

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 *Indonesia Blue Economy Index*

Ekonomi biru merupakan konsep ekonomi makro yang berfokus pada pemanfaatan sumber daya laut secara berkelanjutan, dengan penekanan pada inovasi yang ramah lingkungan dalam produk dan layanan di sektor kelautan (Sarkar et al., 2022). Tujuan utamanya mendorong pertumbuhan ekonomi yang finansial dan bertanggung jawab terhadap lingkungan (Niner et al., 2022). Banyak negara mulai menyadari potensi besar sektor kelautan sebagai penggerak pertumbuhan ekonomi (Fenichel et al., 2020). Forum global seperti *Asia-Pacific Economic Cooperation* (APEC) menjadikan ekonomi biru sebagai agenda utama untuk meningkatkan produksi dan transaksi bisnis, dengan harapan keuntungan ekonomi sejalan dengan kesejahteraan masyarakat pesisir dan pelestarian ekosistem laut.

Konsep ekonomi biru di Indonesia diimplementasikan melalui pengembangan *Indonesia Blue Economy Index* (IBEI) pada tahun 2022 dengan dukungan ARISE+ Indonesia, sebagai alat pemantauan kemajuan penerapan ekonomi biru di berbagai wilayah, fokus pada tiga pilar utama:

- a. Pilar Lingkungan: Menilai kualitas ekosistem laut dan pesisir sebagai indikator keberlanjutan ekonomi berbasis kelautan. Lingkungan yang sehat menjadi dasar untuk pertumbuhan ekonomi berkelanjutan
- b. Pilar Ekonomi: Mengukur kontribusi sektor kelautan terhadap PDB nasional, termasuk perikanan, akuakultur, dan pariwisata, untuk menilai perannya dalam menciptakan lapangan kerja dan pertumbuhan ekonomi.
- c. Pilar Sosial: Berfokus pada inklusi sosial, pilar ini mengevaluasi bagaimana sektor kelautan meningkatkan taraf hidup masyarakat

Indonesia melalui pendapatan, akses pekerjaan layak, dan distribusi manfaat.

Ketiga pilar IBEI memberikan panduan komprehensif untuk mengukur kemajuan ekonomi biru di Indonesia. Indeks ini berfungsi sebagai alat evaluasi dan acuan bagi pembuat kebijakan untuk mengidentifikasi sektor yang memerlukan perhatian lebih demi pembangunan berkelanjutan.

2.1.2 Classification and Regression Trees

Metode *Classification and Regression Trees* (CART) digunakan untuk memprediksi nilai atau klasifikasi data yang hilang, terutama dalam analisis data spasial. CART efektif menangani data numerik dan kategorikal, serta berguna untuk mengisi data hilang (Mariani et al., 2024). Keunggulan CART terletak pada kemudahan interpretasi melalui struktur pohon, fleksibilitas dalam menangani berbagai tipe data, dan ketahanan terhadap data hilang dan outlier (Ozcan & Peker, 2023). Metode ini juga memiliki kelemahan, yaitu risiko overfitting jika pohon yang dibangun terlalu kompleks dan sensitivitas terhadap perubahan kecil dalam data (Hamzah et al., 2021). Prosesnya melibatkan pemisahan data secara rekursif hingga kriteria penghentian tercapai, dengan setiap daun pohon mewakili nilai prediksi atau kelas tertentu.

2.1.3 Penyusunan Indeks Ekonomi Biru

Penghitungan IEB ini akan mengadopsi penghitungan yang dilakukan *United Nation Development Programme* (UNDP) dalam menghitung Indeks Pembangunan Manusia (IPM). Perhitungan IPM terdapat 3 dimensi penyusunnya yaitu dimensi kesehatan, pendidikan, dan pengeluaran. Dimensi IEB juga mengadopsi pilar yang sudah dilakukan oleh IBEI, yaitu sosial, ekonomi, dan lingkungan. Penelitian ini mengadopsi perhitungan sesuai UNDP, kemudian menerapkannya dalam proses perhitungan secara matematis (Kurnia et al., 2023):

1. Normalisasi data menggunakan *min-max*

$$x'_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \quad (2.1)$$

2. Menghitung indeks dimensi dengan rata-rata aritmatik

$$I_j = \frac{\sum_{i=1}^n x'_{ij}}{n} \quad (2.2)$$

3. Melakukan perhitungan indeks dengan rata-rata geometrik

$$Indeks = \left(\prod_{j=1}^n I_j \right)^{\frac{1}{n}} \times 100 \quad (2.3)$$

di mana:

x_{ij} : data variabel ke i dimensi ke j

x'_{ij} : variabel ke i dimensi ke j yang dinormalisasi

I_j : indeks dimensi ke j

n : banyaknya dimensi

2.1.4 Seleksi Variabel

2.1.4.1 Group Least Absolute Shrinkage and Selection Operator

Regresi Linier Berganda (RLB) adalah teknik statistik yang digunakan untuk menganalisis hubungan antara satu variabel respon dan sejumlah variabel prediktor. Model regresi ini didefinisikan sebagai berikut (Sinaga et al., 2022):

$$y_i = \sum_{j=1}^p X_{ij}\beta_j + \varepsilon_i \quad (2.4)$$

di mana y_i merupakan nilai respon untuk sampel ke- i , sementara X_{ij} adalah nilai variabel prediktor ke- j . Koefisien regresi β_j menunjukkan pengaruh masing-masing prediktor terhadap variabel respons, dan ε_i merupakan komponen error yang mencerminkan ketidakpastian suatu model yang diharapkan mengikuti distribusi normal dengan $\mu = 0$ dan sebaran σ^2 atau $\varepsilon_i \sim N(0, \sigma^2)$. Model ini juga mengasumsikan adanya linearitas antara variabel serta independensi antara setiap observasi yang diperoleh.

Apabila terdapat hanya satu variabel prediktor, model tersebut dapat disederhanakan menjadi:

$$y_i = x_i^T \beta + \varepsilon_i \quad (2.5)$$

di mana x_i adalah vektor dari variabel prediktor yang termuat dalam X , dan β adalah vektor koefisien. Notasi $x_i^T \beta$ menggambarkan produk skalar antara dua vektor, yaitu X_i dan β yang menunjukkan adanya hubungan linier di antara keduanya. Model RLB juga dapat disajikan dalam bentuk matriks, Untuk meningkatkan efisiensi, yakni sebagai berikut:

$$y_i = X\beta + \varepsilon_i \quad (2.6)$$

di mana persamaan ini galat diasumsikan memiliki sebaran normal $\varepsilon \sim N(0, \sigma^2 I_n)$ dengan I_n adalah matriks identitas pada $\mathbb{R}^{n \times n}$.

Solusi dari model RLB dapat diformulasikan sebagai masalah optimasi. Solusi ini dinyatakan dalam bentuk:

$$\hat{\beta}^{OLS} = \arg \min \|y - X\beta\|_2^2 \quad (2.7)$$

Hasil solusi *Ordinary Least Square* (OLS) pada persamaan 2.8 dapat dihitung dengan asumsi bahwa X memiliki *rank* kolom penuh. Hal ini berarti $X^T X$ merupakan definit positif dan solusi OLS diberikan oleh rumus:

$$\hat{\beta}^{OLS} = (X^T X)^{-1} X^T y \quad (2.8)$$

namun, jika X tidak memiliki *rank* kolom penuh, solusi tidak dapat ditentukan secara unik. Penting untuk memastikan bahwa variabel prediktor yang dipilih memenuhi syarat ini agar hasil analisis menjadi relevan.

Estimasi koefisien OLS dapat memberikan estimasi varians minimum yang tidak bias dan konsisten. Metode ini tidak dapat digunakan ketika jumlah prediktor k melebihi jumlah observasi n , karena tidak ada solusi unik untuk persamaan tersebut. Oleh karena itu, metode *Least Absolute Shrinkage Selection Operator* (LASSO) menjadi solusi efektif untuk mereduksi dimensi dalam data. LASSO dapat dinyatakan dalam bentuk berikut (Emmert & Dehmer, 2019):

$$\begin{aligned} \hat{\beta} &= \arg \min \left\{ \frac{1}{2n} \sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2 \right\} \\ \text{fungsi kendala: } &\|\beta\|_1 \leq t \end{aligned} \quad (2.9)$$

di mana parameter t menunjukkan batas dari norma L1 untuk koefisien β . Pendekatan ini juga bisa ditulis dalam kerangka Lagrange, sebagai berikut:

$$\hat{\beta}^{L_1} = \arg \min \left\{ \frac{1}{2n} \sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \|\beta\|_1 \right\} \quad (2.10)$$

Bentuk fungsi 2.11 penalti L1 (λ) ditambahkan untuk mengurangi kompleksitas model, dimana besar penalti diatur oleh nilai λ . Kedua bentuk rumus tersebut melalui kondisi yang memastikan ada nilai $\lambda > 0, \forall \epsilon > 0$ yang menghasilkan solusi yang sama. Fungsi penalti L1 ini dinyatakan dalam bentuk *piecewise* untuk menentukan koefisien $\hat{\beta}^{L_1}$ berdasarkan nilai λ . Bentuk fungsi penalti L1 sebagai berikut:

$$\hat{\beta}^{L_1} = \begin{cases} \frac{1}{N} \langle z, y \rangle - \lambda & \text{jika } \frac{1}{N} \langle z, y \rangle > \lambda \\ 0 & \text{jika } \frac{1}{N} |\langle z, y \rangle| \leq \lambda \\ \frac{1}{N} \langle z, y \rangle + \lambda & \text{jika } \frac{1}{N} \langle z, y \rangle < -\lambda \end{cases} \quad (2.11)$$

di mana $\langle z, y \rangle$ merupakan produk skalar antara vektor z dan y . Nilai λ yang besar akan semakin banyak koefisien yang menyusut menjadi nol sehingga menciptakan model yang sederhana.

Parameter λ dalam metode LASSO perlu diestimasi untuk mendapatkan model yang optimal. Estimasi ini dapat dilakukan melalui metode *Cross Validation* (CV). *Mean Squared Error* (MSE) untuk setiap fold F_k dapat dihitung menggunakan:

$$e(\lambda) = \frac{1}{\#F_k} \sum_{j \in F_k} (y_i - \hat{y}_i)^2 \quad (2.12)$$

di mana $\#F_k$ merupakan jumlah sampel dalam himpunan F_k . Rata-rata MSE dari seluruh k -fold dapat ditentukan sebagai berikut:

$$CV(\lambda) = \frac{1}{K} \sum_{k=1}^K e(\lambda)_k \quad (2.13)$$

Nilai λ optimal didapatkan dengan meminimalkan fungsi $CV(\lambda)$ yaitu sebagai berikut:

$$\hat{\lambda}_{min} = \arg \min CV(\lambda) \quad (2.14)$$

Pengembangan dari LASSO yaitu *Group LASSO* yang berfokus pada dimensi/kelompok variabel (Huang et al., 2024). *Group LASSO* dirancang untuk situasi di mana sekelompok prediktor harus memiliki koefisien non-nol atau nol secara simultan. Mari kita asumsikan bahwa p prediktor dibagi menjadi G dimensi, dengan p_g sebagai jumlah prediktor dalam dimensi $g \in \{1, \dots, G\}$. Matriks $X_g \in \mathbb{R}^{n \times p_g}$ merupakan prediktor yang sesuai dengan dimensi g serta vektor koefisien regresi diberikan oleh $\beta_g \in \mathbb{R}^{p_g}$. *Group LASSO* menyelesaikan masalah optimasi konveks yang dinyatakan sebagai berikut:

$$\hat{\beta} = \arg \min \left\{ \frac{1}{2n} \|y - \sum_{g=1}^G X_g \beta_g\|_2^2 + \lambda \sum_{g=1}^G \sqrt{p_g} \|\beta_g\|_2 \right\} \quad (2.15)$$

di mana parameter p_g memperhitungkan variasi jumlah prediktor dalam masing-masing dimensi. Parameter p_g apabila bernilai 1 untuk semua dimensi g , maka *Group LASSO* berfungsi sama seperti LASSO biasa, sedangkan apabila parameter $p_g > 1$, *Group LASSO* memungkinkan penanganan lebih efektif dengan memperhatikan struktur dimensi.

2.1.4.2 Elastic Net

LASSO mampu menurunkan beberapa koefisien menjadi nol, namun metode ini dapat menjadi tidak stabil apabila terdapat banyak fitur yang berkorelasi tinggi. Oleh karena itu, dikembangkanlah metode Elastic Net yang menggabungkan regularisasi L1 dari LASSO dengan regularisasi L2 dari Ridge (Araveeporn, 2021). Ridge menyusutkan koefisien dengan penalti L2 yang dapat dinyatakan sebagai berikut:

$$\begin{aligned} \hat{\beta}^{RR} &= \arg \min \left\{ \frac{1}{2n} \sum_{i=1}^n \left(y_i - \sum_j x_{ij} \beta_j \right)^2 + \lambda \|\beta\|_2^2 \right\} \\ &= \arg \min \left\{ \frac{1}{2n} \text{RSS}(\beta) + \lambda \|\beta\|_2^2 \right\} \\ &= \arg \min \left\{ \frac{1}{2n} \|y - X\beta\|_2^2 + \lambda \|\beta\|_2^2 \right\} \end{aligned} \quad (2.16)$$

di mana $RSS(\beta)$ adalah *residual sum of squares*, $\lambda \|\beta\|_2^2$ adalah penalti regularisasi, sedangkan λ adalah parameter untuk mengontrol penyusutan koefisien.

Metode Ridge memiliki kekurangan yaitu tidak dapat mengatur koefisien menjadi nol, seluruh variabel akan dipertahankan dengan penyusutan koefisien saja. Hal ini dapat menyebabkan terjadinya *overfitting* model. Oleh karena itu, Elastic Net mampu mengatasi kekurangan dari LASSO dan Ridge dengan menggabungkan regularisasi L1 dan L2 yang dapat dinyatakan melalui penalti $P_\alpha(\beta)$ sebagai berikut:

$$P_\alpha(\beta) = \alpha \|\beta\|_2^2 + (1 - \alpha) \|\beta\|_1 \quad (2.17)$$

Asumsi bahwa prediktor dan respon telah distandarisasi, Elastic Net dapat dinyatakan sebagai berikut:

$$\hat{\beta} = \arg \min \left\{ \frac{1}{2n} \sum_{i=1}^N (y_i - \sum_j \beta_j x_{ij})^2 + \lambda P_\alpha(\beta) \right\} \quad (2.18)$$

Elastic Net sangat berguna ketika jumlah prediktor p jauh lebih besar dibandingkan jumlah observasi n , memungkinkan pemilihan lebih dari n prediktor. Hal ini juga membantu mendorong efek pengelompokan dengan mempertahankan kelompok prediktor yang saling berkorelasi, yang tidak dilakukan oleh LASSO maupun Ridge.

Parameter λ_1 (LASSO) dan λ_2 (Ridge) dalam persamaan Elastic Net sebagai berikut:

$$\alpha = \frac{\lambda_1}{\lambda_1 + \lambda_2} \quad (2.19)$$

Hasilnya adalah bentuk alternatif dari Elastic Net yaitu sebagai berikut:

$$\hat{\beta} = \arg \min \left\{ \frac{1}{2n} \|y - X\beta\|_2^2 + \lambda_2 \|\beta\|_2^2 + \lambda_1 \|\beta\|_1 \right\} \quad (2.20)$$

2.1.5 Standarisasi Data

Standarisasi data merupakan proses transformasi data ke dalam skala yang sama, sehingga memudahkan perbandingan antar variabel dan meningkatkan akurasi model analisis (Ngoc et al., 2022). Salah satu metode

standarisasi yang umum digunakan adalah Z-Score (Pambudi et al., 2023). Metode ini mengubah data menjadi nilai standar dengan mean 0 dan standar deviasi 1. Rumus Z-Score sebagai berikut:

$$z = \frac{x - \mu}{\sigma} \quad (2.21)$$

di mana z merupakan nilai standarisasi Z-Score; x merupakan nilai data observasi; serta μ dan σ didefinisikan sebagai nilai mean dan standar deviasi dari data.

2.1.6 Regresi Terboboti Geografis Multiskala

Analisis spasial memerlukan pemahaman bagaimana hubungan antar variabel dapat bervariasi di berbagai lokasi. Model regresi konvensional yang dikenal sebagai model global mengasumsikan bahwa hubungan tersebut tetap konstan di seluruh area pengamatan. Penerapan model RTG memiliki bentuk fungsi sebagai berikut (Cellmer et al., 2020):

$$y_i = \beta_{i0} + \sum_{j=1}^k \beta_{ij} x_{ij} + \epsilon_i, \quad i = 1, \dots, n \quad (2.22)$$

di mana y_i merepresentasikan pengamatan terhadap variabel respon di lokasi ke- i ; β_{i0} adalah *intersept*; x_{ij} adalah variabel prediktor ke- j di lokasi tersebut; β_{ij} adalah koefisien lokal, dan ϵ_i adalah galat. Vektor parameter β dihitung menggunakan OLS yang memiliki bentuk matriks sebagai berikut:

$$\hat{\beta} = [X^T X]^{-1} X^T y \quad (2.23)$$

di mana X merupakan matriks berukuran $m \times k$ yang berisi variabel prediktor, sedangkan y merupakan vektor observasi yang berukuran $m \times 1$.

Estimasi OLS ini tidak terindeks oleh i , sehingga mencerminkan hubungan antar variabel prediktor dan respon untuk seluruh lokasi. Kalibrasi model dilakukan dengan menggunakan data dari lokasi-lokasi tetangga untuk menyusun matriks bobot spasial W berukuran $n \times n$. Penyusunan W penting untuk menentukan skema peminjaman data termasuk pemilihan fungsi kernel dan parameter *bandwidth*. *Bandwidth* merupakan ukuran ketetanggaan dimana semakin besar nilai *bandwidth*

maka semakin banyak lokasi yang digunakan dalam analisis. Oleh karena itu, estimasi parameter untuk lokasi ke- i dapat diperoleh dengan menerapkan bobot dari estimasi OLS:

$$\hat{\beta}_i = [X^T W_i X]^{-1} X^T W_i y \quad (2.24)$$

di mana W_i merupakan matriks pembobot yang berukuran $n \times n$ didefinisikan sebagai $W_i = \text{diag}[w_{i1}, \dots, w_{in}]$. Matriks ini memberikan bobot pada setiap pengamatan berdasarkan jarak dari lokasi ke- i , yang dihitung menggunakan fungsi kernel dan *bandwidth* tertentu. β merupakan koefisien regresi sebagai vektor berdimensi $k \times 1$, sedangkan y adalah vektor dari variabel terikat berukuran $n \times 1$. Nilai prediksi untuk observasi pada lokasi ke- i diberikan sebagai berikut:

$$\hat{y}_i = X_i [X^T W_i X]^{-1} X^T W_i y \quad (2.25)$$

di mana hal ini setara dengan mengalikan estimasi parameter $\hat{\beta}_i$ untuk setiap lokasi dengan X_i . Seluruh vektor nilai prediksi dapat ditulis ulang sebagai berikut:

$$\begin{aligned} \hat{y} &= S y \\ S &= X_i [X^T W_i X]^{-1} X^T W_i \end{aligned} \quad (2.26)$$

di mana S sebagai matriks proyeksi yang menggambarkan bagaimana nilai prediksi dihasilkan untuk variabel respon. Matriks varian kovarian untuk parameter V_i dapat dibangun sebagai berikut:

$$\begin{aligned} C_i &= [X^T W_i X]^{-1} X^T W_i \\ V_i &= C_i C_i^T \hat{\sigma}^2 \end{aligned} \quad (2.27)$$

di mana $\hat{\sigma}^2$ adalah estimasi standar deviasi yang didefinisikan sebagai:

$$\hat{\sigma}^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{m - \text{tr}(S)} \quad (2.28)$$

di mana m adalah jumlah observasi dalam sampel. *Trace* dalam matriks proyeksi $\text{tr}(S)$ memberikan nilai penting yang disebut sebagai *Effective Number of Parameters* (ENP). ENP berfungsi sebagai pengatur jumlah

derajat kebebasan yang dipakai dalam model. Kesalahan standar untuk estimasi parameter dapat dinyatakan sebagai berikut:

$$se(\hat{\beta}_i) = \sqrt{diag(V_i)} \quad (2.29)$$

di mana $diag(V_i)$ adalah elemen diagonal matriks V_i .

Salah satu keterbatasan utama dari metode RTG adalah mengasumsikan bahwa setiap hubungan antar variabel dalam model memiliki ukuran ketetanggaan atau *bandwidth* yang seragam. Keterbatasan ini dapat diatasi melalui model Regresi Terboboti Geografis Multiskala (RTGM) yang menerapkan *bandwidth* yang berbeda untuk setiap hubungan variabel dalam model. Keunggulan dari RTGM adalah kemampuannya untuk menjelaskan heterogenitas spasial dengan lebih tepat dan mengurangi bias yang muncul selama proses estimasi parameter model RTGM dapat dijelaskan sebagai berikut:

$$y_i = b_{w0}(\beta_{0i}) + \sum_{j=1}^k b_{wj}(\beta_{ij}x_{ij}) + \epsilon_i, i = 1, \dots, n \quad (2.30)$$

di mana $b_{w0}, b_{w1}, \dots, b_{wk}$ adalah *bandwidth* yang digunakan untuk setiap komponen model. RTGM juga dapat disusun dalam bentuk *Generalized Additive Model* (GAM) yang dinyatakan sebagai:

$$y = f_{b_{w0}}(\beta_{0i}) + \sum_{j=1}^k f_{b_{wk}}(X_k) + \epsilon_i, i = 1, \dots, n \quad (2.31)$$

Fungsi $f_{b_{wk}} = b_{wk}(\beta_k)X_k$ adalah fungsi penghalus yang diterapkan pada variabel prediktor ke- j . Persamaan 2.31 akan diatur ulang untuk menentukan fungsi penghalus $f_{b_{wk}}$ dan menghitung parameter β_j dengan bentuk fungsinya sebagai berikut:

$$\begin{aligned}
f_{b_{wk}} &= E(y - \sum_{p \neq k} f_{b_{wp}} - \varepsilon | X_k) \\
&= E(y - \sum_{p \neq k} f_{b_{wp}} | X_k) \\
&= A_k(y - \sum_{p \neq k} f_{b_{wp}})
\end{aligned} \tag{2.32}$$

di mana A_k adalah matriks topi dari RTG univariat untuk X_k . Pembentukan matriks A_k ditunjukkan sebagai berikut:

$$A_k = \begin{pmatrix} x_{1k}(X_k^T W_1 X_k)^{-1} X_k^T W_1 \\ \vdots \\ x_{nk}(X_k^T W_n X_k)^{-1} X_k^T W_n \end{pmatrix}_{n \times n} \tag{2.33}$$

di mana W_k adalah matriks bobot spasial yang dihitung berdasarkan *bandwidth* spesifik untuk setiap variabel. Semua komponen additif akan dimasukkan ke dalam format matriks yang ditunjukkan sebagai berikut:

$$\begin{bmatrix} I & A_1 & \dots & A_1 \\ A_2 & I & \dots & A_2 \\ \dots & \dots & \dots & \dots \\ A_k & A_k & \dots & I \end{bmatrix} \begin{bmatrix} f_{bw_1} \\ f_{bw_2} \\ \dots \\ f_{bw_k} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_k \end{bmatrix} y \tag{2.34}$$

Matriks 2.34 dapat ditulis kembali sebagai:

$$Pf = Qy \tag{2.35}$$

di mana P adalah matriks berukuran $nk \times nk$ dan Q adalah matriks berukuran $nk \times n$. Matriks ini akan direstrukturasasi dengan memberikan komponen additif $f_{bw_1}, f_{bw_2}, \dots, f_{bw_k}$ sebagai berikut:

$$f = \begin{bmatrix} f_{bw_1} \\ f_{bw_2} \\ \dots \\ f_{bw_k} \end{bmatrix} = P^{-1}Qy \tag{2.36}$$

Matriks topi kovariat dapat diperoleh dengan:

$$R = \begin{bmatrix} R_1 \\ R_2 \\ \dots \\ R_k \end{bmatrix} = P^{-1}Q \tag{2.37}$$

Persamaan 2.35 memberikan bentuk tertutup untuk estimasi parameter RTGM dan inferensinya. Solusi dari hal ini akan diterapkan prosedur *backfitting*. Proses ini menghasilkan parameter *bandwidth* untuk proses k yang dimodelkan. Hal ini mengisyaratkan bahwa setiap bagian model terdiri dari model RTG univariat dan setiap iterasi dalam prosedur *backfitting* melibatkan kalibrasi RTG.

Kalibrasi *backfitting* RTGM dimulai dengan menetapkan nilai awal untuk estimasi parameter lokal $\hat{\beta}_i^{RTGM}$. Ketepatan nilai awal $\hat{\beta}$ dibandingkan dengan nilai konvergen akhir berpengaruh pada jumlah iterasi yang diperlukan. Nilai awal yang semakin dekat dengan nilai akhir yang diinginkan, semakin sedikit iterasi yang digunakan. Nilai-nilai awal akan ditentukan, selanjutnya nilai prediksi $\hat{y}$ dan residual $\hat{\varepsilon}$ dihitung. Iterasi dalam proses *backfitting*, nilai terkini dari $\hat{f}_k + \hat{\varepsilon}$ diregresikan terhadap X_k . Hal ini menghasilkan nilai sementara untuk *bandwidth* optimal bw_k serta memperbarui estimasi parameter lokal $\hat{\beta}_k$, fungsi penghalus $\hat{f}_k$ dan residual $\hat{\varepsilon}$ diperbarui.

Setiap iterasi menghasilkan nilai-nilai terkini dari $\hat{y}$ dan $\hat{\varepsilon}$ digunakan untuk mengevaluasi apakah perbedaan antara nilai-nilai tersebut dan nilai sebelumnya sudah cukup kecil untuk menunjukkan bahwa algoritma *backfitting* telah mencapai konvergensi. Kriteria *Score of Change* (SOC) dapat diterapkan untuk mengukur konvergensi ini, yang didasarkan pada *residual sum of squares* (RSS) model RTGM. Kriteria ini ditunjukkan sebagai berikut:

$$SOC_{RSS} = \frac{RSS_{baru} - RSS_{lama}}{RSS_{baru}} \quad (2.38)$$

Konvergensi tercapai ketika SOC menjadi lebih kecil dari nilai ambang tertentu η yang biasa diatur sebesar 10^{-5} .

Matriks topi spesifik kovariat R_k memetakan variabel respon y ke setiap komponen model yang diestimasi $\hat{f}_k$ dalam prosedur *backfitting* dengan persamaan sebagai berikut:

$$\hat{f}_k = R_k y \quad (2.39)$$

di mana R_k dihitung sebagai:

$$R_k = A_k (I - \sum_{p \neq k} R_p) \quad (2.40)$$

Perhitungan R_k , A_k , dan R_p perlu dilakukan dan diperbarui setelah setiap iterasi *backfitting*. Nilai akhir dari R_k didapat setelah konvergensi tercapai, kemudian dapat dijumlahkan untuk memperoleh matriks topi keseluruhan S untuk model RTGM:

$$S = \sum_k R_k \quad (2.41)$$

Matriks topi R_k dapat digunakan untuk menghitung ENP yang spesifik sebagai berikut:

$$ENP_k = tr(R_k) \quad (2.42)$$

ENP model secara keseluruhan dapat dihitung dengan:

$$ENP_{model} = \sum_k ENP_k \quad (2.43)$$

2.1.7 Regresi Terboboti Geografis dan Temporal Multiskala

RTG dan RTGM telah memberikan pengertian yang baik tentang hubungan spasial dan variabilitas dalam data, namun keduanya masih memiliki keterbatasan terutama dalam memahami dinamika yang melibatkan aspek temporal. Regresi Terboboti Geografis dan Temporal (RTGT) muncul sebagai solusi untuk mengatasi masalah tersebut. Model RTGT dirumuskan sebagai berikut (Liu & Dong, 2021):

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{j=1}^p \beta_k(u_i, v_i, t_i) x_{ij} + \varepsilon_i \quad (2.44)$$

di mana y_i adalah nilai yang diamati dari variabel respon untuk lokasi observasi (u_i, v_i) dan waktu t_i ; parameter $\beta_0(u_i, v_i, t_i)$ adalah konstanta;

parameter $\beta_k(u_i, v_i, t_i)$ adalah koefisien regresi untuk variabel ke- k pada lokasi (u_i, v_i) dan waktu t_i ; parameter x_{ij} adalah variabel prediktor ke- j pada pengamatan ke- i ; dan ε adalah kesalahan pada pengamatan ke- i yang diasumsikan $\varepsilon_i \sim N(0, \sigma^2)$.

Koefisien regresi $\hat{\beta}(u_i, v_i, t_i)$ pada titik ke- i dapat diperoleh dengan menggunakan *Weighted Least Square* (WLS). Parameter yang diestimasi dari model RTGT sebagai berikut:

$$\hat{\beta}(u_i, v_i, t_i) = (X^T W(u_i, v_i, t_i) X)^{-1} X^T W(u_i, v_i, t_i) Y \quad (2.45)$$

di mana matriks bobot $W(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{in})$ adalah matriks bobot pada lokasi observasi (u_i, v_i) dan waktu t_i . Elemen diagonal w_{ij} (dengan $1 \leq j \leq n$) adalah fungsi jarak *spatio-temporal* di titik observasi (u_i, v_i, t_i) .

Fungsi jarak *spatio-temporal* merupakan gabungan dari fungsi jarak spasial dan temporal menggunakan jarak *Euclidean* yang ditunjukkan sebagai berikut (Sifriyani et al., 2022):

$$\begin{cases} (d_{ij}^s)^2 = (u_i - u_j)^2 + (v_i - v_j)^2 \end{cases} \quad (2.46)$$

$$\begin{cases} (d_{ij}^t)^2 = (t_i - t_j)^2 \end{cases} \quad (2.47)$$

$$\begin{cases} (d_{ij}^{sT})^2 = \phi^s [(u_i - u_j)^2 + (v_i - v_j)^2] + \phi^T [(t_i - t_j)^2] \end{cases} \quad (2.48)$$

di mana ϕ^s dan ϕ^T adalah faktor-faktor yang mempengaruhi untuk menyeimbangkan efek yang berbeda saat mengukur jarak *spatio-temporal*. Berdasarkan fungsi jarak tersebut, bobot geografis diperoleh sebagai berikut:

$$\begin{aligned} w_{ij} &= \exp \left\{ - \left(\frac{\phi^s [(u_i - u_j)^2 + (v_i - v_j)^2] + \phi^T [(t_i - t_j)^2]}{b_{sT}^2} \right) \right\} \\ &= \exp \left\{ - \left(\frac{[(u_i - u_j)^2 + (v_i - v_j)^2]}{b_s^2} + \frac{[(t_i - t_j)^2]}{b_T^2} \right) \right\} \end{aligned} \quad (2.49)$$

Nilai dari b_s^2 dan b_T^2 kemudian digunakan untuk menghasilkan persamaan:

$$w_{ij} = \exp \left\{ - \left(\frac{(d_{ij}^s)^2}{b_s^2} + \frac{(d_{ij}^T)^2}{b_T^2} \right) \right\} \quad (2.50)$$

di mana b_s adalah *bandwidth* spasial dan b_T adalah *bandwidth* temporal. Nilai ϕ^s dan ϕ^T tidak sama dengan nol. Parameter nilai τ adalah rasio dari $\frac{\phi^s}{\phi^T}$ dengan $\phi^s \neq 0$. Persamaan parameter tersebut ditunjukkan sebagai berikut:

$$\frac{(d_{ij}^{sT})^2}{\phi^s} = (u_i - u_j)^2 + (v_i - v_j)^2 + \tau [(t_i - t_j)^2] \quad (2.51)$$

Asumsikan $\phi^s = 1$ untuk mengurangi parameter yang tidak diketahui. Terdapat satu parameter yang tidak diketahui yaitu τ . Parameter τ berfungsi untuk meningkatkan atau mengurangi efek jarak temporal terhadap jarak spasial. Parameter ini diperoleh dari validasi silang sebagai berikut:

$$(\tau) = \sum_i (y_i - \hat{y}_{\neq i}(\tau))^2 \quad (2.52)$$

Matriks pembobot W_{ij} ditentukan oleh jarak *spatio-temporal* dan *bandwidth spatio-temporal*. Nilai prediksi dapat dibentuk seperti berikut:

$$\begin{aligned} \hat{y} &= \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} X_1^T (X^T W(u_1, v_1, t_1) X)^{-1} X^T W(u_1, v_1, t_1) \\ X_2^T (X^T W(u_2, v_2, t_2) X)^{-1} X^T W(u_2, v_2, t_2) \\ \vdots \\ X_n^T (X^T W(u_n, v_n, t_n) X)^{-1} X^T W(u_n, v_n, t_n) \end{bmatrix} y \\ &= Sy \end{aligned} \quad (2.53)$$

Tabel 2.1 menunjukkan fungsi kernel yang akan digunakan terbagi menjadi beberapa jenis, yaitu sebagai berikut (Al-Hasani et al., 2021):

Tabel 2. 1 Fungsi Kernel

Fungsi	Spesifikasi	
<i>Gaussian</i>	$w_{ij} = \exp \left(-\frac{1}{2} \left(\frac{ d_{ij}^{ST} }{b} \right)^2 \right)$	(2.54)
<i>Uniform</i>	$w_{ij} = \begin{cases} 1, & \text{jika } d_{ij} \leq b \\ 0, & \text{lainnya} \end{cases}$	(2.55)
<i>Bisquare</i>	$w_{ij} = \begin{cases} \left(1 - \left(\frac{ d_{ij}^{ST} }{b} \right)^2 \right), & \text{untuk } d_{ij} \leq b \\ 0, & \text{lainnya} \end{cases}$	(2.56)
<i>Exponential</i>	$w_{ij} = \exp \left(-\left(\frac{ d_{ij}^{ST} }{b} \right) \right)$	(2.57)

Sumber: Al-Hasani et al., 2021

Fungsi kernel dalam metode ini digunakan untuk mengukur hubungan antar variabel dalam model. Kernel ini memberikan bobot pada titik data berdasarkan jarak antar variabel, yang memungkinkan model menangkap variabilitas antar lokasi dan waktu. Tabel 2.1 menunjukkan empat fungsi kernel yang digunakan dalam penelitian ini: *gaussian*, *uniform*, *bisquare*, dan *exponential*. Setiap kernel memberikan bobot yang berbeda sesuai dengan jarak antar titik data. Penelitian ini menguji keempat kernel tersebut untuk memastikan pemilihan kernel terbaik yang dapat memberikan hasil optimal dalam model.

RTGT memiliki kelemahan yang sama dengan RTG, yaitu mengasumsikan bahwa setiap hubungan antar variabel dalam model mempunyai ukuran *bandwidth* yang seragam. Hal ini dapat menghambat kemampuan model untuk menangkap variasi yang ada antara lokasi dan waktu. Regresi Terboboti Geografis dan Temporal Multiskala (RTGTM) dikembangkan untuk mengatasi keterbatasan ini (Liu et al., 2021). Metode ini menggunakan fungsi bobot yang bervariasi berdasarkan skala spasial dan temporal, sehingga memungkinkan untuk mengidentifikasi pola hubungan yang berbeda pada skala yang berbeda. Bentuk persamaan RTGTM sebagai berikut:

$$Y = \beta_{bwt_0s_0} + \beta_{bwt_1s_1} \otimes x_1 + \cdots + \beta_{bwt_ps_p} \otimes x_p + \varepsilon \quad (2.58)$$

di mana $\otimes$ merupakan perkalian *element-wise* matriks, dan bwt_ps_p merupakan spasial dan temporal *bandwidth* yang digunakan untuk penentuan koefisien variabel. Konsep algoritma *backfitting* RTGTM sama dengan yang dilakukan RTGM.

Awalan dari algoritma *backfitting* RTGTM ini dengan membuat perkiraan semua item additif seperti $\hat{f}_0^0, \hat{f}_1^1, \dots, \hat{f}_p^p$. Kesalahan awal akan didefinisikan sebagai:

$$\varepsilon = y - \hat{f}_0^0 - \hat{f}_1^1 - \cdots - \hat{f}_p^p \quad (2.59)$$

di sini akan ditetapkan nilai ambang η sebesar 10^{-5} . Perubahan SOC antara setiap iterasi dapat dihitung berdasarkan perubahan *Change in the Additive Terms* (CAT). Perubahan ini akan ditunjukkan sebagai berikut:

$$SOC_{RSS} = \frac{RSS_{baru} - RSS_{lama}}{RSS_{baru}} \quad (2.60)$$

$$SOC_{CAT} = \sqrt{\frac{\sum_{j=1}^p \frac{\sum_{i=1}^n (\hat{f}_{ij}^{(k+1)} - \hat{f}_{ij}^{(k)})^2}{n}}{\sum_{i=1}^n \left(\sum_{j=1}^p \hat{f}_{ij}^{(k+1)} \right)^2}}$$

Seluruh komponen akan diregresikan terhadap variabel yang sesuai. Kesalahan untuk setiap komponen akan diperbarui. Apabila dalam iterasi nilai $SOC > \eta$ maka dapat dilanjutkan untuk memperbarui nilai kesalahan, namun apabila SOC lebih kecil dari η maka proses akan dihentikan. Koefisien estimasi pada iterasi terakhir adalah koefisien yang diperkirakan. Oleh karena itu, nilai fitting akhir dari Y dinyatakan sebagai berikut:

$$\hat{Y}^{k+1} = \hat{f}_0^{(k+1)} + \hat{f}_1^{(k+1)} + \cdots + \hat{f}_p^{(k+1)} \quad (2.61)$$

2.1.7 Goodness of Fit

2.1.7.1 Koefisien Determinasi (R^2)

Koefisien determinasi (R^2) merupakan salah satu ukuran yang digunakan untuk menilai sejauh mana suatu model dapat menjelaskan variasi dalam data yang teramati (Chicco et al., 2021). Parameter ini memberikan gambaran tentang seberapa baik model yang diusulkan cocok dengan data. Nilai R^2 dihitung dengan rumus berikut:

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (2.62)$$

di mana SSE (*Sum of Squares Error*) menunjukkan total kesalahan dalam prediksi model, sedangkan SST (*Total Sum of Squares*) mencerminkan total variasi dalam data. R^2 memberikan informasi tentang proporsi variasi dalam variabel independent yang dapat dijelaskan oleh model yang dibangun.

2.1.7.2 Akaike Information Criteria (AIC)

Akaike Information Criteria (AIC) adalah salah satu metode yang paling sering digunakan dalam pemilihan model dalam analisis statistik (DelSole & Tippet, 2021). Metode ini berakar pada prinsip *maximum likelihood* yang secara konvensional digunakan untuk menaksir parameter model setelah struktur dan ukuran model ditetapkan. Fungsi *maximum likelihood* yang menilai seberapa besar kemungkinan data diamati sesuai parameter model ditulis sebagai:

$$\mathcal{L}(\mu, \sigma | y_1, \dots, y_n) = f(y_1, \dots, y_n | \mu, \sigma) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\sum_{i=1}^n \frac{(y_i - \mu)^2}{2\sigma^2}} \quad (2.63)$$

di mana f adalah fungsi probabilitas yang menggambarkan peluang pengamatan y berdasarkan parameter μ (rata-rata) dan σ (simpangan baku). Fungsi ini dioptimalkan dengan log-likelihood untuk memaksimalkan kecocokan model, yang dirumuskan sebagai:

$$\ln \mathcal{L}(\hat{\mu}_{MLE}, \hat{\sigma} | y_1, \dots, y_n) = \max_{\mu, \sigma} \ln \mathcal{L}(\mu, \sigma | y_1, \dots, y_n) \quad (2.64)$$

Gagasan utama Akaike adalah menggabungkan estimasi parameter dengan proses pemilihan model berdasarkan kompleksitasnya. AIC dihitung melalui rumus berikut:

$$AIC = -2 \ln \left(\mathcal{L}(\hat{\theta}_{MLE} | y) \right) + 2k \quad (2.65)$$

di mana:

$\mathcal{L}$: fungsi *likelihood*

$\hat{\theta}_{MLE}$: nilai parameter maksimum *likelihood*

k : jumlah parameter bebas dalam model



2.2 Hasil Penelitian Terdahulu

Beberapa penelitian terdahulu yang relevan dengan metode yang digunakan akan ditunjukkan pada Tabel 2.2, yaitu sebagai berikut:

Tabel 2. 2 Hasil Penelitian Terdahulu

No	Penulis (Tahun)	Judul Penelitian	Metode	Hasil
1	(Aghnyn et al., 2024)	<i>Mapping Blue Economy Potential Using Spatial Statistical Downscaling Model: Analysis of the Impact of Climate Change on Freshwater Fish Resources</i>	Digunakan analisis Spatial Statistical Downscaling with INLA tentang ekonomi biru	Kecamatan Kalipuro memiliki potensi ekonomi biru tertinggi dan Kalibaru terendah, dengan temuan yang didasarkan pada analisis data satelit dan metode <i>Statistical Downscaling</i> , yang mengungkapkan variasi signifikan antar wilayah serta pentingnya integrasi data dan pendekatan berkelanjutan dalam pengelolaan sumber daya kelautan untuk kebijakan yang lebih efektif.

No	Penulis (Tahun)	Judul Penelitian	Metode	Hasil
2	(Huang et al., 2024)	<i>Lasso and Group Lasso with Categorical Predictors: Impact of Coding Strategy on Variable Selection and Prediction</i>	Menggunakan 7 variabel yang terbagi menjadi beberapa kategori kemudian di seleksi menggunakan LASSO dan Group LASSO	<i>Group</i> LASSO lebih konsisten dalam pemilihan variabel dibandingkan LASSO, meskipun akurasi prediksi tetap dipengaruhi oleh strategi pengkodean, dengan solusi yang meliputi penggunaan pendekatan transparan dalam pelaporan pengkodean dan pertimbangan antara akurasi prediksi dan seleksi variabel.
3	(Bastiaan et al., 2022)	<i>Comparing the Performance of Prediction Model of Ridge and Elastic Net in Correlated Dataset</i>	Menggunakan 95 variabel independen yang kemudian diseleksi menggunakan Ridge dan Elastic Net	Elastic Net lebih efektif daripada Ridge regression dan OLS dalam menangani multikolinieritas pada dataset dengan korelasi tinggi, menghasilkan MSE lebih rendah dan model prediksi yang lebih akurat.

No	Penulis (Tahun)	Judul Penelitian	Metode	Hasil
4	(Liu et al., 2021)	<i>Prediction of PM2.5 concentrations at unsampled points using multiscale geographically and temporally weighted regression</i>	Membandingkan metode RTGT dan RTGTM dengan menggunakan kernel <i>bisquare</i>	Penelitian ini mengembangkan model RTGTM untuk memprediksi konsentrasi PM2.5 di titik tanpa sampel, menghasilkan $R^2 = 0,94$ dan RMSE rendah, serta pengaruh skala yang berbeda pada kovariat.
5	(Wu et al., 2019)	<i>Multiscale geographically and temporally weighted regression: exploring the spatiotemporal determinants of housing prices</i>	Membandingkan metode RTGT dan RTGTM dengan menggunakan kernel <i>gaussian</i>	mengembangkan model RTGTM dengan kernel <i>gaussian</i> untuk memetakan faktor-faktor yang mempengaruhi harga perumahan di Shenzhen, yang meningkatkan R^2 dan mengurangi AICc, serta memberikan rekomendasi kebijakan untuk perencanaan perkotaan.

Sumber: diolah oleh peneliti

Penelitian Aghnyn et al. (2024) mengkaji konsep ekonomi biru sebagai pendekatan berkelanjutan dalam memanfaatkan sumber daya di ekosistem laut, pesisir, dan daratan untuk mendukung pertumbuhan ekonomi yang seimbang dan ramah lingkungan. Penelitian tersebut menggunakan metode *Statistical Downscaling* untuk mengevaluasi dampak perubahan iklim terhadap sumber daya ikan air tawar dan memetakan potensi ekonomi biru melalui analisis data satelit, terutama curah hujan rata-rata. Hasilnya menunjukkan bahwa Kecamatan Kalipuro memiliki potensi ekonomi biru tertinggi, sementara Kalibaru terendah, mencerminkan variasi signifikan antar wilayah. Temuan tersebut menunjukkan bahwa penelitian Aghnyn et al. (2024) belum mengembangkan analisis di wilayah lain dengan karakteristik data yang berbeda, seperti data spasial dan temporal. Oleh karena itu, penelitian ini menggunakan pemodelan *spatio-temporal* dengan metode RTGTM untuk mendapatkan representasi spasial dan temporal yang lebih komprehensif mengenai potensi ekonomi biru di Pulau Sumatera, khususnya pada tingkat lokal.

Penelitian Huang et al. (2024) mengkaji pengaruh pemilihan strategi pengkodean terhadap hasil model LASSO dan *Group LASSO*, terutama dalam hal pemilihan variabel. LASSO sangat dipengaruhi oleh strategi pengkodean untuk menghasilkan variasi dalam pemilihan variabel dan akurasi prediksi. *Group LASSO* menunjukkan pemilihan variabel yang lebih konsisten meskipun akurasi prediksi tetap dipengaruhi oleh pilihan pengkodean. Bastiaan et al. (2022) menyoroti keunggulan Elastic Net dalam menangani multikolinieritas pada dataset dengan korelasi tinggi antar variabel independen yang menghasilkan nilai *Mean Squared Error* (MSE) lebih rendah dibandingkan dengan Ridge regression dan OLS. Elastic Net lebih efektif dalam menghasilkan model prediksi yang akurat pada data dengan korelasi tinggi antar variabel. Dua penelitian tersebut belum membandingkan secara langsung antara *Group LASSO* dan Elastic Net dalam konteks seleksi variabel. Oleh karena itu, penelitian ini

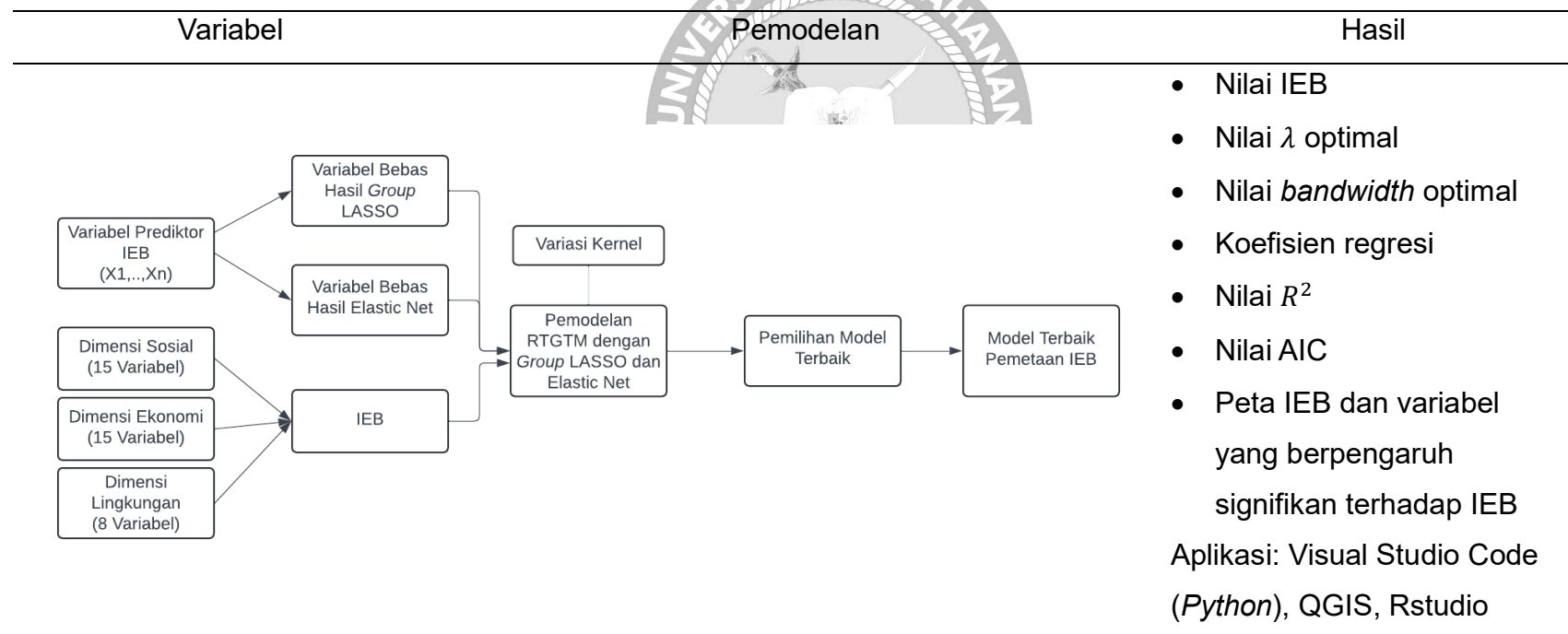
membandingkan *Group* LASSO dan Elastic Net untuk menentukan metode yang lebih efektif dalam memilih variabel yang signifikan. Hasil seleksi variabel dari kedua metode tersebut akan digunakan dalam pemodelan RTGTM untuk menghasilkan representasi spasial dan temporal yang lebih komprehensif mengenai potensi ekonomi biru dengan mempertimbangkan faktor-faktor dinamis yang mempengaruhi wilayah tersebut.

Penelitian Liu et al. (2021) mengembangkan model RTGTM untuk memprediksi konsentrasi PM2.5 di titik yang tidak teramati, mengatasi masalah ketidakmampuan model sebelumnya dalam menangani skala berbeda pada setiap kovariat, dan memberikan analisis ketidakpastian posterior. Hasil penelitian tersebut menunjukkan bahwa model RTGTM dengan kernel *bisquare* menunjukkan nilai R^2 yang sangat tinggi (0,94) dan RMSE yang rendah. Penelitian Wu et al. (2019) mengembangkan model RTGTM untuk mengeksplorasi pengaruh faktor-faktor yang signifikan terhadap harga perumahan di Shenzhen secara *spatio-temporal* dengan hasil yang lebih akurat menggunakan kernel *gaussian*. Dua penelitian tersebut hanya menggunakan satu kernel dalam pemodelan dan tidak melibatkan proses seleksi variabel untuk mengatasi multikolinearitas sebelum pemodelan. Oleh karena itu, penelitian ini membandingkan empat jenis kernel (*uniform*, *gaussian*, *bisquare*, dan *exponential*) serta menggunakan metode seleksi variabel *Group* LASSO dan Elastic Net untuk meningkatkan akurasi model RTGTM dalam pemodelan *spatio-temporal*.

2.3 Kerangka Berpikir

Berikut kerangka berpikir dalam penelitian ini yang diilustrasikan melalui diagram alur sederhana pada Tabel 2.3 untuk memvisualisasikan apa yang akan dilakukan dalam penelitian ini:

Tabel 2. 3 Kerangka Berpikir



Sumber: diolah oleh peneliti

Bagan pada Tabel 2.3 menggambarkan kerangka berpikir untuk menyusun IEB, yang dimulai dengan identifikasi variabel prediktor dari dimensi sosial, ekonomi, dan lingkungan. Langkah berikutnya adalah mencari variabel prediktor tambahan di luar IEB dan melakukan seleksi variabel menggunakan metode *Group LASSO* dan *Elastic Net*. Hasil seleksi dari kedua metode ini digunakan untuk membangun model RTGTM dengan variasi kernel, yang mengintegrasikan variabel spasial dan temporal dengan variasi kernel untuk meningkatkan akurasi. Proses ini menghasilkan pemilihan model terbaik yang mengoptimalkan nilai λ dan *bandwidth*, serta menghitung koefisien regresi. Hasil akhir mencakup nilai IEB, nilai R^2 , nilai AIC, serta peta IEB dan variabel-variabel yang berpengaruh signifikan terhadap IEB, seluruh analisis dilaksanakan menggunakan aplikasi Visual Studio Code dengan bahasa pemrograman *Python*.