

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Definisi Spatio-Temporal

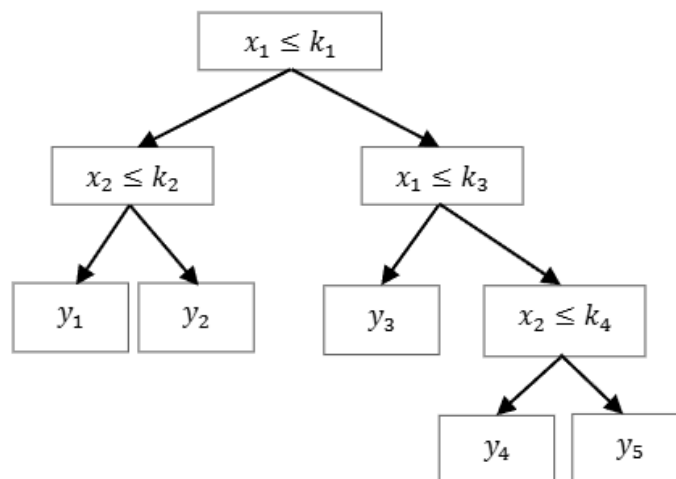
Statistik spatio-temporal merupakan metode statistik yang digunakan untuk menganalisis data dengan informasi lokasi dan waktu atau yang biasa disebut dengan data spatio-temporal. Data spatio-temporal merupakan data yang menggabungkan konsep data *cross-section* dan *time series*. Data *cross-section* merupakan data yang diambil pada satu titik waktu tertentu untuk berbagai entitas. Sementara itu, data *time series* adalah data yang diambil dalam periode waktu tertentu untuk satu entitas. Kombinasi antara dua konsep data tersebut biasanya disebut sebagai data panel (Baginski, 2019; Filgrima & Achmad, 2022).

2.1.2 *Classification and Regression Trees*

Algoritma *machine learning* yang dikenal sebagai metode *Classification and Regression Trees* (CART) digunakan untuk mengklasifikasikan atau memprediksi data dengan menggunakan konsep pohon keputusan. CART memiliki banyak keunggulan, seperti kemudahan interpretasi dan visualisasi struktur pohonnya, fleksibilitas dalam menangani berbagai jenis data, serta ketahanan terhadap data yang hilang dan *outlier*. Meskipun CART memiliki banyak keunggulan, metode tersebut juga memiliki kelemahan dalam hal *overfitting* jika pohonnya terlalu kompleks dan sensitif terhadap perubahan kecil dalam data.

CART menggunakan aturan pemisahan yang ditentukan oleh variabel prediktor untuk membagi data secara rekursif menjadi sub-sampel yang lebih kecil. Ketika kriteria penghentian dicapai dalam proses pemisahan tersebut, struktur pohon dalam algoritma CART dibentuk. Dalam regresi, setiap daun pohon menunjukkan nilai prediksi untuk variabel respons. Sementara itu, setiap daun pohon tersebut dalam klasifikasi

menunjukkan kelas tertentu. Ilustrasi lebih jelas untuk kerangka kerja metode CART dapat dilihat pada Gambar 2.1 (Hamzah et al., 2021).



Gambar 2.1 Kerangka Kerja Metode CART

Sumber: (Hamzah et al., 2021)

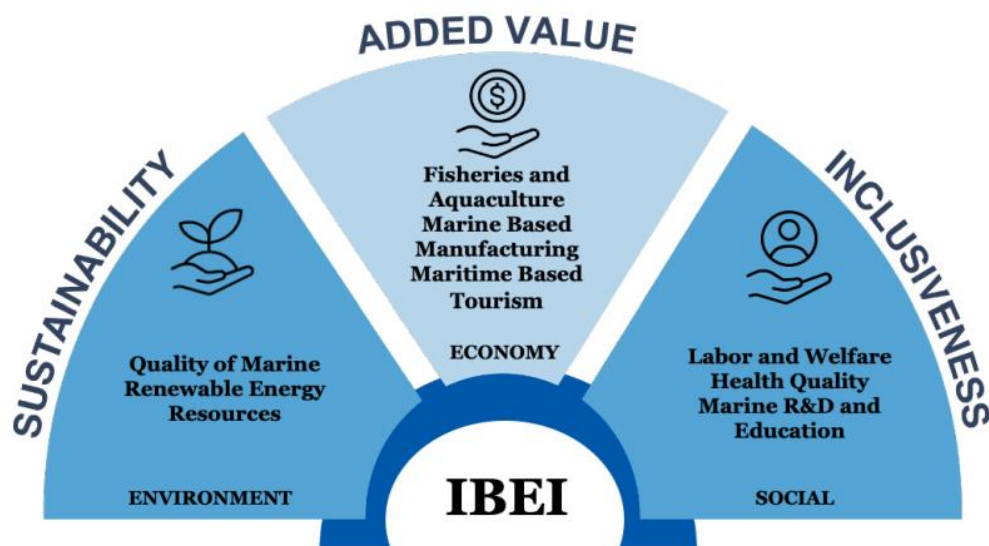
2.1.3 Nilai Indeks Ekonomi Biru

Nilai indeks ekonomi biru (IEB) berfungsi sebagai alat ukur untuk memantau potensi ekonomi biru di setiap kabupaten dan kota di Indonesia. Dimensi penyusun indeks tersebut mengacu pada kerangka kerja *Indonesia's Blue Economy Index* (IBEI) yang didukung oleh program ARISE+ Indonesia. Berdasarkan tujuan pembangunan berkelanjutan atau *Sustainable Development Goals* (SDGs), indeks ekonomi biru dapat dijelaskan oleh tiga dimensi utama, yakni lingkungan, ekonomi, dan sosial.

Pertama, dimensi lingkungan berfokus pada pengelolaan kualitas ekosistem laut sebagai indikator keberlanjutan ekonomi laut. Dimensi tersebut mencakup dua sektor utama, yaitu energi terbarukan serta sumber daya dan konservasi laut. Terdapat lima poin SDGs yang dapat dicapai oleh dimensi ini, yakni SDGs 6 (*Clean Water and Sanitation*), SDGs 7 (*Affordable and Clean Energy*), SDGs 12 (*Responsible Consumption and Production*), SDGs 13 (*Climate Action*), dan SDGs 14 (*Life Below Water*).

Kedua, dimensi ekonomi menunjukkan kontribusi sektor kelautan terhadap perekonomian Indonesia. Empat sektor utama yang diukur dalam dimensi tersebut meliputi perikanan tangkap dan budidaya (akuakultur), industri berbasis kelautan, pariwisata, serta perdagangan, transportasi, dan logistik. Dua poin SDGs yang dapat dicapai oleh dimensi ekonomi dalam aspek ekonomi biru adalah SDGs 8 (*Decent Work and Economic Growth*) dan SDGs 9 (*Industry, Innovation, and Infrastructure*).

Ketiga, dimensi sosial mencerminkan aspek inklusivitas dan keadilan sosial dalam pemanfaatan sumber daya laut. Tiga sektor utama yang dapat menjelaskan dimensi tersebut antara lain kesejahteraan, kesehatan, dan pendidikan. Dimensi sosial dalam aspek ekonomi biru dapat mendukung tercapainya lima poin SDGs, yakni SDGs 1 (*No Poverty*), SDGs 2 (*Zero Hunger*), SDGs 3 (*Good Health and Well-being*), SDGs 4 (*Quality Education*), dan SDGs 10 (*Reduced Inequalities*) (Kementerian PPN, 2023a).



Gambar 2.2 Dimensi Penyusun Indonesian Blue Economy Index (ARISE+ Indonesia)

Sumber: (Kementerian PPN, 2023a)

Setelah mengetahui dimensi penyusunnya seperti yang terlihat pada Gambar 2.2, penyusunan indeks ekonomi biru dihitung dengan mengadopsi konsep penyusunan Indeks Pembangunan Manusia (IPM) yang dikeluarkan oleh *United Nations Development Programme* (UNDP). Proses penyusunan indeks tersebut meliputi (Dasic et al., 2020; Harya, 2020):

1. Normalisasi *min-max* untuk menstandarisasi data
2. Metode rata-rata aritmatik untuk menghitung indeks dimensi penyusun IEB
3. Metode rata-rata geometrik untuk menghitung IEB

Secara matematis dapat dijelaskan sebagai berikut.

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

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

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

di mana:

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

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

$\max(x_{ij})$: nilai maksimum data variabel ke i dimensi ke j

$\min(x_{ij})$: nilai minimum data variabel ke i dimensi ke j

I_j : indeks dimensi ke j

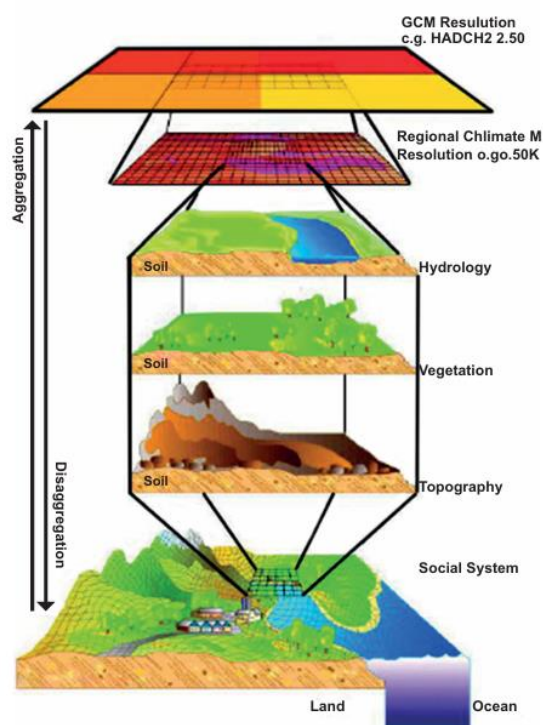
IEB : indeks ekonomi biru

2.1.4 Statistical Downscaling

Statistical Downscaling (SD) merupakan teknik statistik yang digunakan dalam pemodelan iklim untuk menghubungkan data iklim global dengan kondisi lokal sehingga prediksi perubahan iklim dengan membangun hubungan antara data observasi dengan faktor meteorologi regional mungkin dilakukan (Su et al., 2020). Metode tersebut membantu mengurangi kesenjangan antara skala global dan regional dengan

menyesuaikan sifat temporal, spasial, dan multivariat dalam data iklim (Robin & Vrac, 2021). SD melibatkan perhitungan yang relatif kecil, perumusan model yang sederhana, dan pengoreksian kesalahan sistematis dari *Global Climate Model* (GCM) yang merupakan data skala global yang biasa digunakan sebagai penjelas dalam merepresentasikan berbagai macam sistem yang terdapat di bumi, seperti atmosfer, lautan, permukaan daratan, dan es laut yang sangat berguna untuk penelitian perubahan dan variabilitas iklim (Djuraidah et al., 2021).

SD menggunakan data GCM sebagai dasar dalam peramalan iklim lokal. Pemilihan variabel yang digunakan sebagai prediktor serta penentuan domain (lokasi dan jumlah grid) diperlukan untuk mendapatkan peramalan yang stabil. Metode SD mengasumsikan kondisi fisiografik regional dan kondisi iklim global dengan resolusi rendah mampu mengendalikan iklim regional dengan resolusi tinggi. Secara lebih jelas, konsep *statistical downscaling* dapat dilihat pada ilustrasi berikut.



Gambar 2.3 Konsep Statistical Downscaling

Sumber: (Rachmawati et al., 2019)

Berdasarkan ilustrasi Gambar 2.3, berikut adalah bentuk umum dari model *statistical downscaling* (Santri & Hanike, 2020).

$$y_t = f(X_{t \times g}) \quad (2.4)$$

dengan :

y_t : variabel iklim lokal

$X_{t \times g}$: variabel luaran GCM

t : banyaknya waktu

g : banyaknya grid domain GCM

Penentuan jumlah grid GCM yang akan digunakan sebagai input model merupakan langkah penting yang perlu dilakukan dalam proses peramalan SD. Namun, data GCM biasanya memiliki karakteristik yang kompleks, yaitu berdimensi tinggi (banyak variabel) dan multikolinearitas (antar variabel saling berkorelasi). Teknik statistik yang efektif untuk mengatasi masalah tersebut adalah reduksi dimensi dengan *principal component analysis* (PCA) (Carreau & Guinot, 2021; Rachmawati et al., 2019).

2.1.5 *Principal Component Analysis*

Karl Pearson pada tahun 1901 menemukan sebuah teknik untuk reduksi dimensi yang diberi nama *Principal Component Analysis* (PCA). PCA atau yang biasa disebut dengan analisis komponen utama merupakan sebuah teknik statistik nonparametrik dengan menggunakan konsep transformasi orthogonal untuk membentuk suatu kombinasi linear. Melalui kombinasi linear tersebut, PCA mengubah sekelompok variabel yang berkorelasi menjadi sekelompok variabel yang tidak berkorelasi dengan nama faktor. Akibatnya, PCA mampu meringkas informasi penting dari data berdimensi tinggi agar memiliki representasi yang lebih sederhana tanpa kehilangan informasi yang signifikan (Reddy et al., 2020).

Proses reduksi dimensi dengan PCA dapat dijelaskan sebagai berikut. Asumsikan bahwa dataset $x_1, x_2, \dots, x_n$ memiliki n dimensi. Kemudian, data berdimensi n tersebut akan direduksi menjadi dimensi k

($k \ll n$) dengan menggunakan PCA. Langkah-langkah tersebut antara lain (Saputri et al., 2023; Susanti et al., 2023):

1. Standarisasi data dengan Z-score

Standarisasi data dilakukan dengan mengubah semua variabel dataset ke dalam skala yang sama untuk menghindari adanya kesalahan interpretasi dikarenakan dataset yang digunakan sering kali memiliki skala satuan yang berbeda. Proses tersebut melibatkan pengurangan setiap titik data (x_i) dengan rata-rata data ($\bar{x}_i$) dan kemudian membagi hasilnya dengan standar deviasi data (σ_i). Secara matematis, proses tersebut dapat dituliskan sebagai berikut.

$$x'_i = \frac{x_i - \bar{x}_i}{\sigma_i} \quad (2.5)$$

2. Menghitung matriks kovariansi

Setelah data distandarisasi, langkah selanjutnya adalah menganalisis hubungan antar variabel dalam dataset melalui matriks kovariansi. Matriks kovariansi digunakan sebagai alat untuk mengukur tingkat dan arah hubungan linear setiap pasang variabel dalam dataset yang dapat dihitung dengan formula berikut.

$$\text{cov}(x', y') = \frac{1}{n-1} \sum_{i=1}^n (x'_i - \bar{x}') (y'_i - \bar{y}') \quad (2.6)$$

$$A = \begin{bmatrix} \text{cov}(x', x') & \text{cov}(x', y') \\ \text{cov}(y', x') & \text{cov}(y', y') \end{bmatrix} \quad (2.7)$$

3. Menghitung nilai eigen dan vektor eigen

Jika A merupakan matriks kovariansi dengan ordo $n \times n$, maka skalar λ dan vektor-vektor $x_{n \times 1} \neq 0$ disebut nilai eigen dan vektor eigen ketika memenuhi persamaan:

$$A_{n \times n} x = \lambda x \quad (2.8)$$

Sekumpulan nilai eigen $\lambda_1, \lambda_2, \dots, \lambda_n$ dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ jika dan hanya jika memenuhi persamaan:

$$\det(A_{n \times n} - \lambda I_n) = 0 \quad (2.9)$$

Setelah diperoleh sekumpulan nilai eigen, vektor eigen x dari setiap nilai eigen dapat diperoleh melalui persamaan:

$$(A_{n \times n} - \lambda I_n)x = 0 \quad (2.10)$$

4. Menentukan faktor

Terdapat tiga kriteria umum yang dapat digunakan sebagai pedoman dalam menentukan jumlah faktor yang optimal untuk digunakan pada analisis lebih lanjut. Tiga kriteria tersebut antara lain:

a. Nilai eigen

Kaiser Criterion menyarankan bahwa faktor-faktor yang memiliki nilai eigen lebih besar dari satu harus dipertahankan untuk analisis lebih lanjut. Nilai eigen suatu faktor menunjukkan kontribusi faktor tersebut terhadap proporsi varians total. Faktor yang memiliki nilai eigen lebih dari satu menunjukkan bahwa faktor tersebut dapat menjelaskan lebih banyak varians dalam data dibandingkan dengan variabel asli. Alhasil, faktor yang harus dipertahankan untuk digunakan pada analisis lanjutan adalah faktor yang memiliki nilai eigen lebih dari satu.

b. Scree plot

Scree plot merupakan representasi grafis yang menunjukkan nilai eigen setiap faktor, diurutkan dari yang terbesar ke yang terkecil, dengan sumbu vertikal dan horizontal secara berturut-turut mewakili nilai eigen dan nomor faktor. Titik-titik pada *scree plot* dihubungkan dengan garis untuk menunjukkan tren penurunan nilai eigen. Titik *elbow* atau titik siku pada *scree plot* digunakan untuk menentukan jumlah faktor yang optimal di mana titik tersebut menunjukkan adanya penurunan nilai eigen menjadi lebih landai setelah penurunan yang tajam di awal. Faktor yang berada sebelum titik tersebut menunjukkan bahwa faktor tersebut dapat menjelaskan sebagian besar varians dalam

data sehingga layak untuk dipertahankan dalam analisis lebih lanjut.

c. Proporsi kumulatif varians

Proporsi kumulatif varians, yang dihitung dengan menjumlahkan varians yang dijelaskan oleh setiap faktor hingga titik tertentu, menunjukkan seberapa besar varians total dalam data yang dapat dijelaskan oleh sejumlah faktor tersebut. Proporsi kumulatif varians yang tinggi menunjukkan bahwa faktor-faktor yang dipilih mampu mewakili sebagian besar informasi dalam data asli. Secara umum, proporsi kumulatif varians berkisar antara 70% dan 90% dianggap cukup untuk menangkap sebagian besar informasi penting dalam data asli sehingga reduksi dimensi dapat dilakukan tanpa kehilangan informasi yang signifikan. Alhasil, analisis lanjutan dan interpretasi hasil dapat lebih mudah.

5. Transformasi data

Transformasi data dalam PCA bertujuan untuk mempertahankan informasi penting dari data asli dengan membentuk representasi baru yang memiliki dimensi lebih rendah. Proses transformasi melibatkan pembentukan kombinasi linear dari variabel asli yang dikenal sebagai komponen utama atau faktor. Kombinasi linear tersebut diperoleh dengan mengalikan matriks data hasil standarisasi yang diasumsikan dalam matriks D dengan matriks vektor eigen X dan kemudian disimpan dalam matriks F yang merupakan hasil reduksi dimensi dengan PCA. Proses perkalian matriks ini dapat dituliskan sebagai berikut.

$$F = DX \quad (2.11)$$

Data transformasi tersebut yang akan digunakan untuk analisis lebih lanjut karena data tersebut telah direpresentasikan dalam ruang dimensi yang lebih rendah sehingga memudahkan analisis dan interpretasi.

2.1.6 Regresi Spatio-Temporal

Analisis data spasial dan temporal yang melibatkan informasi geografis dan waktu memerlukan pertimbangan khusus terhadap autokorelasi serta heterogenitas spasial dan temporal. Regresi spatio-temporal yang merupakan pengembangan dari regresi konvensional dirancang dengan tujuan mendapatkan kesimpulan yang akurat dan dapat diandalkan. Pengabaian faktor autokorelasi serta heterogenitas spasial dan temporal dalam regresi konvensional yang disebabkan oleh asumsi independensi data yang tidak terpenuhi mengakibatkan bias pada kesimpulan yang dihasilkan. Model regresi spatio-temporal yang berbeda-beda telah banyak dikembangkan untuk mengatasi hal tersebut, seperti model yang melibatkan model Besag-York-Mollie 2 dan *Random Walk 2*.

Model spatio-temporal telah banyak diaplikasikan dalam berbagai penelitian, terutama untuk mengidentifikasi pola spasial dan temporal dari suatu kasus. Terdapat dua pendekatan komponen temporal yang dapat digunakan dalam membangun model ini yakni parametrik dan nonparametrik. Struktur parametrik menggunakan persamaan matematika tertentu untuk menggambarkan perubahan kasus seiring dengan berjalannya waktu. Sementara itu, struktur nonparametrik memiliki fleksibilitas yang lebih besar karena struktur tersebut menggunakan data untuk menentukan pola perubahan dan tidak mengasumsikan bentuk tertentu.

2.1.6.1 Model Besag-York-Mollie 2

Model Besag-York-Mollie 2 (BYM2) merupakan pengembangan dari model BYM yang dirancang untuk mengatasi dua keterbatasan utama dari model sebelumnya. Pertama, model BYM tidak dapat mengidentifikasi secara jelas perbedaan pengaruh antara efek acak terstruktur dan tidak terstruktur. Kedua, hiperparameter presisi yang mengontrol tingkat variabilitas efek acak tidak memiliki skala yang sama sehingga sulit untuk membandingkan tingkat variabilitas dua komponen tersebut (Simkin et al.,

2022). Model BYM2 mengatasi keterbatasan tersebut dengan memodifikasi kombinasi efek acak spasial $u_i + v_i$ dari model BYM menjadi bentuk berikut.

$$\frac{1}{\sqrt{\tau_b}}(\sqrt{1-\phi}v_i + \sqrt{\phi}u_i) \quad (2.12)$$

dengan :

- u_i : efek spasial terstruktur
- v_i : efek spasial tidak terstruktur
- τ_b : parameter presisi yang menjelaskan kontribusi varians dari penjumlahan efek acak u_i dan v_i
- $1 - \phi$: parameter pencampuran yang mengontrol kontribusi varians dari efek acak spasial tidak terstruktur v_i
- ϕ : parameter pencampuran yang mengontrol kontribusi varians dari efek acak spasial terstruktur u_i

Parameter τ_b dan ϕ dalam formula tersebut ditentukan menggunakan prior *penalized complexity* (PC). Prior tersebut terbukti dapat menjadi pilihan yang tepat dalam model spasial Bayesian BYM2 karena dapat mendukung model yang tidak terlalu kompleks dan memungkinkan interpretasi parameter yang lebih jelas. Akibatnya, model BYM2 mampu membedakan dampak ketergantungan spasial dan efek variabilitas atau overdispersi yang ada dalam data sehingga memungkinkan pemahaman yang lebih jelas bagaimana faktor spasial dan faktor variabilitas lainnya berkontribusi pada estimasi parameter model (Morales-Otero & Núñez-Antón, 2021; Morris et al., 2019).

2.1.6.2 Model *Random Walk 2*

Semua parameter yang belum diketahui dalam model *age-period-cohort* (APC) dianggap sebagai variabel acak yang memiliki distribusi prior dalam inferensi Bayesian. Asumsi bahwa efek yang berdekatan dalam waktu cenderung serupa mendorong penggunaan *smoothing prior* yang bertujuan menghasilkan estimasi dan prediksi yang lebih halus dan realistis. Berdasarkan fleksibilitasnya dibandingkan dengan berbagai model

pemulusan temporal, model *random walk* 2 (RW2) telah menjadi pilihan standar dan sangat populer untuk digunakan (Du et al., 2020).

Model RW2 merupakan generalisasi dari model RW1 di mana model RW1 hanya dapat memodelkan data yang memiliki tren konstan. Sementara itu, model RW2 dapat mengakomodasi data yang memiliki tren linear. Oleh karena itu, model RW2 lebih fleksibel untuk memodelkan data yang menunjukkan perubahan gradual dan konsisten dari waktu ke waktu. Meskipun model RW2 lebih fleksibel untuk tren waktu linear, perubahan fluktuatif yang tidak sesuai dengan asumsi tersebut tidak diabaikan tetapi diberikan probabilitas yang lebih rendah. Akibatnya, model yang diperoleh cenderung mengikuti tren linear yang lebih halus (Riebler & Held, 2017).

Model RW2 mengasumsikan bahwa perbedaan kedua dari semua efek waktu mengikuti distribusi normal dengan mean bernilai nol dan independen. Perbedaan kedua dalam model tersebut mencerminkan ketergantungan setiap nilai pada tetangga terdekatnya dan tetangga terdekat tetangga tersebut. Jumlah tetangga yang lebih banyak tersebut mengakibatkan pola temporal yang lebih halus karena model cenderung menghasilkan efek waktu yang mengikuti tren linear lebih mulus (Blangiardo et al., 2020; Li et al., 2019; Nazia et al., 2022). Secara matematis, model RW2 dapat dituliskan sebagai berikut.

$$\gamma_{it} | \gamma_{i,t-1}, \gamma_{i,t-2} \sim \text{Normal}(2\gamma_{i,t-1} + \gamma_{i,t-2}, \tau_\gamma^{-1}) \quad (2.13)$$

2.1.6.3 Regresi Spatio-Temporal Parametrik

Bernardinelli dkk. (1995) menunjukkan tren parametrik dari komponen temporal dengan bentuk prediktor linearnya dapat dituliskan sebagai berikut (Blangiardo & Cameletti, 2015).

$$\eta_{it} = \alpha + \sum_{m=1}^M b_m PC_m + \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} v_i + \sqrt{\phi} u_i) + (\beta + \delta_i) \times t \quad (2.14)$$

di mana :

η_{it}	: <i>link function</i> (fungsi penghubung) variabel respons
α	: konstanta
$b_1, b_2, \dots$	: koefisien variabel prediktor
$PC_1, PC_2, \dots$	: variabel prediktor dari hasil reduksi dimensi dengan PCA
$\frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi}v_i + \sqrt{\phi}u_i)$	: komponen spasial dengan model BYM2, u_i efek acak spasial terstruktur, v_i efek acak spasial tidak terstruktur
β	: tren temporal global
δ_i	: tren diferensial area i diasumsikan $Normal(0, \sigma_\delta^2)$, menjelaskan perbedaan antara tren temporal global dengan tren temporal di area i
t	: variabel waktu

2.1.6.4 Regresi Spatio-Temporal Nonparametrik

Meskipun model regresi spatio-temporal parametrik telah menawarkan kerangka kerja yang baik untuk menganalisis pola spasial dan temporal, model ini memiliki keterbatasan linearitas pada tren temporal diferensial δ_i . Namun, asumsi tersebut mungkin tidak selalu sesuai dengan kenyataan. Hal tersebut terjadi dalam kasus di mana perubahan temporalnya memiliki pola yang kompleks atau fluktuatif. Keterbatasan tersebut dapat diatasi menggunakan regresi spatio-temporal nonparametrik. Formula prediktor linear regresi spatio-temporal nonparametrik dapat dituliskan sebagai berikut.

$$\eta_{it} = \alpha + \sum_{m=1}^M b_m PC_m + \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi}v_i + \sqrt{\phi}u_i) + \gamma_t + \phi_t \quad (2.15)$$

Model ini mempertahankan struktur dasar model parametrik yang menggunakan spesifikasi spasial BYM2, namun efek tren temporal global dan diferensial diganti dengan dua istilah baru. Istilah-istilah tersebut adalah γ_t sebagai efek acak temporal terstruktur dan ϕ_t sebagai efek acak temporal tidak terstruktur. Efek acak temporal terstruktur dimodelkan menggunakan model RW2, sedangkan efek acak temporal tidak terstruktur diasumsikan berdistribusi $Normal\left(0, \frac{1}{\tau_\phi}\right)$ (Rachmawati et al., 2018).

2.1.6.5 Regresi Spatio-Temporal Nonparametrik dengan Interaksi

Meskipun model regresi spatio-temporal nonparametrik lebih mampu menangkap pola perubahan temporal yang kompleks, model tersebut belum mempertimbangkan potensi interaksi spatio-temporal. Keterbatasan tersebut dapat diatasi melalui adanya model regresi spatio-temporal nonparametrik yang telah dikembangkan dengan mengintegrasikan komponen interaksi. Pengembangan ini memungkinkan untuk mengidentifikasi pola interaksi antara spasial dan temporal yang terjadi sehingga pemahaman yang komprehensif tentang bagaimana variabel respons berubah dalam ruang dan waktu dapat diperoleh. Model regresi spatio-temporal nonparametrik dengan interaksi memiliki bentuk prediktor linear sebagai berikut (Rachmawati & Puspongoro, 2018).

$$\eta_{it} = \alpha + \sum_{m=1}^M b_m PC_m + \frac{1}{\sqrt{\tau_b}} (\sqrt{1-\phi} v_i + \sqrt{\phi} u_i) + \gamma_t + \phi_t + \delta_{it} \quad (2.16)$$

Lima elemen pertama dari komponen regresi pada persamaan (2.16) memiliki spesifikasi yang sama dengan model nonparametrik. Perbedaan utamanya terletak pada penambahan komponen δ_{it} yang secara eksplisit menjelaskan interaksi antara efek acak spasial dan temporal. Interaksi antara spasial dan temporal terbagi menjadi empat jenis seperti yang terdapat pada Tabel 2.1 (Sani et al., 2023).

Tabel 2.1 Tipe Interaksi Spatio-Temporal

Tipe	Efek Spasial	Efek Temporal	Matriks Korelasi S_δ
1	Tidak Terstruktur	Tidak Terstruktur	$S_v \otimes S_\phi = I \otimes I = I$
2	Tidak Terstruktur	Terstruktur	$S_v \otimes S_\gamma$
3	Terstruktur	Tidak Terstruktur	$S_u \otimes S_\phi$
4	Terstruktur	Terstruktur	$S_u \otimes S_\gamma$

Sumber: (Sani et al., 2023)

2.1.7 Inferensi Bayesian

Paradigma Bayesian menggunakan asumsi bahwa semua parameter yang tidak diketahui diasumsikan sebagai variabel acak untuk tujuan mengestimasi distribusi posterior gabungan. Distribusi tersebut seringkali rumit, terutama dalam model multivariat yang melibatkan banyak parameter yang saling terkait. Kompleksitas tersebut dapat diatasi dengan mengestimasi distribusi posterior gabungan menggunakan konsep teorema Bayes:

$$\pi(\theta|y) = \frac{\pi(y|\theta)\pi(\theta)}{\pi(y)} \quad (2.17)$$

Namun, proses ini dipersulit oleh keberadaan kemungkinan marjinal $\pi(y)$. Kemungkinan marjinal merupakan konstanta normalisasi yang mewakili probabilitas mengamati data tanpa mempertimbangkan parameter θ . Penghitungan kemungkinan marjinal dalam model multivariat menjadi rumit karena melibatkan integrasi atas seluruh ruang parameter yang dapat dituliskan sebagai berikut.

$$\int_{\Theta} \pi(y|\theta)\pi(\theta)d\theta \quad (2.18)$$

Oleh karena itu, distribusi posterior $\pi(\theta|y)$ seringkali disederhanakan dengan menskalakan ulang hasil kali antara distribusi likelihood dan distribusi prior yang dapat dijelaskan dalam formula berikut.

$$\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta) \quad (2.19)$$

Distribusi likelihood $\pi(\theta|y)$ mengukur seberapa baik model dengan parameter θ tertentu menjelaskan data y yang diamati. Sementara itu, distribusi prior $\pi(\theta)$ mengukur keyakinan awal tentang parameter θ sebelum melihat data y . Melalui integrasi informasi dari likelihood dan prior, distribusi posterior dapat memperbarui keyakinan tentang parameter θ setelah melihat data y .

Ketika distribusi posterior gabungan untuk semua parameter telah diketahui, distribusi posterior marginal untuk setiap parameter juga dapat diperoleh. Distribusi posterior marginal tersebut dihitung dengan mengintegrasikan distribusi posterior gabungan atas semua parameter lainnya yang dapat dinyatakan sebagai berikut.

$$\pi(\theta_i|y) = \int \pi(\theta|y) d\theta_{-i} \quad (2.20)$$

Semua parameter dalam model, kecuali parameter ke- i (θ_i) digambarkan oleh θ_{-i} . Namun, penghitungan integral tersebut secara analitis seringkali sulit dalam praktiknya, terutama untuk model yang kompleks. Oleh karena itu, teknik aproksimasi seperti aproksimasi *Laplace* diperlukan untuk mengestimasi distribusi posterior (Gómez-Rubio, 2020).

2.1.8 Algoritma *Integrated Nested Laplace Approximation*

Asumsikan bahwa y mewakili beberapa data, ψ mewakili vektor dari semua variabel berdistribusi Gaussian, dan θ merupakan vektor hiperparameter yang belum tentu berdistribusi Gaussian. Fokus utama dari inferensi Bayesian adalah distribusi posterior marginal. Distribusi posterior marginal dapat dituliskan sebagai berikut.

$$\pi(\psi_k|y) = \int_{\theta} \int_{\psi_k} \pi(\psi, \theta|y) d\psi_{-k} d\theta \quad (2.21)$$

dengan ψ_k mewakili ψ selain komponen ke- k .

Distribusi posterior dari ψ_k dan θ_i secara berturut-turut dapat dituliskan sebagai :

$$\pi(\psi_k|y) = \int \pi(\psi_k|\theta, y) \pi(\theta|y) d\theta \quad (2.22)$$

$$\pi(\theta_i|y) = \int \pi(\theta|y) d\theta_{-i} \quad (2.23)$$

Kemudian distribusi posterior marginal untuk $\pi(\theta|y)$ dapat diaproksimasi dengan cara berikut.

$$\begin{aligned} \pi(\theta|y) &= \frac{\pi(\psi, \theta|y)}{\pi(\psi|\theta, y)} \propto \frac{\pi(\theta)\pi(\psi|\theta)\pi(y|\psi)}{\pi(\psi|\theta, y)} \\ &\approx \frac{\pi(\theta)\pi(\psi|\theta)\pi(y|\psi)}{\tilde{\pi}(\psi|\theta, y)} \Big|_{\psi=\psi^*(\theta)} =: \tilde{\pi}(\theta|y) \end{aligned} \quad (2.24)$$

dengan $\tilde{\pi}(\psi|\theta, y)$ merupakan aproksimasi Gaussian dari $\pi(\psi|\theta, y)$ dan $\psi^*(\theta)$ adalah titik puncaknya.

Umumnya, aproksimasi memiliki sifat yang tidak terlalu akurat. Oleh karena itu, vektor parameter ψ diuraikan menjadi $\psi = (\psi_k, \psi_{-k})$ dan aproksimasi *Laplace* diterapkan lagi sebagai berikut.

$$\begin{aligned} \pi(\psi_k|\theta, y) &= \frac{\pi(\psi_k, \psi_{-k}|\theta, y)}{\pi(\psi_{-k}|\psi_k, \theta, y)} \\ &\approx \frac{\pi(\psi, \theta|y)}{\tilde{\pi}(\psi_{-k}|\psi_k, \theta, y)} \Big|_{\psi_{-k}=\psi_{-k}^*(\psi_k, \theta)} =: \tilde{\pi}(\psi_k|\theta, y) \end{aligned} \quad (2.25)$$

Persamaan (2.25) menjelaskan bahwa aproksimasi Gaussian *Laplace* dari $\pi(\psi_k|\theta, y)$ adalah $\tilde{\pi}(\psi_k|\theta, y)$ dan titik puncak dari parameter tersebut adalah $\psi_{-k}^*(\psi_k, \theta)$. Alhasil, distribusi posterior marginalnya dapat diperoleh dengan cara berikut.

$$\tilde{\pi}(\psi_k|y) \approx \int \tilde{\pi}(\psi_k|\theta, y) \tilde{\pi}(\theta|y) d\theta \quad (2.26)$$

Persamaan (2.26) menggunakan ekspansi deret Taylor untuk mengaproksimasi *Laplace* secara numerik dengan tujuan mendapatkan hasil yang cukup akurat dan komputasi yang efisien (Blangiardo & Cameletti, 2015; Gómez-Rubio, 2020; Rachmawati & Puspongoro, 2018;

Sani et al., 2023). Secara matematis, hal tersebut dapat dituliskan sebagai berikut.

$$\tilde{\pi}(\psi_k|y) \approx \sum_j \tilde{\pi}(\psi_k|\theta^{(j)}, y) \tilde{\pi}(\theta^{(j)}|y) \Delta_j \quad (2.27)$$

Persamaan (2.27) menyatakan bahwa beberapa titik $\{\theta^{(j)}\}$ berkorespondensi dengan $\{\Delta_j\}$.

Secara operasional, persamaan (2.27) dapat diperoleh dengan langkah-langkah berikut.

- a. Mengeksplorasi secara nonparametrik distribusi posterior gabungan hyperparameter $\tilde{\pi}(\theta|y)$ untuk mengetahui titik-titik yang akan digunakan dalam integrasi numerik. Eksplorasi tersebut dilakukan dengan melakukan reparameterisasi ruang θ melalui langkah-langkah berikut.
 - 1) Mencari nilai θ yang paling mungkin terjadi dengan memaksimalkan logaritma dari distribusi posterior $\tilde{\pi}(\theta|y)$ menggunakan metode Newton-Raphson.
 - 2) Menghitung matriks Hessian negatif H pada titik puncak θ^* untuk mengetahui bentuk distribusi posterior di sekitar titik puncak tersebut.
 - 3) Menghitung dekomposisi eigen $\Sigma = V\Lambda^{1/2}V'$ dengan Σ merupakan invers dari matriks Hessian negatif untuk mengetahui bentuk dan skala distribusi posterior di sekitar titik puncak θ^* .
 - 4) Mendefinisikan variabel baru z yang telah distandarisasi dan saling orthogonal sehingga $\theta(z) = \theta^* + V\Lambda^{\frac{1}{2}}z$
 Variabel baru z tersebut dilakukan untuk menyederhanakan proses eksplorasi ruang parameter θ dan menghitung distribusi posterior marjinal dengan lebih efisien.

- b. Menghitung posterior bersyarat $\tilde{\pi}(\psi_k|\theta^{(j)}, y)$ untuk setiap titik $\theta^{(j)}$ dan kemudian melakukan integrasi numerik untuk mendapatkan posterior marginal $\tilde{\pi}(\psi_k|y)$.

2.1.9 Watanabe-Akaike Information Criterion

Watanabe-Akaike *Information Criterion* (WAIC) merupakan kriteria informasi prediksi yang menggeneralisasikan AIC untuk model Bayesian. WAIC berfungsi sebagai alternatif dari DIC dalam pemilihan model. Meskipun WAIC dan DIC dalam praktiknya sering memberikan hasil yang serupa, WAIC memiliki landasan teoritis yang lebih kuat dan akurat dalam menangani model dengan distribusi posterior yang kompleks seperti model hierarkis Bayesian, model spatio-temporal, dan model yang menggunakan data tidak biasa. Secara matematis, WAIC dapat dihitung dengan menggunakan formula berikut.

$$WAIC = -2 \sum_i E[\log(\pi(y|\theta))] + 2 \sum_i Var[\log(\pi(y|\theta))] \quad (2.28)$$

Komponen pertama dan komponen kedua dari persamaan tersebut secara berturut-turut merupakan rata-rata log-likelihood posterior dan varians log-likelihood posterior. Dengan demikian, model dengan nilai WAIC yang lebih rendah merupakan hasil yang diharapkan karena menunjukkan kinerja model yang lebih baik (Aisah et al., 2020; Evans, 2019).

2.2 Hasil Penelitian Terdahulu

Penelitian ini berfokus pada dampak perubahan iklim terhadap potensi *blue economy*. Pemahaman mengenai konteks dan relevansi penelitian ini diperoleh penulis dengan melakukan studi literatur terhadap penelitian-penelitian terdahulu. Tujuannya adalah untuk mengidentifikasi celah penelitian dalam pemodelan SD dengan regresi spatio-temporal menggunakan inferensi INLA di mana hal tersebut akan menjadi dasar untuk pengembangan penelitian ini. Beberapa penelitian terdahulu yang relevan dengan topik ini akan disajikan pada Tabel 2.2.

Tabel 2.2 Hasil Penelitian Terdahulu

No	Penulis (Tahun)	Judul Penelitian	Metode dan Pemodelan	Hasil
1	Mutiara Aghnyn et al. (2024)	Mapping Blue Economy Potential Using Spatial Statistical Downscaling Model: Analysis of the Impact of Climate Change on Freshwater Fish Resources	Pemodelan regresi spasial dengan model efek acak spasial berupa BYM2, inferensi INLA	Berdasarkan pemodelan SD, curah hujan menyebabkan penurunan potensi <i>blue economy</i> di Kabupaten Banyuwangi. Model SD memiliki kemampuan untuk mendekati data aktual dan memiliki potensi besar dalam aplikasi praktis, terutama ketika data lokal terbatas atau tidak lengkap.

2	Marwa Elsherif (2024)	Nexus between Blue Economy, Renewable Energy and Environmental Sustainability in the MENA Region: Evidence from Panel Threshold Regression	Model <i>Stochastic Regression on Population, Affluence and Technology</i> (STIRPAT) dengan berbagai teknik ekonometrik (analisis korelasi, uji multikolinearitas, uji ketergantungan <i>cross-sectional</i> , uji kointegrasi, kuantil melalui momen, regresi ambang batas, dan uji kausalitas Granger panel)	Investasi dalam energi terbarukan dan ekonomi biru menunjukkan adanya potensi besar untuk meningkatkan keberlanjutan lingkungan. Meskipun keduanya memberikan manfaat yang positif, dampaknya terhadap lingkungan masih bervariasi dan lebih dapat dirasakan oleh negara-negara dengan tingkat ekonomi yang lebih tinggi.
---	-----------------------	--	--	---

3	Luis A. Barboza et al. (2023)	Spatio-temporal downscaling emulator for regional climate models	Spatio-temporal downscaling emulator dengan INLA dan pendekatan Dambon, <i>fitting</i> model dengan <i>Stochastic Partial Differential Equation</i> (SPDE)	Metode downscaling lebih akurat dalam memprediksi suhu dengan resolusi yang lebih tinggi meskipun membutuhkan waktu yang lebih lama. Emulator berhasil memprediksi pola dan tren suhu secara umum tetapi kesulitan dalam menangkap suhu yang lebih tinggi selama bulan-bulan musim panas. Selain itu, emulator terkadang memprediksi suhu yang lebih tinggi daripada yang diamati ketika suhu sebenarnya lebih dingin.
4	Anik Djuraidah et al. (2021)	Extreme data analysis using spatio-temporal bayes regression with INLA in statistical downscaling model	Pemodelan SD regresi spatio-temporal Bayes menggunakan inferensi INLA dengan tiga distribusi berbeda, yaitu Gamma, Bernoulli, dan	Curah hujan untuk lokasi yang tidak teramati dapat diestimasi menggunakan model Bayes dengan meminjam karakteristik dari lokasi yang teramati. Efek acak temporal menunjukkan efek yang signifikan dalam memperkirakan curah hujan ekstrem. Distribusi Generalized Pareto merupakan distribusi yang tepat dalam menggambarkan karakteristik curah hujan di wilayah Jawa Barat.

			<i>generalized Pareto</i> yang melibatkan <i>spatial smoothing</i> <i>parameters</i> , <i>threshold</i> <i>assesment</i> , dan <i>tail</i> <i>index parameter</i>	
--	--	--	---	--

Sumber: diolah oleh peneliti



Seperti yang telah ditunjukkan pada Tabel 2.2, dapat disimpulkan bahwa penelitian sebelumnya tentang ekonomi biru dengan pendekatan kuantitatif menggunakan teknik pemodelan spatio-temporal SD belum banyak dilakukan. Mutiara Aghnyn et al. (2024) melakukan pemodelan SD menggunakan regresi spasial dengan INLA untuk mengetahui pengaruh iklim terhadap potensi ekonomi biru di Kabupaten Banyuwangi. Selain itu, penelitian ini menggunakan variabel jumlah penangkapan ikan untuk menjelaskan potensi ekonomi biru di Kabupaten Banyuwangi. Temuan dari penelitian tersebut mengindikasikan bahwa peningkatan curah hujan berdampak pada berkurangnya hasil tangkapan ikan yang pada akhirnya dapat mengancam keberlanjutan potensi ekonomi biru di Kabupaten Banyuwangi. Berdasarkan hasil penelitian, dapat diketahui bahwa penelitian tersebut terbatas pada wilayah Kabupaten Banyuwangi, belum mempertimbangkan aspek temporal, dan hanya menggunakan jumlah penangkapan ikan sebagai indikator potensi ekonomi biru. Sebagai tindak lanjut atas hal tersebut, penelitian ini secara spesifik membentuk indeks ekonomi biru (IEB) pada tingkat kabupaten dan kota di Pulau Sumatera dengan cakupan analisis yang dibatasi pada wilayah tersebut. Selain itu, pemodelan yang digunakan dalam penelitian ini adalah pemodelan SD pada regresi spatio-temporal Bayesian untuk mengidentifikasi pengaruh iklim global terhadap IEB di Pulau Sumatera pada periode 2019 hingga 2022.

Marwa Elsherif (2024) menyatakan bahwa investasi dalam energi terbarukan dan ekonomi biru memiliki potensi besar untuk meningkatkan keberlanjutan lingkungan berdasarkan hasil pemodelan dengan STIRPAT. Namun, hal tersebut belum dapat dirasakan oleh seluruh wilayah di area MENA secara merata. Model STIRPAT melibatkan berbagai teknik ekonometrik yang memiliki fungsi masing-masing untuk membentuk model tersebut, diantaranya adalah analisis korelasi, uji multikolinearitas, uji ketergantungan *cross-sectional*, uji kointegrasi, kuantil melalui regresi

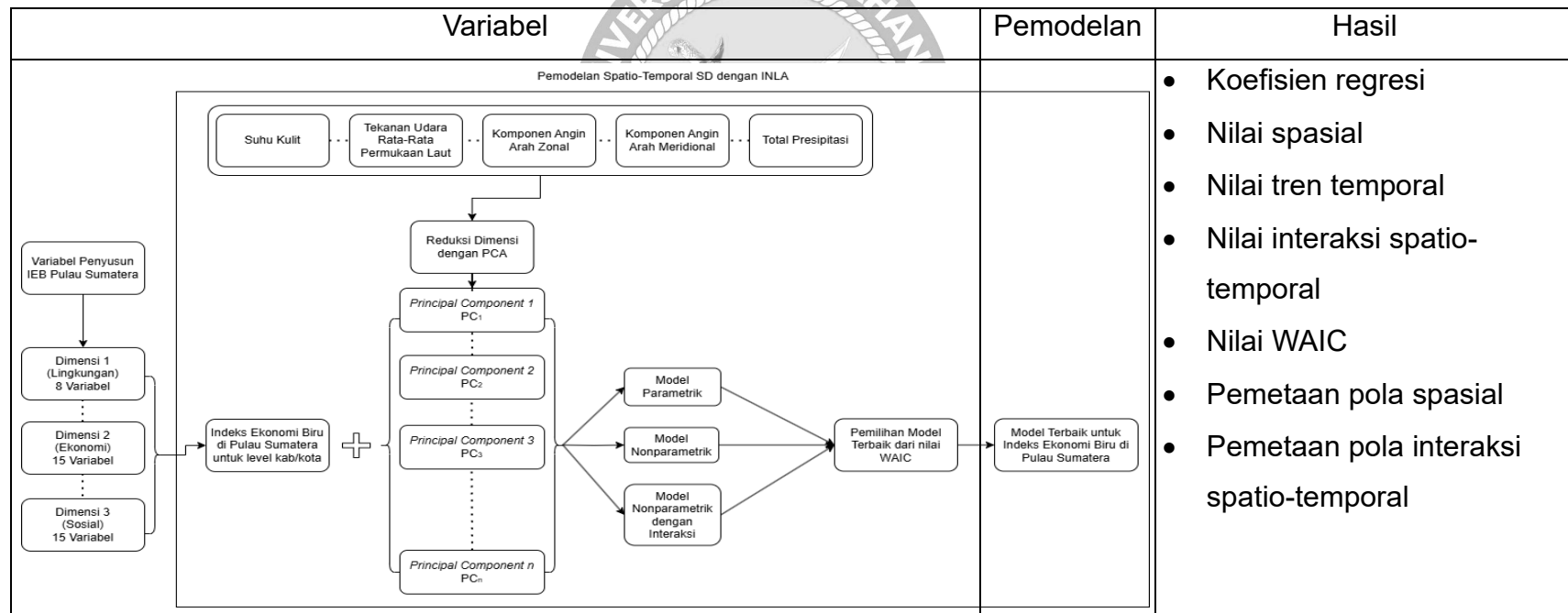
momen dan ambang batas, serta uji kausalitas Granger panel. Analisis dengan berbagai teknik ekonometrik tersebut membantu mengidentifikasi potensi hubungan antara kegiatan ekonomi biru, energi terbarukan, dan jejak ekologi dari data panel dalam model STIRPAT dari waktu ke waktu serta mengeksplorasi efek berbagai faktor pada jejak ekologi di berbagai tingkat. Penelitian tersebut terbatas pada analisis hubungan ekonometrik dalam model STIRPAT untuk memahami keterkaitan antara ekonomi biru dan lingkungan di mana ekonomi biru diposisikan sebagai variabel prediktor. Sebaliknya, penelitian ini menggunakan ekonomi biru sebagai variabel respons dan menerapkan pemodelan SD pada regresi spatio-temporal Bayesian dengan pendekatan INLA untuk menganalisis hubungan antara iklim global dengan ekonomi biru.

Penelitian yang dilakukan oleh Luis A. Barboza et al. (2023) dan Anik Djuraidah et al. (2021) memiliki beberapa kesamaan, salah satunya adalah menggunakan data dengan karakteristik berupa titik yang dikombinasikan dengan pemodelan spatio-temporal menggunakan pendekatan INLA. Meskipun memiliki kesamaan, adapun perbedaan antara dua penelitian tersebut adalah ruang lingkup wilayah geografis dan periode waktu yang digunakan sebagai fokus penelitian. Selain itu, Anik Djuraidah et al. (2021) melibatkan proses reduksi dimensi PCA untuk menangani masalah multikolinearitas dan mengurangi dimensi variabel GCM pada model SD sehingga memungkinkan prediksi skala lokal dengan tetap mempertahankan variasi data awal. Sementara itu, Luis A. Barboza et al. (2023) tidak melakukan hal tersebut. Dua penelitian tersebut belum mengembangkan penelitian di wilayah lain dengan menggunakan karakteristik data yang berbeda, misalnya data dengan karakteristik areal. Oleh karena itu, penelitian ini melakukan pemodelan SD pada regresi spatio-temporal Bayesian menggunakan data dengan karakteristik areal untuk mendapatkan representatif spasial yang lebih komprehensif mengenai potensi ekonomi biru di Pulau Sumatera.

2.3 Kerangka Berpikir

Tabel 2.3 menunjukkan kerangka berpikir yang dibentuk oleh penulis sebagai pedoman penelitian. Kerangka berpikir tersebut dirancang untuk mencapai tujuan penelitian dan menghindari penyimpangan dari topik utama penelitian.

Tabel 2.3 Kerangka Berpikir



Sumber: diolah oleh peneliti