regresi logistik biner mengikuti distribusi Bernoulli dengan fungsi probabilitas seperti yang ditunjukkan pada Persamaan (2.1) [39].

$$f(y_i) = \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (2.1)$$

$\pi(x_i)$  adalah peluang kejadian ke- $i$ , dan  $y_i$  adalah peubah acak ke- $i$  yang terdiri dari 0 dan 1.

Dalam regresi logistik, variabel respon dituliskan sebagai  $y = \pi(x_i) + \varepsilon$ , dengan  $\varepsilon$  mempunyai salah satu dari kemungkinan dua nilai, yaitu apabila  $y = 1$  maka  $\varepsilon = 1 - \pi(x_i)$  dengan peluang  $\pi(x_i)$  dan jika  $y = 0$  maka  $\varepsilon = -\pi(x_i)$  dengan peluang  $1 - \pi(x_i)$  [4]. Adapun model probabilitas regresi logistik biner dapat dituliskan sebagaimana pada Persamaan (2.2).

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} = \frac{\exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)}{1 + \exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)} \quad (2.2)$$

Apabila Persamaan (2.2) dituliskan menjadi model logit, maka diperoleh Persamaan (2.3).

$$\text{logit}[\pi(x)] = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \sum_{k=1}^p \beta_k x_k \quad (2.3)$$

Nilai  $\beta$  adalah parameter model dan  $x_k, k = 1, 2, \dots, p$  merupakan variabel prediktor ke- $k$  [2].

Pendugaan parameter pada regresi logistik biner dilakukan dengan menggunakan metode *maximum likelihood estimation* (MLE). Metode MLE menghasilkan nilai parameter dengan memaksimalkan fungsi peluang dari data yang diamati dengan menggunakan fungsi *log-likelihood* yang ditampilkan pada Persamaan (2.4) [46].

$$L(\boldsymbol{\beta}) = \sum_{i=1}^n [y_i(\boldsymbol{\beta}^T \mathbf{x}_i) - \ln(1 + \exp(\boldsymbol{\beta}^T \mathbf{x}_i))] \quad (2.4)$$

Untuk memperoleh nilai estimasi  $\boldsymbol{\beta}$ , maka Persamaan (2.4) diturunkan terhadap  $\boldsymbol{\beta}$  dan hasil turunan tersebut disamakan dengan nol, sehingga diperoleh Persamaan (2.5) [20].

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = X^T(\mathbf{y} - \boldsymbol{\pi}(\mathbf{x})) = \mathbf{0} \quad (2.5)$$

Berdasarkan Persamaan (2.5), hasil turunan pertama fungsi  $L(\boldsymbol{\beta})$  merupakan Persamaan non linier, sehingga solusi eksak untuk mendapatkan penaksir  $\boldsymbol{\beta}$  tidak dapat diperoleh secara analitik. Metode alternatif untuk memperoleh penaksir  $\boldsymbol{\beta}$  adalah metode iteratif *Newton-Raphson* [20].

$$\hat{\boldsymbol{\beta}}^{(t+1)} = \hat{\boldsymbol{\beta}}^{(t)} - \left(H(\hat{\boldsymbol{\beta}}^{(t)})\right)^{-1} \mathbf{g}(\hat{\boldsymbol{\beta}}^{(t)}) \quad (2.6)$$

$t = 1, 2, \dots$  sampai konvergen.

$\mathbf{g}(\hat{\boldsymbol{\beta}}^{(t)}) = \frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$  merupakan vektor hasil turunan pertama fungsi *log-likelihood* sesuai Persamaan (2.5).  $H(\hat{\boldsymbol{\beta}}^{(t)})$  adalah matriks Hessian yang merupakan matriks hasil turunan kedua dari fungsi  $L(\boldsymbol{\beta})$  dengan  $H(\hat{\boldsymbol{\beta}}^{(t)}) = -X^T V X$ .  $V$  merupakan matriks diagonal berukuran  $n \times n$  dengan elemen diagonal:

$$V = \begin{bmatrix} \pi(x_1)(1 - \pi(x_1)) & 0 & \dots & 0 \\ 0 & \pi(x_2)(1 - \pi(x_2)) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi(x_n)(1 - \pi(x_n)) \end{bmatrix}$$

### 2.3 Multikolinearitas

Multikolinearitas merupakan suatu kondisi adanya hubungan linier atau korelasi yang tinggi antara variabel prediktor pada model regresi. Adanya multikolinearitas dapat mempengaruhi interpretasi koefisien regresi sehingga memungkinkan terjadinya kesalahan dalam pengambilan keputusan serta pengujian signifikansi parsial menjadi tidak signifikan [14]. Pendeteksian multikolinearitas dapat dilakukan dengan menggunakan nilai *variance inflation factor* (VIF). Nilai VIF yang lebih besar dari 10 mengindikasikan bahwa terdapat multikolinearitas pada model. Nilai VIF dapat diidentifikasi menggunakan Persamaan (2.7).

$$VIF_k = \frac{1}{1 - R_k^2}, k = 1, 2, \dots, p \quad (2.7)$$

Nilai  $R_k^2$  adalah koefisien determinasi antara  $x_k$  dengan variabel prediktor lainnya pada model [41].

### 2.4 Least Absolute Shrinkage and Selection Operator

*Least absolute shrinkage and selection operator* (LASSO) merupakan metode untuk mengatasi multikolinearitas dengan cara mereduksi koefisien parameter  $\boldsymbol{\beta}$  mendekati nol atau

menjadi nol sehingga menghasilkan parameter  $\beta$  yang lebih sedikit dan model yang dihasilkan mudah diinterpretasikan dibandingkan dengan model awal [13]. Penduga koefisien LASSO dituliskan pada Persamaan (2.8).

$$\hat{\beta}^{LASSO} = \arg \min_{\beta} \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{k=1}^p x_{ki} \beta_k \right)^2, \sum_{k=1}^p |\beta_k| \leq t \quad (2.8)$$

Nilai  $t$  merupakan parameter yang mengontrol penyusutan pada estimasi parameter model, dengan nilai  $t \geq 0$  [12]. Misalkan  $\hat{\beta}_k$  merupakan nilai estimasi parameter dengan metode *Ordinary Least Squares* (OLS), jika  $t < \sum_{k=1}^p |\hat{\beta}_k|$ , maka  $\hat{\beta}^{LASSO}$  akan mengecil mendekati nol atau tepat nol. Jika nilai  $t \geq \sum_{k=1}^p |\hat{\beta}_k|$  maka hasil estimasi parameter dengan LASSO akan sama dengan nilai estimasi parameter dengan metode *Ordinary Least Squares* (OLS) [47]. Persamaan (2.8) juga dapat dituliskan dalam Persamaan lagrange sebagaimana dituliskan pada Persamaan (2.9).

$$\hat{\beta}^{LASSO} = \arg \min \left\{ \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{k=1}^p x_{ki} \beta_k \right)^2 + \lambda \sum_{k=1}^p |\beta_k| \right\} \quad (2.9)$$

Nilai  $\lambda$  pada Persamaan (2.9) didapatkan dengan menggunakan metode *k-fold Cross-Validation* (CV), yaitu dengan membagi data secara acak menjadi  $k$ -kelipatan (*fold*) yang berukuran sama. Diantara  $k$ -*fold*, satu *fold* disimpan sebagai data *test* untuk menguji model. Kemudian sisa  $k - 1$  *fold* akan digunakan sebagai data *train* untuk menyesuaikan model dengan nilai dari parameter penyusutan. Nilai  $k$ -*fold* CV untuk masing-masing  $\lambda$  didefinisikan pada Persamaan (2.10) [13].

$$k - CV_{(\lambda)} = \frac{1}{k} \sum_{i=1}^k \left( y_i - \hat{y}_{i(\lambda)}^{-k(i)} \right)^2 \quad (2.10)$$

$\hat{y}_{i(\lambda)}^{-k(i)}$  adalah nilai dugaan  $y$  pada saat *fold* ke- $k$  tidak digunakan dalam pendugaan parameter [6]. Nilai  $\lambda$  optimal merupakan  $\lambda$  yang memiliki nilai  $k$ -*fold* CV minimum [13].

Metode LASSO dapat diterapkan pada regresi logistik. Penerapan LASSO pada regresi logistik dilakukan dengan memasukkan pinalti LASSO pada fungsi *log-likelihood* dalam pendugaan parameter model, sehingga estimasi parameter regresi logistik dengan LASSO diperoleh Persamaan (2.11) [45].

$$\hat{\beta}^{LASSO} = \arg \max \left\{ L(\beta) - \lambda \sum_{k=1}^p |\beta_k| \right\} \quad (2.11)$$

$L(\beta)$  merupakan fungsi *log-likelihood* untuk regresi logistik biner yang ditampilkan pada Persamaan (2.4).

## 2.5 Algoritma Cyclic Coordinate Descent

Untuk mendapatkan hasil dari estimasi parameter regresi logistik dengan LASSO maka diperlukan suatu algoritma, yaitu algoritma *Cyclic Coordinate Descent* (CCD). Tujuan dari algoritma CCD adalah membentuk nilai parameter baru ( $\beta_k^{baru}$ ) yang dapat memaksimalkan nilai *penalized log-likelihood* seperti pada Persamaan (2.11). Nilai dari parameter baru yang terbentuk ditampilkan pada Persamaan (2.12).

$$\beta_k^{baru} = \begin{cases} \beta_k - \Delta_k, & \text{jika } \Delta v_k < -\Delta_k \\ \beta_k + \Delta v_k, & \text{jika } -\Delta_k \leq \Delta v_k \leq \Delta_k \\ \beta_k + \Delta_k, & \text{jika } \Delta v_k > \Delta_k \end{cases} \quad (2.12)$$

Nilai  $\beta_k$  adalah penduga parameter awal,  $\Delta_k$  adalah hampiran penduga  $\beta_k$ , dan  $\Delta v_k$  adalah nilai penduga terbaru [34]. Nilai  $\Delta_k$  dimisalkan dalam  $[1, \infty)$  atau dapat pula menggunakan Persamaan (2.13).

$$\Delta_k = \max\left(2|\beta_k|, \frac{\Delta_k}{2}\right) \quad (2.13)$$

Nilai interval  $(\beta_k - \Delta_k, \beta_k + \Delta_k)$  berubah-ubah tergantung nilai penduga awal dan nilai  $\Delta v_k$  sebelumnya. Adapun nilai  $\Delta v_k$  dapat diperoleh dengan Persamaan (2.14).

$$\Delta v_k = -\frac{S_k(\beta) - \lambda s}{Q(\beta_k, \Delta_k)} \quad (2.14)$$

$S_k(\beta)$  adalah turunan pertama fungsi *log-likelihood* untuk regresi logistik biner,  $Q(\beta_k, \Delta_k)$  adalah turunan kedua fungsi *log-likelihood* untuk regresi logistik biner, dan nilai  $s = \beta_k/|\beta_k|$ .

Adapun prosedur dari algoritma CCD adalah sebagai berikut [45].

1. Menentukan nilai awal  $\beta_k = 0$  dan  $\Delta_k = 1$  untuk  $k = 1, 2, \dots, p$
2. Menghitung nilai  $\Delta v_k$  dengan ketentuan:
  - a. Apabila  $\beta_k = 0$ , maka  $s = 1$ , kemudian hitung  $\Delta v_k$  menggunakan Persamaan (2.14). apabila  $\Delta v_k < 0$  maka  $s = -1$  kemudian hitung  $\Delta v_k$ . Namun apabila  $\Delta v_k \geq 0$  maka  $\Delta v_k = 0$
  - b. Apabila  $\beta_k \neq 0$ , maka  $s = \frac{\beta_k}{|\beta_k|}$ , kemudian hitung  $\Delta v_k$  menggunakan Persamaan (2.14). apabila  $s(\beta_k + \Delta v_k) \leq 0$  maka  $\Delta v_k = -\beta_k$ .
3. Menghitung nilai  $\Delta_k$  menggunakan Persamaan (2.13)
4. Menghitung nilai  $\beta_k^{baru}$  menggunakan Persamaan (2.12)
5. Mengulangi Langkah 2 hingga Langkah 4 hingga mencapai kondisi konvergen.

## 2.6 Uji Heterogenitas Spasial

Uji heterogenitas spasial diperlukan untuk mengetahui adanya keragaman spasial pada data. Adanya heterogenitas spasial menunjukkan bahwa data dapat dimodelkan dengan metode *geographically weighted logistic regression* [40]. Uji heterogenitas spasial dapat dilakukan dengan menggunakan uji *Breusch-Pagan* dengan hipotesis sebagai berikut.

$H_0$  : tidak terdapat heterogenitas spasial

$H_1$  : terdapat heterogenitas spasial

Adapun statistik uji yang digunakan adalah:

$$BP = \frac{1}{2} \mathbf{f}^T Z(Z^T Z)^{-1} Z^T \mathbf{f} \sim \chi_{p,\alpha}^2 \quad (2.15)$$

$i = 1, 2, \dots, n$  adalah banyaknya observasi,  $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$ ,  $f_i = \left(\frac{e_i^2}{\sigma_e^2} - 1\right)$ , dengan  $e_i = y_i - \hat{y}_i$ ,  $Z$  adalah matriks berukuran  $n \times (p + 1)$  dari variabel prediktor yang sudah distandardkan, dengan  $p$  adalah banyaknya variabel prediktor. Hipotesis  $H_0$  ditolak apabila nilai  $BP > \chi_{p,\alpha}^2$  atau  $p$ -value  $< \alpha$  [35].

## 2.7 Matriks Pembobot Spasial

Pembobotan dalam analisis spasial merupakan hal yang penting karena nilai pembobotan mewakili letak data observasi satu terhadap data observasi yang lain [48]. Pembobot berperan untuk memberikan pendugaan parameter yang berbeda di setiap lokasi pengamatan [9]. Matriks pembobot

merupakan matriks diagonal yang berukuran  $(n \times n)$  yang elemen diagonalnya merupakan nilai pembobot untuk lokasi pengamatan ke- $i$ . Matriks pembobot dituliskan sebagai berikut [30].

$$W(u_i, v_i) = \begin{bmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{bmatrix} \quad (2.16)$$

Matriks pembobot tersebut perlu dihitung untuk setiap lokasi pengamatan. Besarnya nilai pembobot ditentukan oleh fungsi pembobot. Fungsi pembobot yang dapat digunakan adalah fungsi pembobot *fixed gaussian* yang diperoleh seperti pada Persamaan (2.17) [28].

$$w_{ij} = \exp \left\{ -\frac{1}{2} \left( \frac{d_{ij}}{b} \right)^2 \right\} \quad (2.17)$$

$w_{ij}$  adalah nilai pembobot pada lokasi ke- $i$  terhadap lokasi ke- $j$ ,  $d_{ij}$  merupakan jarak antara lokasi ke- $i$  dan lokasi ke- $j$ , dan  $b$  merupakan nilai *bandwidth*.

Nilai *bandwidth* yang digunakan pada Persamaan (2.17) merupakan nilai *bandwidth* optimum. Nilai *bandwidth* optimum dapat dilakukan dengan menggunakan nilai *Cross Validation* (CV) seperti pada persamaan (2.18) [24].

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)] \quad (2.18)$$

Nilai  $\hat{y}_{\neq i}(b)$  merupakan nilai taksiran  $y_i$  dengan menggunakan *bandwidth*  $b$  dengan pengamatan di lokasi ke- $i$  tidak dimasukkan dalam perhitungan, yaitu dengan membuat bobot lokasi ke- $i$  bernilai nol. *bandwidth* optimum akan diperoleh apabila nilai CV yang dihasilkan minimum [48].

## 2.8 Geographically Weighted Logistic Regression

*Geographically weighted logistic regression* (GWLRL) merupakan pengembangan dari regresi logistik biner yang diterapkan pada data spasial, yaitu data yang mempertimbangkan unsur lokasi. Pemodelan data spasial dengan regresi logistik biasa akan menyebabkan model menjadi tidak valid akibat adanya hubungan antara data variabel respon dan lokasi pengamatan yang menyebabkan nilai estimasi parameter tiap lokasi pengamatan berbeda [29]. Adapun model GWLRL untuk setiap lokasi ke- $i$  dituliskan pada Persamaan (2.19).

$$\begin{aligned} \pi(u_i, v_i) &= \frac{\exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \cdots + \beta_p(u_i, v_i)x_{ip})}{1 + \exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \cdots + \beta_p(u_i, v_i)x_{ip})} \\ &= \frac{\exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i)}{1 + \exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_i)} \end{aligned} \quad (2.19)$$

$\mathbf{x}_i^T$  merupakan matriks dengan elemen  $(1 \ x_{i1} \ x_{i2} \ \cdots \ x_{ip})$  dan  $\boldsymbol{\beta}^T(u_i, v_i) = (\beta_0(u_i, v_i) \ \beta_1(u_i, v_i) \ \beta_2(u_i, v_i) \ \cdots \ \beta_p(u_i, v_i))$ . Kemudian, dengan melakukan transformasi logit pada Persamaan (2.19), maka diperoleh Persamaan (2.20) [40].

$$\text{logit}[\pi(u_i, v_i)] = \ln \left( \frac{\pi(u_i, v_i)}{1 - \pi(u_i, v_i)} \right) = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} \quad (2.20)$$

Pendugaan parameter pada model GWLRL dilakukan dengan menggunakan metode *maximum likelihood estimation* (MLE). Metode MLE menghasilkan nilai parameter dengan memaksimalkan fungsi peluang dari data yang diamati dengan menggunakan fungsi *log-likelihood*. Fungsi *log-likelihood* dengan pembobot spasial untuk pengestimasi parameter GWLRL pada lokasi ke- $i$  ditampilkan pada Persamaan (2.21) [46].



$$L(\boldsymbol{\beta}(u_i, v_i)) = \sum_{j=1}^n w_{ij} (y_j \boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_j - \ln(1 + \exp(\boldsymbol{\beta}^T(u_i, v_i) \mathbf{x}_j))) \quad (2.21)$$

Selanjutnya untuk memperoleh fungsi  $\ln(L(\boldsymbol{\beta}(u_i, v_i)))$  yang dimaksimumkan, maka  $L(\boldsymbol{\beta}(u_i, v_i))$  diturunkan terhadap parameter  $\beta_k(u_i, v_i)$  dan hasil turunan tersebut disamakan dengan nol, sehingga diperoleh Persamaan (2.22) [21].

$$\frac{\partial L(\boldsymbol{\beta}(u_i, v_i))}{\partial \boldsymbol{\beta}(u_i, v_i)} = X^T W(u_i, v_i) (\mathbf{y} - \boldsymbol{\pi}(u_i, v_i)) = \mathbf{0} \quad (2.22)$$

Untuk mendapatkan nilai taksiran  $\beta_k$  dari fungsi  $\ln(L(\boldsymbol{\beta}(u_i, v_i)))$ , maka perhitungan MLE dapat dilakukan dengan menggunakan pendekatan numerik yaitu dengan iterasi Newton-Raphson [21].

$$\widehat{\boldsymbol{\beta}}^{(t+1)}(u_i, v_i) = \widehat{\boldsymbol{\beta}}^{(t)}(u_i, v_i) - \left( H(\widehat{\boldsymbol{\beta}}^{(t)}(u_i, v_i)) \right)^{-1} \mathbf{g}(\widehat{\boldsymbol{\beta}}^{(t)}(u_i, v_i)) \quad (2.23)$$

$t = 1, 2, \dots$  sampai konvergen,  $\mathbf{g}(\boldsymbol{\beta}^{(t)}(u_i, v_i)) = \frac{\partial L(\boldsymbol{\beta}(u_i, v_i))}{\partial \boldsymbol{\beta}(u_i, v_i)}$  merupakan vektor hasil turunan pertama fungsi log-likelihood sesuai Persamaan (2.22),  $H(\widehat{\boldsymbol{\beta}}^{(t)}(u_i, v_i))$  adalah matriks Hessian yang merupakan matriks hasil turunan kedua dari fungsi  $\partial L(\boldsymbol{\beta}(u_i, v_i))$ . dengan  $H(\boldsymbol{\beta}^{(t)}(u_i, v_i)) = -X^T W(u_i, v_i) V(u_i, v_i) X$ .  $V(u_i, v_i)$  merupakan matriks diagonal berukuran  $n \times n$  dengan elemen diagonal:

$$V(u_i, v_i) = \begin{bmatrix} \pi_{i1}(u_i, v_i)(1 - \pi_{i1}(u_i, v_i)) & 0 & \dots & 0 \\ 0 & \pi_{i2}(u_i, v_i)(1 - \pi_{i2}(u_i, v_i)) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi_{in}(u_i, v_i)(1 - \pi_{in}(u_i, v_i)) \end{bmatrix}$$

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 telah konvergen, yaitu jika  $\|\boldsymbol{\beta}^{(t+1)}(u_i, v_i) - \boldsymbol{\beta}^{(t)}(u_i, v_i)\| \leq \varepsilon$  dengan  $\varepsilon$  merupakan bilangan yang sangat kecil. Penaksir parameter model GWLR merupakan nilai  $\widehat{\boldsymbol{\beta}}^{(t+1)}(u_i, v_i)$  pada iterasi terakhir. Proses iterasi ini dilakukan pada setiap lokasi ke- $i$ , sehingga dapat diperoleh penaksir parameter model GWLR [21].

## 2.9 Uji Signifikansi Parameter

Uji signifikansi parameter dilakukan untuk mengetahui parameter-parameter apa saja yang memiliki pengaruh yang signifikan terhadap variabel respon [5]. Uji signifikansi parameter dilakukan dengan *wald test* dengan hipotesis:

$H_0 : \beta_k(u_i, v_i) = 0; k = 1, 2, \dots, p$  (variabel prediktor  $\beta_k(u_i, v_i)$  tidak berpengaruh signifikan terhadap variabel respon)

$H_1 : \beta_k(u_i, v_i) \neq 0$  (variabel prediktor  $\beta_k(u_i, v_i)$  berpengaruh signifikan terhadap variabel respon)  
Statistik uji yang digunakan adalah:

$$W = \frac{\widehat{\beta}_k(u_i, v_i)}{SE(\widehat{\beta}_k(u_i, v_i))} \quad (2.24)$$

Nilai  $\hat{\beta}_k(u_i, v_i)$  adalah nilai estimasi parameter  $\beta_k(u_i, v_i)$  pada lokasi ke- $i$  dan  $SE(\hat{\beta}_k(u_i, v_i)) = \sqrt{\text{Var}(\hat{\beta}_k(u_i, v_i))}$ . Adapun kriteria uji yang digunakan adalah  $H_0$  ditolak apabila nilai  $p$ -value lebih kecil dari  $\alpha$  atau  $|W|$  lebih besar dari  $Z_{\alpha/2}$  [1].

### 2.10 Uji Kesesuaian Model

Uji kesesuaian model GWLR dilakukan untuk melihat apakah faktor geografis memiliki pengaruh yang signifikan atau tidak terhadap model [22]. Pengujian ini dilakukan dengan membandingkan nilai *deviance* model regresi logistik dan model GWLR dengan hipotesis:

$H_0$  :  $\beta_k(u_i, v_i) = \beta_k; i = 1, 2, \dots, n; k = 1, 2, \dots, p$  (tidak ada perbedaan yang signifikan antara model regresi logistik dan model GWLR)

$H_1$  : paling sedikit ada satu  $\beta_k(u_i, v_i) \neq \beta_k$  (Ada perbedaan yang signifikan antara model regresi logistik dan model GWLR)

Nilai *deviance* model regresi logistik adalah  $D(\hat{\beta})$  dengan derajat bebas  $df_1 = n$  dan nilai *deviance* model GWLR adalah  $D(\hat{\beta}(u_i, v_i))$  dengan derajat bebas  $df_2 = np$ . Uji statistik yang digunakan seperti pada Persamaan (2.25) [32]:

$$F_{hitung} = \frac{D(\hat{\beta})/df_1}{D(\hat{\beta}(u_i, v_i))/df_2} \quad (2.25)$$

$F_{hitung}$  akan mengikuti distribusi  $F$  dengan derajat bebas  $df_1$  dan  $df_2$ . Kriteria uji yang digunakan adalah  $H_0$  ditolak apabila nilai  $F_{hitung}$  lebih besar dari  $F_{\alpha; df_1; df_2}$  atau  $p$ -value lebih kecil dari  $\alpha$  [32]. Penolakan  $H_0$  menandakan bahwa faktor geografis memiliki pengaruh yang signifikan pada model GWLR [22].

### 2.11 Interpretasi Model

Interpretasi model GWLR dilakukan dengan menggunakan nilai *odds ratio* (OR). Nilai OR digunakan untuk melihat kecenderungan variabel prediktor yang signifikan memengaruhi variabel respon. Nilai OR menggambarkan perubahan kecenderungan setiap penambahan satu unit variabel prediktor [33]. Nilai OR pada model GWLR dapat dihitung menggunakan Persamaan (2.26) [1].

$$OR(u_i, v_i) = \exp(\hat{\beta}(u_i, v_i)) \quad (2.26)$$

Apabila nilai  $OR = 1$ , berarti variabel prediktor dan variabel respon tidak memiliki hubungan. Kemudian apabila nilai  $OR < 1$ , berarti variabel prediktor memiliki hubungan negatif terhadap variabel respon. Nilai  $OR > 1$  menyatakan variabel prediktor memiliki hubungan positif terhadap variabel respon [10].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Uji Mutikolinearitas

Perhitungan nilai VIF dilakukan menggunakan Persamaan (2.7). Adapun nilai VIF yang diperoleh dapat dilihat pada Tabel 3.1.

Tabel 3.1. Nilai VIF

Variabel	VIF	Keterangan
$x_1$	2,259	Tidak Terdapat Multikolinearitas
$x_2$	1,858	Tidak Terdapat Multikolinearitas

$x_3$	2,113	Tidak Terdapat Multikolinearitas
$x_4$	23,756	Terdapat Multikolinearitas
$x_5$	42,838	Terdapat Multikolinearitas
$x_6$	2,338	Tidak Terdapat Multikolinearitas
$x_7$	6,272	Tidak Terdapat Multikolinearitas

Dari Tabel 3.1 dapat dilihat bahwa variabel  $x_4$  yaitu banyak wanita bekerja dan variabel  $x_5$  yaitu banyak wanita usia subur memiliki nilai VIF yang lebih besar dari 10, yaitu masing-masing sebesar 23,756 dan 42,838. dengan demikian dapat disimpulkan bahwa terdapat multikolinearitas pada data *unmet need* KB. Untuk mengatasi hal tersebut, maka digunakan metode *least absolute shrinkage and selection operator* (LASSO).

### 3.2 Uji Heterogenitas Spasial

Pengujian heterogenitas spasial dilakukan dengan menggunakan uji *Breusch-Pagan* dengan hipotesis sebagai berikut.

$H_0$  : tidak terdapat heterogenitas spasial

$H_1$  : terdapat heterogenitas spasial

Pengujian heterogenitas spasial dilakukan dengan menggunakan Persamaan (2.15) sehingga diperoleh nilai statistik uji BP, yaitu 16,314. Nilai tersebut lebih besar dari nilai  $\chi^2_{7;0,05}$  yaitu 14,067. Sehingga kriteria keputusan adalah tolak  $H_0$ , yang artinya terdapat heterogenitas spasial pada data *unmet need* KB. Oleh karena itu, model yang digunakan adalah model yang memperhatikan faktor lokasi, yaitu model GWLR.

### 3.3 Pemodelan *Geographically Weighted Logistic Regression* dengan LASSO

Pendugaan parameter model GWLR dengan LASSO diawali dengan menentukan matriks pembobot. Adapun matriks pembobot yang digunakan adalah pembobot *fixed gaussian*. Penentuan matriks pembobot yang digunakan melibatkan jarak ( $d_{ij}$ ) pada setiap lokasi pengamatan dan nilai *bandwith* ( $b$ ) optimum. Jarak yang digunakan merupakan jarak dengan rute terpendek antar Kantor Bupati/Walikota di setiap Kabupaten/Kota di Provinsi Sulawesi Selatan yang diperoleh dari aplikasi *Google Maps*. Adapun nilai *bandwith* dapat diperoleh berdasarkan nilai CV. *Bandwith* optimum akan diperoleh apabila nilai CV yang dihasilkan minimum. Dengan menggunakan Persamaan (2.18) dan bantuan *software* R-4.2.3 diperoleh nilai *bandwith* optimum adalah 592,132. Nilai jarak dan *bandwith* yang telah diperoleh selanjutnya digunakan pada Perhitungan matriks pembobot dengan menggunakan Persamaan (2.17). Matriks pembobot yang telah diperoleh kemudian digunakan untuk menduga parameter model GWLR dengan LASSO di setiap Kabupaten/Kota di Sulawesi Selatan.

Penerapan LASSO pada GWLR dilakukan dengan memasukkan pinalti LASSO pada fungsi *log-likelihood* dalam proses pendugaan parameter GWLR sehingga diperoleh Persamaan (3.1).

$$\ln \left( L(\boldsymbol{\beta}^{LASSO}(u_i, v_i)) \right) = L(\boldsymbol{\beta}(u_i, v_i)) - \lambda \sum_{k=1}^p |\beta_k(u_i, v_i)| \quad (3.1)$$

Fungsi  $L(\boldsymbol{\beta}(u_i, v_i))$  pada Persamaan (3.1) merupakan fungsi *log-likelihood* sesuai Persamaan (2.21). Penduga parameter GWLR dengan metode LASSO dapat diperoleh dengan memaksimalkan fungsi *log-likelihood* pada Persamaan (3.1) sehingga penduga parameter GWLR dengan metode LASSO dapat diperoleh pada Persamaan (3.2).

$$\hat{\boldsymbol{\beta}}^{LASSO}(u_i, v_i) = \arg \max \left\{ \ln \left( L(\boldsymbol{\beta}^{LASSO}(u_i, v_i)) \right) \right\}$$

$$= \arg \max \left\{ L(\boldsymbol{\beta}(u_i, v_i)) - \lambda \sum_{k=1}^p |\beta_k(u_i, v_i)| \right\} \quad (3.2)$$

Nilai  $\lambda$  pada Persamaan (3.2) diperoleh dengan menggunakan metode *k-fold Cross-Validation* (*k-fold CV*). Kemudian untuk mendapatkan hasil dari Persamaan (3.2) maka digunakan algoritma *Cyclic Coordinate Descent* (CCD). Penerapan algoritma CCD diawali dengan mengalikan matriks pembobot  $W(u_i, v_i)$  dengan matriks variabel prediktor yang telah distandarisasi, sehingga variabel respon dan variabel prediktor standar yang digunakan berturut-turut adalah  $\mathbf{y}^T = (y_1 \ y_2 \ \dots \ y_{24})$  dan  $Z_{W(u_i, v_i)} = W(u_i, v_i)Z$ .  $\mathbf{y}$  dan  $Z_{W(u_i, v_i)}$  kemudian diterapkan algoritma CCD untuk memperoleh nilai  $\hat{\boldsymbol{\beta}}^{LASSO}(u_i, v_i)$ .

Model GWLR dengan LASSO yang telah diperoleh selanjutnya dilakukan uji signifikansi parameter dan uji kesesuaian model. Uji signifikansi parameter dilakukan dengan *wald test* dengan hipotesis yang digunakan yaitu:

$H_0 : \beta_k(u_i, v_i) = 0, k = 1, 2, \dots, 7$ ; (variabel prediktor  $\beta_k(u_i, v_i)$  tidak berpengaruh signifikan terhadap variabel respon)

$H_1 : \beta_k(u_i, v_i) \neq 0$  (variabel prediktor  $\beta_k(u_i, v_i)$  berpengaruh signifikan terhadap variabel respon)

Statistik uji yang digunakan yaitu statistik uji  $W$  pada Persamaan (2.24). Pengujian ini dilakukan untuk setiap lokasi dengan hasil uji signifikansi parameter untuk setiap Kabupaten/Kota dapat dilihat pada Tabel 3.2.

**Tabel 3.2.** Hasil Uji Signifikansi Parameter

Kabupaten/Kota	Variabel yang Signifikan
Bantaeng, Barru, Bone, Bulukumba, Enrekang, Gowa, Jeneponto, Kepulauan Selayar, Kota Makassar, Kota Palopo, Luwu, Luwu Timur, Maros, Pangkajene dan Kepulauan, Pinrang, Sidenreng Rappang, Sinjai, Soppeng, Takalar, Tana Toraja, Toraja Utara, Wajo.	$x_2, x_5, x_6, x_7$
Kota Parepare, Luwu Utara	$x_5, x_6, x_7$

Berdasarkan Tabel 3.2 didapatkan bahwa terdapat 22 Kabupaten/Kota yang status *unmet need* KB di wilayah tersebut dipengaruhi oleh persentase wanita dengan pendidikan SMP/setara atau lebih rendah ( $x_2$ ), banyak wanita usia subur ( $x_5$ ), persentase suami/keluarga yang menolak terhadap KB ( $x_6$ ), dan banyak petugas lapangan KB ( $x_7$ ). Hal ini sesuai dengan penelitian yang dilakukan oleh [43], [37], [44], dan [31]. Adapun 2 Kabupaten/Kota lainnya, status *unmet need* KB di wilayahnya dipengaruhi oleh banyak wanita usia subur ( $x_5$ ), persentase suami/keluarga yang menolak terhadap KB ( $x_6$ ), dan banyak petugas lapangan KB ( $x_7$ ). Hal ini sesuai dengan penelitian yang dilakukan oleh [43], [37], dan [44].

Setelah dilakukan uji signifikansi parameter, selanjutnya dilakukan uji kesesuaian model. Adapun hipotesis yang digunakan yaitu:

$H_0 : \beta_k(u_i, v_i) = \beta_k; i = 1, 2, \dots, 24; k = 1, 2, \dots, 7$  (tidak ada perbedaan yang signifikan antara model regresi logistik dengan LASSO dan model GWLR dengan LASSO)

$H_1 : \text{paling sedikit ada satu } \beta_k(u_i, v_i) \neq \beta_k$  (Ada perbedaan yang signifikan antara model regresi logistik dengan LASSO dan model GWLR dengan LASSO)

Statistik uji yang digunakan yaitu statistik uji  $F$  pada Persamaan (2.25) dengan  $df_1 = 24$  dan  $df_2 = 24 \times 7 = 168$ . Adapun Nilai statistik uji  $F$  untuk model GWLR dengan LASSO adalah 22,769. nilai tersebut lebih besar dari nilai  $F_{0,05;24;168}$  yaitu 1,582. Sehingga kriteria keputusannya adalah tolak  $H_0$ , yang artinya terdapat perbedaan yang signifikan antara model regresi logistik

dengan LASSO dan model GWLR dengan LASSO. Ini juga berarti bahwa faktor lokasi memiliki pengaruh yang signifikan terhadap model.

Berdasarkan hasil pengujian, diperoleh model GWLR dengan LASSO untuk pemodelan *unmet need* KB di Sulawesi Selatan. Sebagai contoh, diberikan model untuk Kabupaten Bantaeng dan Kota Parepare. Model untuk Kabupaten Bantaeng dinyatakan oleh  $\pi(u_1, v_1)$  dan model untuk Kota Parepare dinyatakan oleh  $\pi(u_{11}, v_{11})$ .

$$\pi(u_1, v_1) = \frac{\exp(-0,215 + 0,253x_2 - 0,304x_5 - 0,651x_6 + 0,779x_7)}{1 + \exp(-0,215 + 0,253x_2 - 0,304x_5 - 0,651x_6 + 0,779x_7)}$$

$$\pi(u_{11}, v_{11}) = \frac{\exp(-0,268 - 0,860x_5 - 0,872x_6 + 1,273x_7)}{1 + \exp(-0,268 - 0,860x_5 - 0,872x_6 + 1,273x_7)}$$

Kedua model di atas dapat dituliskan menjadi model logit sebagai berikut.

$$\text{logit}[\pi(u_1, v_1)] = -0,215 + 0,253x_2 - 0,304x_5 - 0,651x_6 + 0,779x_7$$

$$\text{logit}[\pi(u_{11}, v_{11})] = -0,268 - 0,860x_5 - 0,872x_6 + 1,273x_7$$

### 3.4 Interpretasi Model

Model GWLR dengan LASSO yang telah diperoleh selanjutnya diinterpretasi. Interpretasi model dilakukan dengan menggunakan nilai *odds ratio* (OR) dari setiap variabel prediktor. Perhitungan nilai OR dilakukan dengan menggunakan Persamaan (2.26) untuk setiap Kabupaten/Kota. Sebagai contoh, berikut ditampilkan nilai OR untuk Kabupaten Bantaeng pada Tabel 3.3.

**Tabel 3.3.** Nilai *Odds Ratio* Kabupaten Bantaeng

Variabel	$\hat{\beta}^{LASSO}$	Nilai OR
Banyak tempat pelayanan KB ( $x_1$ )	-	-
Persentase wanita dengan pendidikan SMP/setara atau lebih rendah ( $x_2$ )	0,253	1,288
Pengeluaran rata-rata perkapita masyarakat ( $x_3$ )	-	-
Banyak wanita bekerja ( $x_4$ )	-	-
Banyak wanita usia subur ( $x_5$ )	-0,304	0,738
Persentase suami/keluarga yang menolak KB ( $x_6$ )	-0,651	0,521
Banyak petugas lapangan KB ( $x_7$ )	0,779	2,180

Berdasarkan Tabel 3.3, diperoleh bahwa persentase wanita dengan pendidikan SMP/setara atau lebih rendah ( $x_2$ ) dan banyak petugas lapangan KB ( $x_7$ ) memiliki nilai OR yang lebih besar dari satu. Ini berarti kedua variabel tersebut berpengaruh positif terhadap persentase *unmet need* KB. Hal ini sesuai dengan penelitian yang dilakukan oleh [38] dan [42]. Selanjutnya, banyak wanita usia subur ( $x_5$ ) dan persentase suami/keluarga yang menolak KB ( $x_6$ ) memiliki nilai OR yang lebih kecil dari satu. Ini mengindikasikan bahwa kedua variabel tersebut berpengaruh negatif terhadap persentase *unmet need* KB. Hal ini berlawanan dengan penelitian yang dilakukan oleh [26]. Adapun interpretasi dari nilai OR pada Tabel 3.3 adalah sebagai berikut:

1. Variabel  $x_2$  yaitu persentase wanita dengan pendidikan SMP/setara atau lebih rendah memiliki nilai OR sebesar 1,288. ini berarti setiap kenaikan persentase wanita dengan pendidikan SMP/setara atau lebih rendah sebesar 1 persen cenderung untuk meningkatkan persentase *unmet need* KB di Kabupaten Bantaeng menjadi 1,288 kali dibandingkan persentase *met need* KB.
2. Variabel  $x_5$  yaitu banyak wanita usia subur memiliki nilai OR sebesar 0,738. ini berarti setiap penambahan 1 orang wanita usia subur cenderung untuk menurunkan persentase *unmet need* KB di Kabupaten Bantaeng menjadi 0,738 kali dibandingkan persentase *met need* KB.

3. Variabel  $x_6$  yaitu persentase suami/keluarga yang menolak KB memiliki nilai OR sebesar 0,521. ini berarti setiap kenaikan persentase suami/keluarga yang menolak KB sebesar 1 persen cenderung untuk menurunkan persentase *unmet need* KB di Kabupaten Bantaeng menjadi 0,521 kali dibandingkan persentase *met need* KB.
4. Variabel  $x_7$  yaitu banyak petugas lapangan KB memiliki nilai OR sebesar 2,180. ini berarti setiap penambahan 1 orang petugas lapangan KB cenderung untuk meningkatkan persentase *unmet need* KB di Kabupaten Bantaeng menjadi 2,180 kali dibandingkan persentase *met need* KB.

#### 4. KESIMPULAN

Dalam mengidentifikasi faktor yang memengaruhi *unmet need* KB di Sulawesi Selatan tahun 2021, model GWLR dengan LASSO menghasilkan model dan jumlah variabel signifikan yang berbeda untuk setiap Kabupaten/Kota di Sulawesi Selatan. Pada 22 Kabupaten/Kota *unmet need* KB di wilayah tersebut dipengaruhi oleh persentase wanita dengan pendidikan SMP/setara atau lebih rendah ( $x_2$ ), banyak wanita usia subur ( $x_5$ ), persentase suami/keluarga yang menolak terhadap KB ( $x_6$ ), dan banyak petugas lapangan KB ( $x_7$ ). Sementara pada 2 Kabupaten/Kota *unmet need* KB di wilayahnya dipengaruhi oleh banyak wanita usia subur ( $x_5$ ), persentase suami/keluarga yang menolak terhadap KB ( $x_6$ ), dan banyak petugas lapangan KB ( $x_7$ ). Berikut model yang terbentuk untuk wilayah Kabupaten Bantaeng dan Kota Parepare yang merupakan bagian dari model GWLR dengan LASSO.

Kabupaten Bantaeng:

$$\pi(u_1, v_1) = \frac{\exp(-0,215 + 0,253x_2 - 0,304x_5 - 0,651x_6 + 0,779x_7)}{1 + \exp(-0,215 + 0,253x_2 - 0,304x_5 - 0,651x_6 + 0,779x_7)}$$

Kota Parepare:

$$\pi(u_{11}, v_{11}) = \frac{\exp(-0,268 - 0,860x_5 - 0,872x_6 + 1,273x_7)}{1 + \exp(-0,268 - 0,860x_5 - 0,872x_6 + 1,273x_7)}$$

Pada penelitian selanjutnya dapat dikaji dan dibandingkan lebih lanjut terkait metode lain yang dapat digunakan untuk mengatasi multikolinearitas pada model GWLR, seperti metode *group LASSO* atau dengan melakukan pemodelan *geographically weighted elastic net logistic regression* (GWENLR). Selain itu, penelitian selanjutnya juga dapat mengkaji lebih lanjut terkait faktor lain yang dianggap memiliki pengaruh terhadap *unmet need* KB di Sulawesi Selatan.

#### DAFTAR PUSTAKA

- [1] Aliu, M. A., Zubedi, F., Yahya, L., & Oroh, F. A., 2022. The Comparison of Kernel Weighting Functions in Geographically Weighted Logistic Regression in Modeling Poverty in Indonesia. *Jurnal Matematika, Statistika dan Komputasi*, 18(3), 362–384. <https://doi.org/10.20956/j.v18i3.19567>
- [2] Alwi, W., Ermawati, & Husain, S., 2018. Analisis Regresi Logistik Biner Untuk Memprediksi Kepuasan Pengunjung Pada Rumah Sakit Umum Daerah Majene. *Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya)*, 6(1), 20–26.

- [3] Amalah, R., 2020. *Pemodelan Geographically Weighted Logistic Regression Dengan Metode Ridge* [Skripsi]. Universitas Hasanuddin.
- [4] Amalah, R., Jaya, A. K., & Sirajang, N., 2023. Pemodelan Geographically Weighted Logistic Regression dengan Metode Ridge. *ESTIMASI: Journal of Statistics and Its Application*, 4(2), 130–143. <https://doi.org/10.20956/ejsa.v4i2.12250>
- [5] Annasiah, F., & Prastuti, M., 2023. Peramalan Konsumsi Energi Listrik untuk Sektor Industri di PT PLN (Persero) Area Gresik Menggunakan Metode Time Series Regression dan ARIMA. *Jurnal SAINS dan Seni ITS*, 12(1), D96–D102.
- [6] Bahmid, N. A., 2018. *Metode Least Absolute Shrinkage and Selection Operator untuk Mengatasi Multikolinearitas pada Regresi Logistik Ordinal* [Skripsi]. Universitas Hasanuddin.
- [7] BKKBN, 2021. *Data Statistik Rutin BKKBN 2019-2021*.
- [8] Chen, S., Notodiputro, K. A., & Rahardiantoro, S., 2020. Penerapan Analisis LASSO Dan Group LASSO Dalam Mengidentifikasi Faktor-Faktor Yang Berhubungan Dengan Tuberkulosis Di Jawa Barat. *Indonesian Journal of Statistics and Its Applications*, 4(1), 39–54.
- [9] Fadliana, A., & Darajat, P. P., 2021. Pemetaan Faktor Risiko Stunting Berbasis Sistem Informasi Geografis Menggunakan Metode Geographically Weighted Regression. *IKRA-ITH Informatika: Jurnal Komputer dan Informatika*, 5(3), 91–102.
- [10] Findasari, & Himayati, A. I. A., 2023. Analisis Regresi Logistik Biner Pada Faktor Resiko Kejadian Tuberkulosis. *Jurnal Matematika, Sains, dan Teknologi*, 24(1), 1–14.
- [11] Hastuti, T., 2022. *Penerapan Model Geographically Weighted Logistic Regression (GWLR) Dengan Fungsi Pembobot Adaptive Gaussian Kernel Pada Data Kemiskinan Di Indonesia* [Skripsi]. Universitas Lampung.
- [12] He, Y., Zhao, Y., & Tsui, K. L., 2019. Exploring influencing factors on transit ridership from a local perspective. *Smart and Resilient Transportation*, 1(1), 2–16. <https://doi.org/10.1108/srt-06-2019-0002>
- [13] Hidayah, I., & Adnan, A., 2021. *Metode Regresi Logistik LASSO Untuk Analisis Gizi Buruk Pada Balita Di Sumatera Tengah*. 1–11.
- [14] Kasim, S. I. A., 2021. *Estimasi Parameter Model Geographically Weighted Logistic Regression Principal Component Analysis Dengan Maksimum Likelihood* [Skripsi]. Universitas Hasanuddin.
- [15] Khariyani, A. M., Kismiantini, & Setiawan, E. P., 2022. Analisis Faktor-Faktor Yang Memengaruhi Jumlah Tuberkulosis menggunakan Geographically Weighted Regression Di Provinsi Jawa Timur. *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, 16–26.
- [16] Kumar, C., Rangappa, K. B., & Suchitra, S., 2022. Effectiveness of online capacity building programs in wholistic development of faculties: an empirical analysis. *Asian Association of Open Universities Journal*, 17(2), 194–211. <https://doi.org/10.1108/AAOUJ-04-2022-0058>
- [17] Lanfranchi, M., Alibrandi, A., Zirilli, A., Sakka, G., & Giannetto, C., 2020. Analysis of the wine consumer's behavior: an inferential statistics approach. *British Food Journal*, 122(3), 884–895. <https://doi.org/10.1108/BFJ-08-2019-0581>

- [18] Lestari, F. D., Kusnandar, D., & Debararaja, N. N., 2020. Estimasi Parameter Model Geographically Weighted Logistic Regression. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, 09(1), 159–164.
- [19] Lestari, S. S. S., Meimela, A., & Revildy, W. D., 2020. Analisis Faktor Tingkat Pengangguran Terbuka dengan Metode Geographically Weighted Lasso (Studi Kasus di Provinsi Jawa Barat Tahun 2019). *Seminar Nasional Official Statistics 2020 : Statistics In The New Normal A Challenge Of Big Data And Official Statistics*, 1286–1293.
- [20] Lestari, V. D., 2020. *Model Geographically Weighted Logistic Regression dengan Fungsi Pembobot Adaptive Tricube (Studi Kasus: Indikator Pencemaran Air Dissolve Oxygen di Daerah Aliran Sungai Mahakam Kalimantan Timur Tahun 2018)*. Universitas Mulawarman.
- [21] Lestari, V. D., Suyitno, & Siringoringo, M., 2021. Analisis Faktor-Faktor Yang Berpengaruh Terhadap Pencemaran Air Sungai Mahakam Menggunakan Pemodelan Geographically Weighted Logistic Regression Pada Data Dissolved Oxygen. *Jurnal EKSPONENSIAL*, 12(1).
- [22] Maulidina, T. P., & Oktora, S. I., 2020. Analisis Spasial Ketertinggalan Daerah Di Indonesia Tahun 2018 Menggunakan Geographically Weighted Logistic Regression. *Indonesian Journal of Statistics and Its Applications*, 4(3), 528–544.
- [23] Mayfield, H. J., Lowry, J. H., Watson, C. H., Kama, M., Nilles, E. J., & Lau, C. L., 2018. Use of Geographically Weighted Logistic Regression to Quantify Spatial Variation in the Environmental and Sociodemographic Drivers of Leptospirosis in Fiji: a Modelling Study. *The lancet Planetary health*, 2(5), e223–e232. [www.thelancet.com/](http://www.thelancet.com/)
- [24] Meimela, A., 2021. MODELING OF THE NUMBER OF TUBERCULOSIS CASES IN INDONESIA. *Jurnal Litbang Sukowati : Media Penelitian dan Pengembangan*, 4(2), 79–85. <https://doi.org/10.32630/sukowati.v4i2.204>
- [25] Nakyeyune, G. K., Bananuka, J., Tumwebaze, Z., & Kezaabu, S., 2022. Knowledge management practices and sustainability reporting: the mediating role of intellectual capital. *Journal of Money and Business*. <https://doi.org/10.1108/jmb-06-2022-0032>
- [26] Nifueki, A. D., Riwu, R. R., & Sir, A. B., 2022. Analisis Pengambilan Keputusan Pasangan Usia Subur Untuk Tidak Berpartisipasi Dalam Program KB. *Cendana Medical Journal*, 23(1), 209–216.
- [27] Pardede, T. T., Sumargo, B., & Rahayu, W., 2022. Penerapan Regresi Least Absolute Shrinkage and Selection Operator (LASSO) Untuk Mengidentifikasi Variabel yang Berpengaruh terhadap Kejadian Stunting di Indonesia. *Jurnal Statistika dan Aplikasinya*, 6(1), 37–48.
- [28] Permai, S. D., Christina, A., & Santoso Gunawan, A. A., 2021. Fiscal decentralization analysis that affect economic performance using geographically weighted regression (GWR). *Procedia Computer Science*, 179, 399–406. <https://doi.org/10.1016/j.procs.2021.01.022>
- [29] Pratiwi, N., Suyitno, & Siringoringo, M., 2020. Penerapan Model Geographically Weighted Logistic Regression Pada Data Status Kesejahteraan Masyarakat di Kalimantan Tahun 2017. *Jurnal EKSPONENSIAL*, 11(1).
- [30] Pratiwi, Y. D., Mariani, S., & Hendikawati, P., 2020. Pemodelan Regresi Spasial Menggunakan Geographically Weighted Regression. *UNNES Journal of Mathematics*, 8(2), 32–41. <http://journal.unnes.ac.id/sju/index.php/ujm>



- [31] Purba, M., Budiati, E., & Djamil, A., 2020. Determinan Perilaku Yang Berhubungan Dengan Terjadinya Unmet Need KB Pada Pasangan Usia Subur (PUS) Di Kota Bandar Lampung. *Manuju: Malahayati Nursing Journal*, 2(3), 491–504.
- [32] Purnatirani, F. T., 2019. *Estimasi Parameter Model Geographically Weighted Logistic Regression (GWLRL) Dengan Fungsi Pembobot Gaussian Kernel* [Skripsi]. Universitas Lampung.
- [33] Rai, A., & Ramadhan, R. R., 2018. Faktor-Faktor Yang Memengaruhi Unmet Need KB Di Provinsi Bengkulu Tahun 2015 Dengan Pemodelan Regresi Logistik Biner. *Indonesian Journal of Statistics and Its Applications*, 2(1), 46–55.
- [34] Ramadhan, H. W., 2017. *Pendugaan Parameter Regresi Logistik Multinomial Menggunakan Metode Least Absolute Shrinkage And Selection Operator (LASSO) (Studi Kasus Pilihan Sekolah SMA Sederajat Oleh Siswa SMP dan MTs Negeri di Kabupaten Trenggalek)*.
- [35] Reskia, A., & Harison., 2022. *Analisis Spasial Tingkat Kesejahteraan Di Indonesia Menggunakan Geographically Weighted Logistic Regression*. 1–8.
- [36] Salsabila, R., Putra, A. A., Amalita, N., & Fitri, F., 2023. Analysis of Factors Influencing the Population Growth Rate in West Sumatra Using Geographically Weighted Logistic Regression. *UNP JOURNAL OF STATISTICS AND DATA SCIENCE*, 1(3), 196–202. <https://doi.org/10.24036/ujsds/vol1-iss3/59>
- [37] Sholicha, C. N., Budiantara, I. N., & Ratna, M., 2018. Regresi Nonparametrik Spline Truncated untuk Memodelkan Persentase Unmet Need di Kabupaten Gresik. *Jurnal Sains dan Seni ITS*, 7(2), D61–D68.
- [38] Sholicha, C. N., Budiantara, I. N., & Ratna, M., 2018. Regresi Nonparametrik Spline Truncated untuk Memodelkan Persentase Unmet Need di Kabupaten Gresik. *Jurnal Sains dan Seni ITS*, 7(2), D61–D68.
- [39] Sofiyat, A. I., Tjalla, A., & Mahdiyah., 2023. Pemodelan Regresi Logistik Biner Terhadap Penerimaan Pegawai Di PT XYZ Jakarta. *Matematika Sains*, 1(1), 1–11.
- [40] Solekha, N. A., & Qudratullah, M. F., 2022. Pemodelan Geographically Weighted Logistic Regression dengan Fungsi Adaptive Gaussian Kernel Terhadap Kemiskinan di Provinsi NTT. *Jambura Journal of Mathematics*, 4(1), 17–32. <https://doi.org/10.34312/jjom.v4i1.11452>
- [41] Sriningsih, M., Hatidja, D., & Prang, J. D., 2018. Penanganan multikolinearitas dengan menggunakan analisis regresi komponen utama pada kasus impor beras di Provinsi Sulut. *Jurnal Ilmiah Sains*, 18(1), 18–24.
- [42] Sulistiawan, D., Gustina, E., Matahari, R., & Marthasari, V., 2020. Profil Sosiodemografis Unmet Need Keluarga Berencana Pada Wanita Kawin Di Daerah Istimewa Yogyakarta. *Jurnal Keluarga Berencana*, 5(02), 1–9.
- [43] Sulistiawan, D., Gustina, E., Matahari, R., & Marthasari, V. M. U. A. D. Y. F., 2020. Profil Sosiodemografis Unmet Need Keluarga Berencana Pada Wanita Kawin Di Daerah Istimewa Yogyakarta. *Jurnal Keluarga Berencana*, 5(02), 1–9.
- [44] Uljanah, K., Winarni, S., & Mawarni, A., 2016. Hubungan Faktor Risiko Kejadian Unmet Need KB (Keluarga Berencana) Di Desa Adiwerna, Kecamatan Adiwerna, Kabupaten Tegal, Triwulan III Tahun 2016. *Jurnal Kesehatan Masyarakat*, 4(4), 204–2012. <http://ejournal-s1.undip.ac.id/index.php/jkm>

- [45] Wantoro, D. D., 2017. *Pendugaan Parameter Regresi Logistik Biner Dengan Metode Least Absolute Shrinkage And Selection Operator (LASSO) (Studi Pada Persentase Tingkat Kemiskinan Kota/Kabupaten di Jawa Timur Tahun 2014)*. Universitas Brawijaya.
- [46] Wardana, L. O., & Sari, L. K., 2020. Analisis Faktor-Faktor Yang Memengaruhi Eksploitasi Pekerja Anak Di Indonesia Menggunakan Regresi Logistik Biner. *Indonesian Journal of Statistics and Its Applications*, 4(3), 432–447.
- [47] Wirdiastuti, C., Syafitri, U. D., Sumertajaya, I. M., Rohaeti, E., & Rafi, M., 2023. Application Of LASSO For Identification Of Functional Groups With Significant Contributions To Antioxidant Activities Of Centella Asiatica. *Communications in Mathematical Biology and Neuroscience*, 1–17. <https://doi.org/10.28919/cmbn/7843>
- [48] Wulandari, 2018. Geographically Weighted Logistic Regression Dengan Fungsi Kernel Fixed Gaussian Pada Kemiskinan Jawa Tengah. *Indonesian Journal of Statistics and Its Applications*, 2(2), 101–112.
- [49] Yolanda, D., & Destri, N., 2019. Faktor Determinan Yang Mempengaruhi Kejadian Unmet Need KB Pada Pasangan Usia Subur Di Kelurahan Campago Ipuah Kecamatan Mandiangin Koto Selayan Kota Bukittinggi Tahun 2018. *Menara Ilmu*, 13(3), 10–15.