PENERAPAN METODE *GEOGRAPHICALLY WEIGHTED PANEL REGRESSION* PADA INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2017-2022

(Skripsi)

Oleh

DETA ERVIANA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024

ABSTRACT

APPLICATION OF THE GEOGRAPHICALLY WEIGHTED PANEL REGRESSION METHOD ON THE HUMAN DEVELOPMENT INDEX IN INDONESIA FROM 2017 TO 2022

By

DETA ERVIANA

Linear regression is a statistical method to examine the interaction between a response variable as well as one or more predictor variables. In a research, a unit of observation should be studied over multiple time periods, as studying one unit in one time period is not enough. Therefore, a statistical approach called panel regression analysis was created to integrate *cross-section* data and *time series* data. But in reality, differences in conditions between locations are influenced by spatial effects that cause spatial heterogeneity. The *Geographically Weighted Regression* (GWR) method was developed to overcome the problem of spatial heterogeneity. Based on the advantages of both methods, a method that combines panel data regression and GWR was developed, namely *Geographically Weighted Panel Regression* (GWPR). The purpose of this study is to determine the factors that affect the human development index (HDI) in Indonesia in 2017-2022 and determine the best model by comparing global regression and GWPR models. The GWPR model with *adaptive bisquare* weights is the best model with the smallest AIC value and the largest R^2. Overall, all predictor variables used in this study have a significant effect on HDI at the significance level of $\alpha = 0{,}05$. The model equations and variables that have a significant effect generated in GWPR modeling are different for each province. Based on the similarity of variables affecting HDI in provinces that are located close together, 8 groups were formed.

Keywords: panel data regression, GWR, GWPR, HDI

ABSTRAK

PENERAPAN METODE *GEOGRAPHICALLY WEIGHTED PANEL REGRESSION* PADA INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2017-2022

Oleh

DETA ERVIANA

Regresi linier merupakan metode statistik untuk memeriksa hubungan antara variabel respons dan satu atau lebih variabel prediktor. Dalam sebuah penelitian, satu unit observasi harus diteliti selama beberapa periode waktu, karena mempelajari satu unit dalam satu periode waktu tidaklah cukup. Oleh karena itu, sebuah pendekatan statistik yang disebut analisis regresi panel diciptakan untuk mengintegrasikan data *cross-section* dan data *time series*. Namun pada kenyataannya, perbedaan kondisi antar lokasi dipengaruhi oleh efek spasial yang menyebabkan terjadinya heterogenitas spasial. Dikembangkanlah metode *Geographically Weighted Regression* (GWR) untuk mengatasi masalah heterogenitas spasial. Berdasarkan kelebihan kedua metode tersebut maka berkembanglah suatu metode yang menggabungkan antara regresi data panel dan GWR yaitu *Geographically Weighted Panel Regression* (GWPR). Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor yang mempengaruhi indeks pembangunan manusia (IPM) di Indonesia tahun 2017-2022 dan menentukan model terbaik dengan membandingkan model regresi global dan GWPR. Model GWPR dengan pembobot *adaptive bisquare* merupakan model terbaik dengan nilai AIC terkecil dan R^2 terbesar. Secara keseluruhan semua variabel prediktor yang digunakan dalam penelitian berpengaruh signifikan terhadap IPM pada taraf signifikansi $\alpha = 0{,}05$. Persamaan model dan variabel yang berpengaruh signifikan yang dihasilkan dalam pemodelan GWPR berbeda untuk setiap provinsi. Berdasarkan kesamaan variabel yang mempengaruhi IPM di provinsi yang letaknya berdekatan membentuk 8 kelompok.

Kata kunci: regresi data panel, GWR, GWPR, IPM

PENERAPAN METODE *GEOGRAPHICALLY WEIGHTED PANEL REGRESSION* PADA INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2017-2022

Oleh

DETA ERVIANA

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024

Judul : **PENERAPAN METODE *GEOGRAPHICALLY WEIGHTED PANEL REGRESSION* PADA INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2017-2022**

Nama Mahasiswa : **Deta Erviana**

Nomor Pokok Mahasiswa : **2057031005**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. **Komisi** Pembimbing

Prof. Drs. Mustofa Usman, M.A., Ph. D.
NIP.19570101 198403 1 020

Widiarti, S.Si., M.Si.
NIP. 198005022005012003

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si. M.Si.
NIP.19740316 620051 1 001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Prof. Drs. Mustofa Usman, M.A., Ph. D.**	......
Sekretaris	: **Widiarti, S.Si., M.Si.**	....
Penguji Bukan Pembimbing	: **Dr. Khoirin Nisa, S.Si., M.Si.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 19711001 200501 1 002

Tanggal Lulus Ujian Skripsi: **25 Maret 2024**

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan dibawah ini:

Nama Mahasiswa : DETA ERVIANA

Nomor Pokok Mahasiswa : 2057031005

Jurusan : Matematika

Judul Skripsi : PENERAPAN METODE *GEOGRAPHICALLY WEIGHTED PANEL REGRESSION* PADA INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2017-2022

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri, bukan hasil orang lain, dan semua hasil tulisan tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 25 Maret 2024

Penulis

DETA ERVIANA
NPM. 2057031005

RIWAYAT HIDUP

Deta Erviana lahir pada tanggal 07 Agustus 2001 di Tulang Bawang. Penulis merupakan anak kedua dari pasangan Bapak Suwarto dan Ibu Wiwik, serta adik dari Dedi Ermawan.

Penulis memulai pendidikan di TK Mutiara Bunda pada tahun 2006 sampai dengan 2008, yang selanjutnya dilanjutkan dengan menempuh pendidikan di SDN 01 Gedung Rejo Sakti pada tahun 2008 sampai dengan 2013, kemudian melanjutkan pendidikan di SMPN 01 Penawar Aji Tahun pada tahun 2014 sampai dengan 2017, hingga sampai di bangku SMAN 01 Meraksa Aji pada tahun 2017 sampai dengan 2020.

Pada tahun 2020, penulis diterima sebagai mahasiswa (S1) Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SMMPTN. Selama masa perkuliahan, penulis mengikuti organisasi Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) sebagai anggota bidang keilmuan periode 2021.

Pada tahun 2023 penulis melaksanakan Kerja Praktek (KP) di Badan Pusat Statistik (BPS) sebagai bentuk penerapan bidang ilmu di dunia kerja. Kemudian pada pertengahan tahun 2023 penulis mengikuti Kuliah Kerja Nyata (KKN) di Pekon Pardasuka, Kecamatan Kota Agung Pusat, Kabupaten Tanggamus.

KATA INSPIRASI

Ambillah resiko yang lebih besar dari apa yang dipikirkan orang lain aman. Berilah perhgatian lebih dari apa yang orang lain pikir bijak. Bermimpilah lebih dari apa yang orang lain masuk akal.

(Claude T. Bissel)

Janganlah pernah menyerah ketika kamu masih mampu berusaha lagi. Tidak ada kata berakhir sampai kamu berhenti mencoba.

(Brian Dyson)

Jadi bersabarlah. Sungguh, janji Allah adalah kebenaran. Dan janganlah mereka mengganggu kamu yang tidak yakin.

(Q.S. Ar-Rum: 60)

Jangan biarkan ketakutan menghalangi langkahmu menuju impianmu. Percayalah, setiap langkahmu membawamu lebih dekat pada keberhasilan.

(Anonym)

Di setiap titik jatuhmu, ada potensi untuk terbang lebih tinggi daripada sebelumnya.

(Anonym)

PERSEMBAHAN

Alahmdulillahirabbil'alamin, puji syukur saya panjatkan kepada Allah SWT yang telah memberikan kemudahan serta kelancaran sehingga skripsi ini dapat terselesaiakan dengan baik. Dengan ini, saya persembahkan skripsi ini kepada:

Bapak dan Ibu

Malaikat tanpa sayap yang diberikan Allah. Ibu terimakasih atas segala cinta, kasih sayang, nasehat, kesabaran, dukungan serta do'a yang tidak pernah ada ujungnya, serta Almarhum Bapak yang selalu menginginkan kedua anaknya untuk tetap menempuh pendidikan setinggi mungkin dan saya yakin Bapak bangga melihat saya sekarang ini dari surga-Nya. I love you both beyond words.

Kakak

Terimakasih telah mendoa'akan, memberikan dukungan, semangat serta motivasi.

Dosen

Terimakasih untuk dosen-dosen pembimbing dan pembahas yang telah membimbing saya selama ini dengan baik, semoga selalu diberikan kesehatan serta umur yang panjang oleh Allah SWT.

Teman dan Sahabat

Untuk teman dan sahabat saya terimakasih atas semua dukungan, semangat, bantuan dan kerjasamanya selama ini

Almamater kebanggaan, Universitas Lampung.

SANWACANA

Segala puji bagi Allah SWT, yang telah memberikan rahmat, taufik serta hidayah-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul "**Penerapan Metode Geographically Weighted Panel Regression Pada Indeks Pembangunan Manusia di Indonesia Tahun 2017-2022**", sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Jurusan Matematika.

Dalam proses penyusunan skripsi, penulis menyadari bahwa skripsi tidak mungkin terselesaikan tanpa adanya bimbingan, nasehat serta bantuan dari banyak pihak, untuk itu penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D., selaku dosen pembimbing I yang ditengah kesibukannya telah meluangkan waktu untuk memberikan bimbingan, pengarahan, serta saran kepada penulis.
2. Ibu Widiarti, S.Si., M.Si., selaku dosen pembimbing II yang telah banyak memberikan bimbingan, saran dan kritik selama proses menyelesaikan skripsi ini.
3. Ibu Dr. Koirin Nisa, S.Si., M.Si., selaku dosen Pembahas skripsi yang telah memberikan evaluasi dan saran bagi perbaikan skripsi penulis.
4. Ibu Prof. Dra. Wamiliana, M.A., Ph.D. selaku dosen pembimbing akademik.
5. Bapak Dr. Aang Nuryaman, S.Si. M.Si., selaku ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung
7. Seluruh dosen, staf, dan karyawan Jurusan Matematika FMIPA UNILA yang telah memberikan ilmu pengetahuan dan segala bantuan kepada penulis.

8. Keluargaku tersayang, terutama Ibu yang telah mendukung, membimbing, mendoakan, serta memberikan semangat kepada penulis .
9. Sahabat serta teman-temanku seperjuangan yang selalu memberikan dukungan, semangat serta do’a kepada penulis dalam menyelesaikan skripsi ini.
10. Semua pihak terkait lainnya yang telah membantu dalam menyelesaikan skripsi ini yang tidak dapat penulis sebutkan satu per satu.

Penulis sangat menyadari bahwa penulisan skripsi ini jauh dari kata sempurna. Oleh karena itu, semua saran serta kritik yang membangun sangat berarti. Penulis berharap semoga penulisan skripsi ini dapat bermanfaat bagi pembaca.

Bandar Lampung, 25 Maret 2024
Penulis

Deta Erviana

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1. Latar Belakang Masalah

Regresi linier merupakan metode statistika untuk memeriksa hubungan antara respon dan satu atau lebih variabel prediktor. Metode *Ordinary Least Square* (OLS) dapat digunakan untuk menduga parameter dalam regresi linier klasik. Koefisien regresi dalam metode OLS diasumsikan berlaku secara global untuk keseluruhan unit pengamatan. Menurut Fotheringham dkk (2002), "informasi yang akurat untuk wilayah lokal jika tidak ada atau sedikit variasi keragaman antar wilayahnya lokalnya didapatkan dari model persamaan global". Sebuah model regresi dianggap sangat baik jika memenuhi asumsi klasik seperti tidak ada autokorelasi, heterokedastisitas, dan residual model berdistribusi normal (Gujarati, 2004).

Dalam suatu penelitian, perlu dilakukan pengamatan terhadap unit pada berbagai periode waktu, bukan hanya satu periode waktu tertentu saja. Oleh karena itu, berkembanglah suatu metode statistika yang menggabungkan data *cross section* dan *time series* yaitu analisis regresi data panel. Menurut Baltagi (2005), "data panel mempunyai beberapa keunggulan yaitu (1) data lebih infomatif, bervariasi dan efisien, (2) dapat menghindari masalah multikolinearitas, (3) lebih baik dalam mempelajari perubahan yang dinamis, (4) lebih baik dalam mengukur pengaruh-pengaruh yang tidak dapat diamati pada data *cross section* murni dan *time series* murni, dan (5) dengan membuat data tersedia menjadi lebih banyak, maka data panel dapat mengurangi bias yang terjadi pada saat mengagregatkan individu ke dalam agregat yang lebih luas". Data panel lebih baik untuk memahami dinamika perubahan karena data panel memeriksa *cross section* yang berulang dari suatu observasi.

Namun pada kenyataannya, terdapat situasi dimana kondisi disuatu lokasi dapat mempengaruhi kondisi di lokasi lain pada periode tertentu. Kondisi tersebut, dapat menyebabkan terjadinya heterogenitas spasial karena adanya pengaruh efek spasial atau kondisi geografis suatu wilayah yang diamati. Menurut Fotheringham, dkk (2002) "masalah heterogenitas spasial dapat dideteksi menggunakan metode *Geographically Weighted Regression* (GWR)". Dalam model GWR, estimasi parameter dapat dilakukan dengan menggunakan pendekatan "*Weighted Least Square*" (WLS), dimana setiap lokasi pengamatan diberikan unsur pembobot. Oleh karena itu, dalam model GWR nilai parameter yang dihasilakan berbeda di setiap lokasi pengamatan.

Penelitian sebelumnya, seperti Wang (2006), Huang dkk (2010) memperluas ide kedekatan dalam GWR ke dimensi ruang dan waktu, yaitu *"Geographically and Temporally Weighted Regression*" (GTWR). Dalam mengakomodasi adanya heterogenitas secara *spatial* (lokasi) dapat menggunakan berkembanglah model GTWR yang merupakan pengembangan dari model GWR. Untuk mengetahui adanya keragaman spasial dan temporal, model GTWR menggabungkan informasi spasial dan temporal dalam matriks pembobot. Selanjutnya, Yu (2010) melakukan penelitian untuk melihat perkembangan wialayah ekonomi di Beijing dengan menggabungkan model GWR dan model regresi panel untuk pertama kalinya, yaitu model *"Geographically Weighted Panel Regression*" (GWPR) menggunakan "*Fixed Effect Model*", dan memperoleh hasil model GWPR lebih baik dari model data panel dan GWR *cross-section*. Beberapa studi penelitian lain juga menyelidiki model GWPR dan menunjukkan hasil yang sama dengan Yu (2010). Selain itu, Cai dkk (2014) menggunakan model GWPR untuk melihat dampak variasi iklim terhadap produksi jagung di Amerika Serikat.

Menurut Amaluddin, dkk (2018), kemajuan pembangunan manusia dibanyak negara termasuk indonesia, dikatakan sebagai indikator keberhasilan dalam meningkatkan kesejahteraan rakyat. “Indeks Pembangunan Manusia” (IPM) yang dikembangkan dan dikenalkan oleh “*United Nation Development Programme*” (UNDP) serta dipublikasikan secara tahunan dalam *Human Development Report* (HRD), digunakan oleh beberapa lembaga resmi diberbagai negara untuk mengukur kemajuan pembangunan manusia (Yolanda, 2017). Banyak negara yang telah menjadikan IPM sebagai standar untuk memahami pola dari pembangunan sosial ekonomi. Dengan ukuran ini, ditekankan bahwa penilaian terhadap kemajuan suatu negara seharusnya tidak berfokus pada pertumbuhan ekonomi, tetapi juga menekankan manusia dan kemampuan mereka sebagai kriteria utama.

Menurut BPS (2022), pandemi Covid-19 memiliki dampak yang signifikan terhadap pembangunan manusia, termasuk Indonesia. Ditengah pandemi Covid-19, IPM Indonesia hanya tumbuh sebesar 0,03% pada tahun 2020. Namun pada tahun 2021-2022, terjadi peningkatan pertumbuhan IPM Indonesia, dengan pertumbuhan sebesar 0,49% pada tahun 2021 dan 0,86% pada tahun 2022, lebih tinggi dibandingkan sebelum masa pandemi Covid-19. Peningkatan dimensi standar hidup layak yang diwakili oleh indikator pengeluaran riil per kapita yang disesuaikan mampu mendorong peningkatan IPM di Indonesia tahun 2022. Indikator ini pada tahun 2022 tumbuh 2,90%, setelah tahun sebelumnya hanya tumbuh 1,30%. Selain itu, dimensi umur panjang dan hidup sehat yang direpresentasikan oleh pertumbuhan indikator Umur Harapan Hidup (UHH) saat lahir sebasar 0,39%, serta dimensi pengetahuan yang direpresentasikan oleh pertumbuhan indikator Harapan Lama Sekolah (HLS) dan Rata-rata Lama Sekolah (RLS) masing-masing sebesar 0,15% dan 1,76%. Meskipun Papua memiliki IPM terendah sebesar 61,39%, pertumbuhan IPM di provinsi tersebut mencapai 1,27%, yang merupakan pertumbuhan tertinggi dibandingkan provinsi lain. DKI Jakarta tetap menjadi provinsi dengan IPM tetringgi, mencapai 81,65%. Pada tahun 2022, DKI Jakarta bahkan meningkat statusnya dari “tinggi” menjadi “sangat tinggi” dengan capaian IPM sebesar 80,64%.

Indeks Pembangunan Manusia (IPM) mengukur pencapaian dalam meningkatkan kapasitas dasar manusia. IPM terdiri dari tiga dimensi penyusun, yaitu dimensi kesehatan, dimensi pendidikan dan dimensi pengeluaran (BPS, 2015). Letak geografis suatu wilayah dapat berdampak pada nilai IPM, terutama jika nilai IPM berubah dari waktu ke waktu, oleh karena itu pendekatan data panel tepat untuk *study of dynamic adjustment* data IPM (Wibisono, 2005). Faktor-faktor yang mempengaruhi tingkat IPM suatu wilayah diperkirakan akan bervariasi sesuai dengan karakteristik dari masing-masing wilayah.

Secara statistik, salah satu cara untuk meningkatkan IPM di Indonesia adalah mengindentifikasi faktor-faktor yang berpengaruh signifikan terhadap IPM menggunakan pemodelan GWPR. Penelitian sebelumnya dilakukan oleh Qur'ani (2014) membahas tentang pemodelan *Geographically Weighted Regression Panel* (GWPR) sebagai pendekatan Model *Geographically Weighted Regression* (GWR) dengan menggunakan *Fixed Effect Model Time Trend*, dimana model GWPR memberikan hasil yang sangat baik yang ditunjukkan dengan nilai koefisien determinasi model GWPR yang tinggi. Kemudian Melliana & Zain (2013) melalukan penelitian menggunakan regresi data panel untuk mengetahui faktor-faktor yang mempengaruhi indeks pembangunan manusia di Kabupaten/Kota Provinsi Jawa Timur dan memperoleh hasil jumlah fasilitas kesehatan dan tingkat partisipasi angkatan kerja berpengaruh signifikan terhadap IPM. Selanjutnya Ningrum, dkk (2020) melakukan penelitian yang membahas pengaruh kemiskinan, tingkat pengangguran, pertumbuhan ekonomi, dan pengeluaran pemerintah terhadap indeks pembangunan manusia di Indonesia tahun 2014-2018 dalam prespektif islam memperolah hasil bahwa tingkat kemiskinan dan tingkat pengangguran berpengaruh signifikan terhadap IPM.

Berdasarkan permasalahan diatas, pada penelitian ini akan dilakukan penerapan model *Geographically Weighted Panel Regression* pada data IPM di Indonesia tahun 2017-2022, dengan satu variabel respon yaitu IPM dan delapan variabel prediktor yaitu Tingkat Pengangguran Terbuka (TPT), Pengeluaran per Kapita (PP), Jumlah Penduduk Miskin (JKM), Indeks Kemahalan Konstruksi (IKK), Angka Partisipasi Sekolah (APS), Rata-Rata Lama Sekolah (RRLS), Persediaan Sumber Air Minum (PSAM), Gini Ratio (GR). Dimana nantinya akan dilihat dari ke delapan variabel prediktor tersebut, variabel mana yang berpengaruh paling signifikan terhadap IPM.

1.2. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Mengetahui faktor-faktor yang mempengaruhi Indeks Pembangunan Manusia di Indonesia tahun 2017-2022.
2. Mengidentifikasi model terbaik antara model regresi global atau model *Geographically Weighted Panel Regression* untuk data Indeks Pembangunan Manusia di Indonesia tahun 2017-2022.

1.3. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1. Memberikan pengetahuan yang baru bagi penulis tentang model regresi data panel dan regresi spasial yaitu *Geographically Weighted Panel Regression*.
2. Memberikan informasi kepada pemerintah pusat dan daerah mengenai indikator-indikator Indeks Pembangunan Manusia untuk merumuskan dasar penyusunan kebijakan.
3. Memberikan pengetahuan dan informasi kepada pembaca mengenai pembentukan model regresi data panel dan *Geographically Weighted Panel Regression*.

TINJAUAN PUSTAKA

2.1. Regresi Linier

Menurut Gujarati (2004) suatu teknik yang digunakan untuk menentukan seberapa besar pengaruh variabel respon terhadap prediktor adalah analisis regresi. Dalam metode regresi, beberapa asumsi dasar dapat digunakan untuk membangun estimator linear tak bias terbaik dalam model regresi yang diperoleh menggunakan pendekatan kuadrat terkecil (*ordinary least squares*). Asumsi-asumsi dasar tersebut dikenal dengan asumsi klasik yang terdiri atas homoskedastisitas, non-autokorelasi, non-multikolinearitas, dan normalitas residual. Hubungan stokastik antara dua variabel X dan Y merupakan bentuk paling sederhana dalam model regresi linear.

$$y_i = \alpha + \beta X_i + \varepsilon_i, \qquad i = 1, 2, \dots, n \tag{2.1}$$

dengan:

y	: Variabel respon
X	: Variabel prediktor
ε	: Variabel gangguan
α dan β	: Parameter regresi
i	: Pengamatan ke-i

Parameter α dan β ditaksir pada dasar yang tersedia untuk variabel X dan y.

Model analisis regresi linear dengan sampel n dan jumlah prediktor k, secara umum bentuk persamaannya dapat dituliskan sebagai berikut:

$$y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_2 X_{i2} + \beta_k X_{ik} + \varepsilon_i, \quad i = 1,2,\ldots.,n \qquad (2.2)$$

dengan:

y_i : Pengamatan ke-i variabel respon

β_0 : Parameter *intercept* (konstanta)

$\beta_1, \beta_2, \ldots, \beta_k$: Parameter koefisien regresi variabel prediktor

$X_{i1}, X_{i2}, \ldots X_{ik}$: Pengamatan ke-i pada variabel prediktor dimana $j = 1, 2, \ldots, k$ dan $i = 1, 2, \ldots, n$

ε_i : *Error* (galat kesalahan), $\varepsilon_i \sim N(0; \sigma^2)$

k : Banyaknya variabel prediktor

Dalam bentuk matriks, persamaan umum model regresi linier yaitu:

$$\boldsymbol{y} = \boldsymbol{\beta X} + \boldsymbol{\varepsilon} \qquad (2.3)$$

dengan:

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}; \quad \boldsymbol{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1k} \\ 1 & X_{21} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{nk} \end{bmatrix}; \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}; \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan:

$\boldsymbol{y}$: Vektor respon

$\boldsymbol{X}$: Matriks prediktor dengan ukuran $n \times (p + 1)$

$\boldsymbol{\beta}$: Vektor parameter $(\beta_0, \beta_1, \ldots, \beta_k)^T$ dan

$\boldsymbol{\varepsilon}$: Vektor *error* dengan notasi $(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n)^T$, $\varepsilon_i \sim iid\ N(0, \sigma^2)$.

Jumlah kuadrat *error* dapat diminimumkan menggunakan Metode *Ordinary Least Square* (OLS), sehingga diperoleh estimator parameter $\boldsymbol{\beta}$.

$$
\begin{aligned}
\min \boldsymbol{\varepsilon}^{\boldsymbol{T}}\boldsymbol{\varepsilon} &= min \sum_{i}^{n} \varepsilon_i^2 \\
\boldsymbol{\varepsilon}^{\boldsymbol{T}}\boldsymbol{\varepsilon} &= (\boldsymbol{y}-\boldsymbol{X\beta})^{\boldsymbol{T}}(\boldsymbol{y}-\boldsymbol{X\beta}) \\
&= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{X\beta} - (\boldsymbol{X\beta})^{\boldsymbol{T}}\boldsymbol{y} + (\boldsymbol{X\beta})^{\boldsymbol{T}}\boldsymbol{X\beta} \\
&= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{X\beta} - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} + \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X\beta} \\
&= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - (\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{X\beta})^{\boldsymbol{T}} - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X\beta} \\
&= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} + \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X\beta} \\
&= \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{2}\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} + \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X\beta} \qquad (2.4)
\end{aligned}
$$

$\boldsymbol{\varepsilon}^{\boldsymbol{T}}\boldsymbol{\varepsilon}$ diturunkan terhadap $\boldsymbol{\beta}$, kemudia disamakn dengan nol, akan menghasilkan penduga parameter $\boldsymbol{\beta}$ pada model regresi.

$$
\begin{aligned}
\frac{\partial(\boldsymbol{\varepsilon}^{\boldsymbol{T}}\boldsymbol{\varepsilon})}{\partial\boldsymbol{\beta}} &= 0 \\
0 - 2\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} + 2\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}\hat{\boldsymbol{\beta}} &= 0
\end{aligned}
$$

sehingga,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{y} \qquad (2.5)$$

dimana $\hat{\boldsymbol{\beta}}$ merupakan penaksir tak bias untuk $\boldsymbol{\beta}$.

2.1.1. Pengujian Hipotesis Model Regresi Linear

Uji serentak yaitu uji yang digunakan untuk mengetahui kecocokan model dengan hipotesis sebagai berikut:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0$

$H_1: minimal\ ada\ satu\ \beta_k \neq 0, dimana\ k = 1,2, \dots, k$

Tabel 2.1. Analisis Varians Model Regresi

Sumber Keragaman	Jumlah Kuadrat	Derajat Bebas	Rata-rata Kuadrat	F-Hitung
Regresi	$\boldsymbol{\hat{\beta}^T X^T y} - n\bar{y}^2$	p	$MSR = \frac{SSR}{p}$	$F = \frac{MSR}{MSE}$
Eror	$\boldsymbol{y^T y} - \boldsymbol{\hat{\beta}^T X^T y}$	$n - p - 1$	$MSE = \frac{SSE}{n - p - 1}$	
Total	$\boldsymbol{y^T y} - n\bar{y}^2$	$n - 1$		

Sumber: Rencher & Schaalje (2008).

Daerah penolakan H_0 yaitu apabila $F_{hitung} > F_{\alpha,p,n-p-1}$ atau nilai *p-value* $< \alpha$.

Parameter yang berpengaruh signifikan terhadap model dapat diketahui menggunakan uji secara parsial, dengan hipotesis sebagai berikut:

Pengujian secara parsial digunakan untuk melihat parameter mana saja yang signifikan terhadap model dengan hipotesis sebagai berikut:

$H_0: \beta_k = 0$

$H_1: \beta_k \neq 0$

Statistik uji yang digunakan yaitu:

$$t = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)} \tag{2.6}$$

$se(\hat{\beta}_k)$ adalah standar *error* dari koefisien $\hat{\beta}_k$.

Dibawah H_0, t akan mengikuti distribusi t dengan derajat bebas $(n-p-1)$, karena itu, jika diberikan tingkat signifikansi sebesar α maka keputusan yang diambil tolak H_0 jika $|t_{hitung}| > t_{\frac{\alpha}{2};n-p-1}$. Seberapa besar pengaruh variansi variabel respon yang dijabarkan oleh model regresi dapat dihitung menggunakan koefisien determinasi R^2 yaitu:

$$\begin{aligned} R^2 &= 1 - \frac{SSE}{SST} \\ &= 1 - \left(\frac{\boldsymbol{y}^T\boldsymbol{y} - \hat{\boldsymbol{\beta}}^T\boldsymbol{X}^T\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y} - n\bar{y}^2}\right) \end{aligned} \tag{2.7}$$

2.1.2. Asumsi Klasik Model Regresi

Metode OLS (*ordinary least square*) dapat dibangun menerapkan sebagian asumsi klasik, seperti permasalahan multikolinearitas, heteroskedastisitas, autokorelasi dan normalitas residual.

a. Multikolinearitas

   Menurut Gujarati (2004) “hubungan linier antara beberapa atau semua variabel prediktor didalam model regresi disebut multikolinearitas”. Adanya korelasi atau hubungan linier yang signifikan antara banyak variabel prediktor dapat menyebabkan terjadinya multikolinearitas, akibatnya sulit untuk menjelaskan hubungan setiap variabel-variabel secara individual terhadap variabel respon (Pangestika, 2015).

   Multikolinearitas dapat dideteksi menggunakan beberapa indikator sebagaimana yang dijelaskan oleh (Gujarati, 2004):

1. Jika R^2 sangat tinggi tetapi tidak ada koefisien regresi yang signifikan secara statistik pada uji t.
2. Jika koefisien korelasi antara dua variabel prediktor cukup tinggi.
3. Jika R^2 sangat tinggi, namun korelasi parsial rendah.
4. Meregresikan tiap variabel x_k atas sisa variabel x dalam model dan mengetahui koefisien determinasi yang berhubungan dengan R_k^2. Suatu R_k^2 yang tinggi akan menguraikan bahwa x_k sengat berkorelasi dengan variabel x.

Multikolinearitas dapat diketahui dengan menghitung nilai *Variance Inflation Factors* (VIF), rumus sebagai berikut (Gujarati, 2004)

$$(VIF)_j = \frac{1}{1-R_j^2} \tag{2.8}$$

dengan R_j^2 yaitu nilai koefisien determinasi variabel ke j. Data di indikasikan terjadi multikolinearitas jika $(VIF)_j > 10$ (Isbiyantoro, dkk., 2014).

b. Heteroskedastisitas

Heterokedastisitas merupakan asumsi penting dalam model regresi linier, menyatakan *error* yang tampak dalam fungsi regresi bersifat homoskedastik, artinya memiliki varians yang sama (Gujarati, 2004).

$$E(\varepsilon_i^2) = \sigma^2, \quad i = 1, 2, \dots, n \tag{2.9}$$

Kesalahan terhadap asumsi ini dinamakan heteroskedastisitas, artinya varians *error* berbeda. Metode yang dapat digunakan untuk mengidentifikasi masalah heterokedastisitas yaitu Uji *Park*, Uji *Glejser*, Uji *White*, Uji *Breusch-Pagan-Goldfrey*, dan lain-lain. Uji *Glejser* merupakan salah satu metode yang dapat digunakan untuk menngidentifikasi masalah heteroskedastisitas. Nilai absolut dari ε_i atau $|\varepsilon_i|$ diregresikan terhadap semua variabel prediktor yang digunakan. Kasus heterokedastisitas terjadi apabila variabel prediktor yang berpengaruh signifikan mempengaruhi nilai absolut *error* (Isbiyantoro, dkk., 2014).

c. Autokorelasi

Menurut Gujarati (2004), "korelasi antara pengamatan dalam suatu rangkaian yang disusun menurut waktu (seperti dalam data *time series*) atau ruang (seperti dalam data *cross section*) disebut autokorelasi". Autokorelasi seperti itu tidak ditemukan dalam *error* model regresi klasik (Gujarati, 2004).

$$E(\varepsilon_i \varepsilon_j) = 0, \qquad untuk\ j \neq s \tag{2.10}$$

Uji Run adalah salah satu pendekatan untuk mendeteksi adanya autokorelasi sisaan (Gujarati, 2004). Uji Run digunakan untuk melihat apakah data sisaan terjadi secara acak atau tidak, jika terjadi secara acak maka sisaan saling bebas. Hipotesis uji Run yaitu (Walpole, dkk., 2007):

H_0: Sisaan saling bebas
H_1: Sisaan tidak saling bebas

Statistik uji:

$$Z = \frac{R-\mu_R}{\sigma_R} \tag{2.11}$$

dimana: $\mu_R = \frac{2n_1 n_2}{n_1 + n_2} + 1$ dan $\sigma_R^2 = \frac{(2n_1 n_2)(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}$

dengan:

R = Banyaknya jumlah run dengan run yaitu suatu urutan data sisaan dengan simbol/tanda yang sama

n_1 = Banyaknya koefisien sisaan yang bertanda positif

n_2 = Banyaknya koefisien sisaan yang bertanda negatif

μ_R = Nilai harapan run

σ_R = Standar deviasi run

Tolak H_0 jika nilai $|Z| > Z_{\frac{\alpha}{2}}$ atau nilai *p-value* $< \alpha$ artinya sisaan tidak saling bebas atau tidak terjadi autokorelasi.

d. Normalitas

Regresi linear klasik menaksirkan bahwa masing-masing *error* (ε_i) berdistribusi secara normal dengan rata-rata $E(\varepsilon_i) = 0$, varians $E(\varepsilon_i^2) = \sigma^2$, dan $E(\varepsilon_i \varepsilon_j) = 0, \quad untuk\ j \neq s$. Secara ringkas dapat ditulis:

$$\varepsilon_i \sim N(0, \sigma^2) \tag{2.12}$$

Menurut Gujarati (2004) "estimator OLS memiliki sifat tidak bias, efisien dan konsisten, dengan asumsi kenormalan". Uji *Jarque-Bera* (JB) adalah salah satu metode yang dapat digunakan untuk menentukan normalitas, dengan menghitung nilai *skewness* (ukuran kemiringan) dan *kurtosis* (ukuran keruncingan). Hipotesisnya sebagai berikut:

H_0: Sisaan menyebar normal

H_1: Sisaan tidak menyebar normal

Statistik uji:

$$JB = n\left[\frac{S_k^2}{6} + \frac{(K-3)^2}{24}\right] \tag{2.13}$$

dimana:

$$S_k = \frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^3}{(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2)^{3/2}}$$

$$K = \frac{\hat{\mu}_4}{\hat{\mu}_2^{2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^4}{(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2)^{2}}$$

dengan:

n = Banyaknya data

S_k = *Skewness* (kemiringan)

K = *Kurtosis* (keruncingan)

Tolak H_0 jika $JB > \chi^2$ atau *p-value* $< \alpha$ artinya sisaan tidak menyebar secara normal.

2.2. Regresi Data Panel

Kombinasi antara data *cross-section* dan *time series* disebut data panel. Menurut Jaya dan Sunengsih (2009) "analisis regresi data panel merupakan jenis analisis yang menggunakan data panel untuk mengidentifikasi pengaruh antara variabel respon dengan satu atau lebih variabel prediktor". Menurut Pangestika (2015), "model komponen satu arah merupakan model regresi panel yang hanya dipengaruhi oleh salah satu unit saja (unit *cross sectional* atau unit waktu), sedangkan model komponen dua arah yaitu model regresi panel yang dipengaruhi oleh kedua unit (unit *cross section*al atau unit waktu)".

Unit *cross section* dapat mencakup rumah tangga, bisnis, negara, orang dan sebagainya, sedangkan unit *time series* mencakup rentang waktu harian, bulanan, tahunan dan sejenisnya. Menurut Croissant (2008) “setiap unit *cross section* diamati secara berulang-ulang selama beberapa periode waktu, dan data panel memiliki struktur khusus, dimana setiap baris pada data sesuai dengan individu tertentu dan periode waktu”. Jika kita mempunyai individu (dimana $i = 1, 2, 3, \dots, N$) dalam periode pengamatan selama waktu (dimana $t = 1, 2, 3, \dots, T$), maka dengan data panel kita akan memiliki total observasi sebanyak NT. Data panel dibagi menjadi dua kategori, yaitu *balanced panel* dan *unbalanced panel*. Data panel yang jumlah periode waktu sama dengan individunya dinamakan *Balanced panel*, sedangkan data yang jumlah periode waktunya berbeda dengann individunya dinamakan *unbalanced panel*.

Menurut Hsiao (2003) secara umum model regresi data panel dapat dituliskan persamaannya sebagai berikut:

$$y_{it} = \alpha_{it} + \boldsymbol{\beta}^T \boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T \tag{2.14}$$

dengan:

y_{it} : Pengamatan untuk unit *cross section* ke-*i* pada periode waktu ke-*t*

α_{it} : Intersep, merupakan efek group/individu dari unit *cross section* ke-*i* pada periode waktu ke-*t*

$\boldsymbol{X}_{it}^T$: $(X_{1it}, X_{2it}, \dots, X_{pit})$ menunjukkan vektor observasi pada variabel prediktor berukuran $1 \times p$

$\boldsymbol{\beta}^T$: $(\beta_1, \beta_2, \dots, \beta_p)$ adalah vektor konstanta (*slope* koefisien) berukuran $1 \times p$

ε_{it} : *error* regresi dari individu ke-*i* untuk periode waktu ke-*t*

dengan asumsi bahwa ε_{it} tidak berkorelasi terhadap variabel prediktor dan berdistribusi $IIDN(0, \sigma^2)$.

Data panel secara umum akan menghasilkan intersep dan koefisien *slope* yang bervariasi untuk masing-masing individu dan periode waktu. Oleh sebab itu, untuk proses estimasi persamaan (2.14) sangat bergantung terhadap asumsi yang diterapkan pada intersep, koefisien *slope* dan variabel gangguannya (Hsiao, 2003).

Tabel 2.2. Struktur Data Panel Secara Umum

Lokasi (i)	Tahun (t)	Variabel Respon (y_{it})	Variabel Prediktor (X_{1it})	Variabel Prediktor (X_{2it})	...	Variabel Prediktor (X_{pit})
1	1	y_{11}	$X_{1.11}$	$X_{2.11}$	...	$X_{p.11}$
2	1	y_{21}	$X_{1.21}$	$X_{2.21}$	...	$X_{p.12}$
...	...	...	...	...	...	...
N	1	y_{N1}	$X_{1.N1}$	$X_{2.N1}$	...	$X_{p.N1}$
1	2	y_{12}	$X_{1.12}$	$X_{2.12}$	...	$X_{p.12}$
2	2	y_{22}	$X_{1.22}$	$X_{2.22}$	...	$X_{p.22}$
...	...	...	...	...	...	...
N	2	y_{N2}	$X_{1.N2}$	$X_{2.N2}$	...	$X_{p.N2}$
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
1	T	y_{1T}	$X_{1.1T}$	$X_{2.1T}$	...	$X_{p.1T}$
2	T	y_{2T}	$X_{1.2T}$	$X_{2.2T}$	...	$X_{p.2T}$
...	...	...	...	...	...	...
N	T	y_{NT}	$X_{1.NT}$	$X_{2.NT}$	...	$X_{p.NT}$

Sumber: Park (2005)

2.2.1. Pendekatan dan Metode Estimasi pada Model Regresi Data Panel

Model regresi dapat diestimasi menggunakan 3 pendekatan yaitu *common effect model*, *fixed effect model*, dan *random effect model*.

a. *Common Effect Model* (CEM)

Metode yang menggabungkan semua data tanpa mempertimbangkan waktu atau tempat kejadian, dengan asumsi bahwa perilaku data seluruh unit *cross-section* konsisten sepanjang waktu disebut *common effect model* (Rahmadeni & Wulandari, 2017). Menurut Baltagi (2005) "*common effect model* merupakan pendekatan yang menggabungkan (*pooled*) semua data *time series* dan *cross section* menggunakan metode *Ordinary Least Square* (OLS) untuk mengestimasi parameternya". Model CEM memiliki kekurangan yaitu terjadi ketidakcocokan model dengan kondisi sebenarnya, dimana masing-masing objek memiliki kondisi yang bervariasi, bahkan kondisi objek akan bervariasi dari waktu ke waktu (Winarno, 2015). Salah satu metode yang sangat terkenal untuk menngestimasi nilai parameter pada persamaan regresi linear yaitu metode OLS. Menurut Greene (2000) persamaan model CEM secara umum dapat dituliskan sebagai berikut:

$$y_{it} = \alpha + \boldsymbol{\beta}^T \boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \qquad (2.15)$$

dengan:

y_{it} : Variabel respon pada unit observasi ke-*i* dan waktu ke-*t*

$\boldsymbol{X}_{it}$: $(X_{1it}, X_{2it}, \dots, X_{pit})$ variabel prediktor untuk pengamatan ke-*i* pada periode waktu ke-*t* berukuran $1 \times p$

$\boldsymbol{\beta}^T$: Vektor konstanta (*slope* koefisien) berukuran $1 \times p$

α : *Intercept* model regresi

ε_{it} : Galat atau komponen *error* pada unit observasi ke-*i* dan waktu ke-*t*

b. *Fixed Effect Model* (FEM)

Model yang menyatakan bahwa koefisien intersep berbeda tiap individu disebut *Fixed Effect Model* (FEM) (Hsiao, 2003). Metode yang digunakan untuk melihat unit *cross-section* yaitu mengizinkan adanya variasi nilai intersep pada masing-masing unit *cross-section* dengan mengasumsikan koefisien *slope* bernilai tetap (Gujarati, 2004). Model FEM dapat dinyatakan sebagai berikut (Greene, 2000):

$$y_{it} = \alpha_i + \boldsymbol{\beta}^T \boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, T \quad (2.16)$$

Model *fixed effect within transformation* merupakan salah satu metode untuk mengestimasi model *fixed effect* (Wooldridge, 2002). Pendekatan ini dilakukan dengan cara menghilangkan efek unit *cross-section* (α_i), dan merata-ratakan nilai variabel repson dan variabel terikat sepanjang waktu untuk setiap unit *cross-section*. Persamaan berikut ini menyatakan *Fixed effect model within transformation* dan persamaan rata-rata:

$$y_{it} = \alpha_i + \sum_{k=1}^{p} \beta_k X_{kit} + \varepsilon_{it}$$
$$\bar{y}_i = \alpha_i + \sum_{k=1}^{p} \beta_k \bar{X}_{ki} + \bar{\varepsilon}_i$$

Persamaan *fixed effect within transformation* dikurangkan terhadap persamaan rata-rata, sehingga didapatkan persamaan berikut:

$$y_{it} - \bar{y}_i = \alpha_i - \alpha_i + \sum_{k=1}^{p} \beta_k (X_{kit} - \bar{X}_{ki}) + \varepsilon_{it} - \bar{\varepsilon}_i$$
$$y_{it} - \bar{y}_i = \sum_{k=1}^{p} \beta_k (X_{kit} - \bar{X}_{ki}) + \varepsilon_{it} - \bar{\varepsilon}_i$$

Persamaan tersebut dapat disederhanakan menjadi

$$y_{it}^{*} = \sum_{k=1}^{p} \beta_k X_{kit}^{*} + \varepsilon_{it}^{*} \tag{2.17}$$

dengan:

y_{it}^{*} : $y_{it} - \bar{y}_i$

X_{kit}^{*} : $X_{kit} - \bar{X}_{ki}$

ε_{it}^{*} : $\varepsilon_{it} - \bar{\varepsilon}_i$

Metode OLS dapat digunakan untuk mengestimasi model *fixed effect within transformation*.

c. *Random Effect Model* (REM)

Model yang menduga data panel meggunakan metode *Generalized Least Squae* (GLS), dimana masing-masing individu diperlakukan dengan komponen eror yang acak dan setiap variabel prediktor yang diamati tidak saling berkorelasi disebut *Random Effect Model* (REM). Menurut Greene (2000) persamaan REM dapat dinyatakan sebagai berikut:

$$y_{it} = (\alpha + u_i) + \boldsymbol{\beta}^T \boldsymbol{X}_{it} + \varepsilon_{it} \tag{2.18}$$

dengan:

y_{it} : Variabel respon pada unit observasi ke-*i* dan waktu ke-*t*

$\boldsymbol{X}_{it}$: $(X_{1it}, X_{2it}, \dots, X_{pit})$ variabel prediktor untuk pengamatan ke-*i* pada periode waktu ke-*t* berukuran $1 \times p$

$\boldsymbol{\beta}^T$: Vektor konstanta (*slope* koefisien) berukuran $1 \times p$

α : *Intercept* model regresi

ε_{it} : Galat atau komponen *error* pada unit observasi ke-*i* dan waktu ke-*t*

u_i : Komponen *error cross section*

Adapun asumsi yang digunakan untuk komponen eror tersebut yaitu:

$E[\varepsilon_{it}|X] = E[u_i|X] = 0$
$E[\varepsilon_{it}^2|X] = \sigma_\varepsilon^2$
$E[u_{it}^2|X] = \sigma_u^2$
$E[\varepsilon_{it}u_j|X] = 0$ untuk semua $i, t,$ dan j
$E[\varepsilon_{it}\varepsilon_{js}|X] = 0$ jika $t \neq s$ atau $i \neq j$
$E[u_i u_j|X] = 0$ jika $i \neq j$

Error ε_{it} tidak saling berkorelasi dan tidak berautokorelasi antar unit *cross section* maupun antar unit *time series* merupakan asumsi dalam model REM. Metode estimasi model REM adalah *Generalized Least Square* (GLS).

2.2.2. Pemilihan Metode Estimasi Model Regresi Data Panel

Pemilihan model regresi data panel dapat dilakukan menggunakan beberapa uji berikut ini:

1. Uji Chow

   Chow test merupakan pengujian statistik yang digunakan untuk memilih salah satu model pada regresi data panel, yaitu antara model *Common Effect* atau *Fixed Effect* yang paling tepat digunakan dalam mengestimasi data panel. Uji *Chow* merupakan uji perbedaan regresi seperti halnya uji F. Perbandingan metode CEM dan FEM dalam uji *Chow* dilihat dari nilai *residual sum of squares* (RSS). Menurut Greene (2000) hipotesisnya sebagai berikut:

   H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_N = a$ (model CEM)
   H_1: minimal ada satu intersep (α_1) yang tidak sama (model FEM)

Statistik uji (Baltagi, 2005)

$$F_{hitung} = \frac{(RRSS - URSS/(N-1)}{(URSS)/(NT-N-k)} \tag{2.19}$$

jika nilai $F_{hitung} > F_{\alpha,N-1,N(T-1)-k}$ atau *p-value* $< \alpha$ maka tolak H_0, artinya model yang tepat adalah FEM.

2. Uji Hausman

Pengujian statistik yang digunakan untuk menentukan model terbaik antara *Random Effect* dan *Fixed Effect* yaitu *Hausman test*. Menurut Greene (2000) hipotesisnya sebagai berikut:

$H_0 : corr\left(X_{it,}\varepsilon_i\right) = 0$ (model REM)
$H_1 : corr(X_{it,}\varepsilon_i) \neq 0$ (model FEM)

Statistik uji:

$$W = (\widehat{\boldsymbol{\beta}}_{FEM} - \widehat{\boldsymbol{\beta}}_{REM})^T [var(\widehat{\boldsymbol{\beta}}_{FEM}) - var(\widehat{\boldsymbol{\beta}}_{REM})]^{-1} (\widehat{\boldsymbol{\beta}}_{FEM} - \widehat{\boldsymbol{\beta}}_{REM}) \tag{2.20}$$

dengan:

$\widehat{\boldsymbol{\beta}}_{FEM}$: Vektor dari estimasi parameter FEM
$\widehat{\boldsymbol{\beta}}_{REM}$: Vektor dari estimasi parameter REM.

Tolak H_0 apabila nilai $W > \chi^2{}_{(k-1;a)}$ atau *p-value* $< \alpha$, dan disimpulkan model yang terbaik adalah model FEM.

3. Uji *Lagrange Multiplier* (LM)

Uji yang digunakan untuk menetapkan model terbaik antara *Random Effect* dan *Common Effect* yaitu *Langrange Multiplier*. Menurut Greene (2000), hipotesisnya sebagai berikut:

$H_0 : \sigma_u^2 = 0$ (model CEM)
$H_1 : \sigma_u^2 \neq 0$ (model REM)

Statistik uji:

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^{N} [\sum_{t=1}^{T} \varepsilon_{it}]^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{it}^2} - 1 \right]^2 \qquad (2.21)$$

Uji statistik LM mengikuti sebaran *Chi Square* dengan derajat bebas 1. Apabalia nilai uji LM lebih besar daripada $\chi^2_{(1)}$ berarti tolak H_0, artinya model REM lebih layak digunakan.

2.2.3. Pengujian Signifikansi Parameter Data Panel

Uji signifikansi dalam regresi data panel mirip seperti uji signifikansi pada regresi linier berganda. Uji ini berfungsi untuk menentukan adanya pengaruh variabel prediktor terhadap variabel respon, dan membantu untuk mengetahui bagaimana model dapat menunjukkan kondisi data yang sebenarnya. Uji parameter dalam regresi data panel ada dua, yaitu uji F dan uji T.

2.3. Aspek Data Spasial

Analisis spasial digunakan jika data tersebut memiliki aspek spasial, dimana terdapat korelasi antar *error* (*spatial dependence*) dan heterogenitas spasial (*spatial heterogeneity*).

2.3.1. *Spatial Dependence*

Ketergantungan spasial merupakan permasalahan yang terjadi karena perbedaan lokasi. Hukum pertama dalam geografi yang disampaikan oleh Tobler dalam Anselin (1988) menyatakan bahwa "*Everything is related to everything else, but near thing are more related than distant things*". Menurut Anselin (1988) untuk menentukan apakah observasi disuatu wilayah memiliki pengaruh pada observasi di wilayah lain yang berdekatan perlu dilakukan pengujian *spatial dependence*. Uji ini dapat dilakukan dengan menggunakan *Moran's I Statistic*.

$$I = \left[\frac{N}{S}\right] . \left\{\frac{|\boldsymbol{\varepsilon}^{T} \boldsymbol{W} \boldsymbol{\varepsilon}|}{\boldsymbol{\varepsilon}^{T} \boldsymbol{\varepsilon}}\right\} \tag{2.22}$$

dengan:

$\boldsymbol{\varepsilon}$: Vektor residual dari OLS,

$\boldsymbol{W}$: Matriks pembobot spasial

N : Jumlah observasi, dan

S : Faktor standarisasi

$\left[\frac{N}{S}\right]$ bernilai satu jika matriks pembobot yang terstandarisasi. Hipotesis pada statistik *Morans' I* yaitu:

H_0: Tidak terjadi *spatial dependence*

H_1: Terjadi *spatial dependence*

Statistik uji yang digunakan yaitu:

$$Z_I = \frac{(\hat{I} - E(\hat{I}))}{\sqrt{Var(\hat{I})}} \tag{2.23}$$

Tolak H_0 jika $|Z_I| \geq Z_{a/2}$

2.3.2. *Spatial Heterogeneity*

Heterogenitas spasial terjadi ketika terdapat perbedaan kondisi antar lokasi disuatu wilayah. Perbedaan ini dapat terlihat dari segi geografis, sosial budaya atau faktor-faktor lain yang dapat menyebabkan heterogenitas spasial pada lokasi pengamatan (Munikah, dkk., 2014). Uji *Breusch Pagan* merupakan salah satu uji yang dapat digunakan untuk mendeteksi heterogenitas spasial. Hipotesis sebagai berikut:

H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$ (Tidak terjadi heterogenitas spasial)
H_1: minimal ada satu $\sigma_i^2 \neq \sigma^2$ (Terjadi heterogenitas spasial)

Nilai dari *Breusch Pagan test* adalah:

$$BP = \left(\frac{1}{2}\right) \boldsymbol{f}^T \boldsymbol{Z}(\boldsymbol{Z}^T \boldsymbol{Z})^{-\mathbf{1}} \boldsymbol{Z}^T \boldsymbol{f} \qquad \sim \chi^2_{(p)} \tag{2.24}$$

dengan:

$\boldsymbol{f} = (f_1, f_2, \dots, f_n)^T$: $f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$

$e_i = y_i - \hat{y}_i$: *Least square* residual untuk pengamatan ke-i.

Z : Matriks berukuran $n \times (p+1)$ yang berisi vektor yang sudah di normal standarkan untuk setiap pengamatan.

Tolak H_0 bila $BP > \chi^2_{(p)}$ atau jika nilai *p-value* $< \alpha$.

2.4. *Geographically Weighted Regression* (GWR)

GWR adalah evolusi dari model regresi linear OLS ke model regresi terboboti yang memperdulikan efek spasial. Hal ini dapat membentuk estimasi parameter yang dapat digunakan untuk menduga nilai pada masing-masing titik atau lokasi dimana data tersebut diobservasi dan dianalisis (Fotheringham, dkk., 2002). Teknik GWR dinilai lebih sederhana dan lebih berguna secara geografis dalam mengeksplorasi ketidakstasioneran dibandingkan dengan metode lain (Leung, 2000). Nilai parameter dalam model regresi bervariasi disetiap lokasi. Model GWR dapat dinyatakan sebagai berikut:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) X_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, N \qquad (2.25)$$

dengan:

y_i : Nilai observasi variabel respon ke-*i*

X_{ik} : Nilai observasi variabel prediktor ke-*k* pada pengamatan ke-*i*

β : Koefisien regresi

(u_i, v_i) : Titik koordinat lokasi *i*

ε_i : *Error* ke-i

Bentuk *error* $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ diasumsikan independen, identik, dan mengikuti distribusi normal dengan *mean* nol dan variasi konstan ($\varepsilon_i \sim iidN(0, \sigma^2)$).

2.5. Koordinat Spasial

Longitude dan *Latitude* merupakan koordinat spasial untuk pembobotan dalam membentuk model GWR. Menurut Lesage (1999) koordinat *longitude* dan *latitude* pada suatu pengamatan, akan membentuk struktur kedekatan yang disebut dengan *neighbourhood*. Adapun garis membujur yang menyambungkan sisi utara dan selatan bumi (kutub), untuk mengukur sisi barat-timur koordinat suatu titik pada belahan bumi disebut dengan garis bujur (*longitude*). Sedangkan garis melintang yang berada di antara kutub utara dan kutub selatan sebagai penghubung sisi timur dan barat bumi yang digunakan untuk mengukur sisi utara-selatan dari koordinat suatu titik di belahan bumi disebut dengan garis lintang (*latitude*) (Isbiyantoro, dkk, 2014).

2.6. Fungsi Pembobot Spasial

Pembobot pada model GWR memiliki peran yang sangat penting karena nilai pembobot tersebut mewakili letak data observasi satu dengan lainnya. Menurut Gwarda (2018), setiap persamaan disesuaikan menggunakan pembobot yang berbeda dari observasi yang terdapat dalam data. Diasumsikan bahwa pengamatan yang dekat satu sama lain mempunyai pengaruh yang besar untuk setiap estimasi parameter daripada pengamatan yang jaraknya lebih jauh, seperti pada Hukum Tobler. Ada beberapa literatur yang digunakan untuk menentukan besarnya pembobot untuk masing-masing lokasi yang berbeda pada model GWR, diantaranya:

a. Fungsi *Invers* Jarak (*Invers Distance Function*)

Misalkan $1/d_{ij}$ adalah fungsi *invers* jarak yang mewakili pembobot antara lokasi (u_i, v_i) dan lokasi (u_j, v_j) $dimana\ d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak *Euclidean* antara lokasi (u_i, v_i) dan lokasi (u_j, v_j). Pembobot ini dapat ditulis:

$$w_{ij} = \begin{cases} 1, jika\ d_{ij} \leq h \\ 0, jika\ d_{ij} > h \end{cases} \tag{2.26}$$

b. Fungsi Pembobot Kernel (*Kernel Function*)

Pembobot yang dihasilkan dengan memanfaatkan fungsi kernel ini dikategorikan sebagai berikut:

1. Fungsi *Kernel Fixed*, merupakan fungsi kernel yang mempunyai *bandwidth* yang sama untuk masing-masing lokasi pengamatan. Beberapa contoh fungsi kernel dalam kategori ini:

(i) *Gaussian*

Menggunakan kernel *Gaussian*, dimana pembobot data dapat mengalami penurunan sesuai dengan kurva *Gaussian* sejalan dengan peningkatan jarak antara i dan j, yang diukur sebagai d_{ij} (Pijneburg, 2013).

$$w_{ij} = exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right) \tag{2.27}$$

(ii) *Bisquare*

Kernel *Bisquare* bersifat kontinu, apabila *bandwidth* melewati jarak i dan j (d_{ij}), nilai pembobot akan diatur ke nol (Pijnenburg, 2013).

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^2\right)^2, untuk\ d_{ij} \leq h \\ 0, untuk\ lainnya \end{cases} \tag{2.28}$$

2. Fungsi Kernel Adaptif merupakan fungsi kernel yang mempunyai nilai *bandwidth* yang bervariasi untuk masing-masing lokasi pengamatan.

   (i) *Adaptive Gaussian*

   Menurut Susanti, dkk (2016), fungsi *adaptive gaussian* yang diusulkan oleh Brunsdon, dkk (1996) mempunyai bentuk persamaan:

$$w_{ij} = exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{h_i}\right)^2\right) \quad (2.29)$$

   Dimana h_i sebagai parameter non negatif adalah paramater yang sering disebut sebagai *bandwidth*. Nilai bobot pada sebuah data akan mendekati 1 jika jaraknya berdekatan dan akan semakin kecil mendekati 0 jika jaraknya semakin jauh.

   (ii) *Adaptive Bisquare*

   Menurut Susanti, dkk (2016), fungsi *adaptive bisquare* yang diusulkan oleh Brunsdon, dkk (1996) memiliki persamaan:

$$w_{ij} = \begin{cases} \left(1-\left(\frac{d_{ij}}{h_i}\right)^2\right)^2, untuk\ d_{ij} \leq h_i \\ 0, untuk\ lainnya \end{cases} \quad (2.30)$$

   Fungsi pembobot kontinu ini serupa dengan fungsi *gaussian* hingga mencapai jarak h_i pada lokasi observasi ke-*i* dan memiliki nilai 0 pada lokasi data yang mempunyai jarak lebih jauh dibandingkan h_i. Nilai h_i adalah *bandwidth* yang menjelaskan jumlah atau proporsi dari pengamatan yang akan dimasukkan dalam estimasi regresi pada lokasi observasi ke-*i*.

c. *Adaptive Exponential*

Menurut Pamungkas, dkk (2016) fungsi *adaptive kernel exponential* yaitu:

$$w_{ij} = exp\left(\frac{-d_{ij}}{h_i}\right) \tag{2.31}$$

Dimaha h_i sebagai parameter non negatif adalah paramater yang sering disebut sebagai *bandwidth*. Nilai bobot pada sebuah data akan mendekati 1 jika jaraknya berdekatan dan akan semakin kecil mendekati 0 jika jaraknya semakin jauh.

2.7. Penentuan *Bandwidth*

Menurut Baltagi (2005) *bandwidth* dalam GWR dapat disamakan dengan radius lingkaran, sehingga obsevasi yang berada dalam tersebut dianggap mempengaruhi pembentukan parameter pada lokasi observasi tertentu. Menurut Fotheringham, dkk (2002), sebagian besar, pemilihan fungsi pembobot memiliki dampak signifikan pada estimasi parameter dalam GWR. Dalam fungsi pembobot *adaptive kernel gaussian*, jika nilai d meningkat, maka bobot mendekati 0, menjadikan model mendekati OLS. Sebaliknya, apabila nilai d menggambarkan jarak maksimum antar lokasi observasi, model yang dihasilkan serupa dengan model regresi global atau OLS (Susanti, dkk., 2016).

Pembobot (w_{ij}) antar lokasi observasi mendekati 1 saat *bandwidth* (h) mendekati tak terhingga, menyebabkan parameter yang diestimasi menjadi sama, sehingga model GWR yang dihasilkan akan mendekati model OLS. Sedangkan nilai *bandwidth* menurun, estimasi parameter menjadi bergantung pada lokasi observasi yang mempunyai jarak berdekatan dengan lokasi observasi ke-i, mengakibatkan varians yang dihasilkan meningkat. Bagaimana menentukan nilai *bandwidth* atau fungsi pembobot yang tepat dalam pemodelan GWR merupakan tantangan yang harus diselesaikan (Susanti, dkk., 2016).

Menurut Bidanset & Lombard (2014) dalam model GWR, jarak mengoptimalkan ukuran *bandwidth fixed kernel*. Salah satu cara untuk mendapatkan *bandwidth* optimum dapat menggunakan metode *Cross Validation* (CV), dimana nilai CV minimum yang dihasilkan menandakan *bandwidth* yang optimum (Fotheringham, dkk., 2002).

$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(h)]^2 \quad (2.32)$$

Dengan $\hat{y}_{\neq i}(h)$ merupakan nilai estimasi untuk y_i yang didapatkan dengan menghapus pengamatan di titik i selama proses pengujian parameter. Teknik *Golden Section Search* merupakan teknik yang dapat digunakan untuk mencari *bandwidth* yang meminimumkan nilai CV (Fotheringham, dkk., 2002).

2.8. *Geographically Weighted Panel Regression* (GWPR)

Gagasan pokok dalam analisis GWPR mirip dengan analisis GWR *cross sectional*. Model pengembangan yang menggabungkan antara model GWR dengan regresi data panel disebut GWPR. Pada GWPR diduga bahwa runtun waktu (*time series*) pada suatu observasi dalam suatu lokasi geografis adalah hasil dari proses *smooth spatiotemporal*. Proses ini mengikuti distribusi dimana pengamatan yang berdekatan (baik dalam satu lokasi geografis atau waktu) memiliki hubunya yang kuat dibandingkan dengan pengamatan yang lebih jauh. Menurut Yu (2010) tujuan dari analisis GWPR yaitu memadukan semua lokasi (*cross-sectional*) dan observasi. Persamaan model GWPR menggunakan model FEM diperoleh dengan menggabungkan model GWR dan regresi panel *fixed effect within transformation* yaitu sebagai berikut:

$$y_{it} = \beta_0(u_{it}, v_{it}) + \sum_{k=1}^{p} \beta_k(u_{it}, v_{it}) X_{itk} + \varepsilon_{it}$$
$$i = 1, 2, \dots, N \ dan \ t = 1, 2, \dots, T \tag{2.33}$$

dengan:

y_{it}	: Variabel respon di lokasi pengamatan ke-*i* pada waktu ke-*t*
$\beta_0(u_{it}, v_{it})$	: *Intercept* dari persamaan yang terbentuk pada pengamatan ke-*i* dan waktu ke-*t*
$\beta_k(u_{it}, v_{it})$	: Koefisien regresi variabel prediktor ke-*k* di lokasi pengamatan ke-*i* pada waktu ke-*t*
(u_{it}, v_{it})	: Titik koordinat letak geografis lokasi pengamatan ke-*i* dan waktu ke-*t*
k	: Jumlah variabel prediktor
X_{itk}	: Variabel prediktor ke-*k* di lokasi pengamatan ke-*i* waktu ke-*t*
ε_{it}	: Residual pengamatan ke-*i* pada waktu ke-*t*

2.8.1. Estimasi Parameter Model GWPR

Estimasi parameter dalam model GWPR dilakukan dengan pendekatan "*Weighted Least Square*" (WLS), sama dengan pendekatan dalam model GWR, yakni menambahkan unsur pembobot $w_{it}(u_{it}, v_{it})$ seperti persamaan (2.33). Matriks pembobot dalam model GWPR terdiri dari nilai diagonal dari pembobot yang berasal dari pembobot yang telah diperoleh dan di ulang kembali sejumlah unit waktu untuk memperoleh estimasi parameter. Dengan demikian, model yang terbentuk dapat bervariasi di setiap lokasi.

$$\sum_{t=1}^{T}\sum_{i=1}^{N} w_{it}\,(u_{it},v_{it})\varepsilon_{it}^{2}$$
$$= \sum_{t=1}^{T}\sum_{i=1}^{N} w_{it}\,(u_{it},v_{it})\left[y_{it}-\beta_0(u_{it},v_{it})-\sum_{k=1}^{p}\beta_k(u_{it},v_{it})\,x_{itk}\right]^2$$

$$\sum_{t=1}^{T}\sum_{i=1}^{N} w_{it}\,(u_{it},v_{it})\varepsilon_{it}^{2}$$
$$= \sum_{t=1}^{T}\sum_{i=1}^{N} w_{it}\,(u_{it},v_{it})\big[y_{it}-\beta_0(u_{it},v_{it})-\beta_1(u_{it},v_{it})x_{it1}$$
$$-\beta_p(u_{it},v_{it})x_{itp}\big]^2$$

Penyelesaian dalam bentuk matriks adalah sebagai berikut:

$$\begin{aligned}
\boldsymbol{\varepsilon^T W}(u_{it},v_{it})\boldsymbol{\varepsilon} &= [\boldsymbol{y}-\boldsymbol{X\beta}(u_{it},v_{it})]^T\boldsymbol{W}(u_{it},v_{it})[\boldsymbol{y}-\boldsymbol{X\beta}(u_{it},v_{it})]\\
&= \boldsymbol{y}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}-\boldsymbol{y}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X\beta}(u_{it},v_{it})\\
&\quad -\boldsymbol{\beta}^T(u_{it},v_{it})\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}\\
&\quad +\boldsymbol{\beta}^T(u_{it},v_{it})\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X\beta}(u_{it},v_{it})\\
&= \boldsymbol{y}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}-2\boldsymbol{\beta}^T(u_{it},v_{it})\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}\\
&\quad +\boldsymbol{\beta}^T(u_{it},v_{it})\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X\beta}(u_{it},v_{it}) \qquad (2.34)
\end{aligned}$$

Persamaan (2.34) kemudian diturunkan terhadap $\boldsymbol{\beta}^T(u_{it},v_{it})$

$$\frac{\partial\boldsymbol{\varepsilon}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{\varepsilon}}{\partial\boldsymbol{\beta}^T(u_{it},v_{it})}=0$$

$$\frac{\partial\Big(\boldsymbol{y^T W}(u_{it},v_{it})\boldsymbol{y}-2\boldsymbol{\beta}^T(u_{it},v_{it})\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}+\boldsymbol{\beta}^T(u_{it},v_{it})\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X\beta}(u_{it},v_{it})\Big)}{\partial\boldsymbol{\beta}^T(u_{it},v_{it})}$$
$$=0$$

$$\begin{aligned}
0-2\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}+2\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X}\widehat{\boldsymbol{\beta}}(u_{it},v_{it}) &= 0\\
-2\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}+2\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X}\widehat{\boldsymbol{\beta}}(u_{it},v_{it}) &= 0\\
2\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X}\widehat{\boldsymbol{\beta}}(u_{it},v_{it}) &= 2\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y}\\
\boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{X}\widehat{\boldsymbol{\beta}}(u_{it},v_{it}) &= \boldsymbol{X}^T\boldsymbol{W}(u_{it},v_{it})\boldsymbol{y} \qquad (2.35)
\end{aligned}$$

Kedua ruas dalam persamaan (2.35) dikalikan dengan *invers* dari $\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{X}$ sehingga diperoleh:

$$[\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{X}\widehat{\boldsymbol{\beta}}(u_{it}, v_{it})$$
$$= [\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{y}$$
$$\widehat{\boldsymbol{\beta}}(u_{it}, v_{it}) = [\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{y}$$

Estimator dari parameter model yang didapat sebagai berikut:

$$\widehat{\boldsymbol{\beta}}(u_{it}, v_{it}) = [\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_{it}, v_{it})\boldsymbol{y} \qquad (2.36)$$

dengan: $\widehat{\boldsymbol{\beta}}(u_{it}, v_{it}) = (\hat{\beta}_{it0}, \hat{\beta}_{it1}, \hat{\beta}_{it2}, \dots, \hat{\beta}_{itp})^T$ adalah vektor koefisien regresi lokal, dan matriks $\boldsymbol{W}$ berukuran $nt \times nt$ yaitu sebagai berkut:

$$\boldsymbol{W}(u_{it}, v_{it})$$
$$= diag(w_{11}(u_{it}, v_{it}), w_{21}(u_{it}, v_{it}), w_{31}(u_{it}, v_{it}), \dots, w_{n1}(u_{it}, v_{it}), \dots,)$$
$$\dots$$
$$\dots$$
$$w_{1t}(u_{it}, v_{it}), w_{2t}(u_{it}, v_{it}), \dots, w_{nt}(u_{it}, v_{it})) \qquad (2.37)$$

$\boldsymbol{W}(u_{it}, v_{it})$ dapat ditulis menjadi $\boldsymbol{W}(it)$. Pada notasi matriks, $\boldsymbol{\beta}$ merupakan matriks yang berisi parameter lokal dengan struktur:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0(u_{11}, v_{11}) & \beta_1(u_{11}, v_{11}) & \cdots & \beta_p(u_{11}, v_{11}) \\ \beta_0(u_{21}, v_{21}) & \beta_1(u_{21}, v_{21}) & \cdots & \beta_p(u_{21}, v_{21}) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_0(u_{N1}, v_{N1}) & \beta_1(u_{N1}, v_{N1}) & \cdots & \beta_p(u_{N1}, v_{N1}) \\ \beta_0(u_{12}, v_{12}) & \beta_1(u_{12}, v_{12}) & \cdots & \beta_p(u_{12}, v_{12}) \\ \beta_0(u_{22}, v_{22}) & \beta_1(u_{22}, v_{22}) & \cdots & \beta_p(u_{22}, v_{22}) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_0(u_{N2}, v_{N2}) & \beta_1(u_{N2}, v_{N2}) & \cdots & \beta_p(u_{N2}, v_{N2}) \\ \vdots & \vdots & \vdots & \vdots \\ \beta_0(u_{1T}, v_{1T}) & \beta_1(u_{1T}, v_{1T}) & \cdots & \beta_p(u_{1T}, v_{1T}) \\ \beta_0(u_{2T}, v_{2T}) & \beta_1(u_{2T}, v_{2T}) & \ddots & \beta_p(u_{2T}, v_{2T}) \\ \vdots & \vdots & \cdots & \vdots \\ \beta_0(u_{NT}, v_{NT}) & \beta_1(u_{NT}, v_{NT}) & & \beta_p(u_{NT}, v_{NT}) \end{bmatrix}$$

Estimasi dalam parameter untuk masing-masing baris pada matriks tersebut dapat dituliskan sebagai berikut:

$$\hat{\boldsymbol{\beta}}(it) = [\boldsymbol{X}^T \boldsymbol{W}(it) \boldsymbol{X}]^{-1} \boldsymbol{X}^T \boldsymbol{W}(it) \boldsymbol{y} \qquad (2.38)$$

dengan:

it : Lokasi ke-*i* dan waktu ke-*t* pada matriks β

$\boldsymbol{W}(it)$: Matriks pembobot spasial untuk lokasi pengamatan ke-*i* dan waktu ke-*t*

2.8.2. Pengujian Kecocokan Model GWPR

Pengujian kecocokan model GWPR hampir mirip pengujian model GWR, yang mekibatkan uji kecocokan parameter secara serentak. Menurut Fotheringham, dkk (2002), hipotesis untuk uji kecocokan model sebagai berikut:

H_0: $\beta_k(u_i, v_i) = \beta_k$ untuk setiap $k = 1, 2, \dots, p$ dan $i = 1,2, \dots, n$ (tidak ada perbedaan yang signifikan antara model regresi global dan GWPR)
H_1: minimal terdapat satu $\beta_k(u_i, v_i) \neq \beta_k$ untuk setiap $k = 1, 2, \dots, p$ dan $i = 1,2, \dots, n$ (terdapat perbedaan yang signifikan antara model regresi global dan GWPR)

Statistik uji:

$$F_{hit} = \frac{RSS(H_1)/df_2}{RSS(H_0)/df_1} \tag{2.39}$$

dengan:

$RSS(H_0) = \boldsymbol{y}^T(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{y}$: *Residual Sum of Square* model regresi data panel dimana $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$

$RSS(H_1) = \boldsymbol{y}^T(\boldsymbol{I} - \boldsymbol{L})^T(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{y}$: *Residual Sum of Square* model regresi GWPR

$df_1 = n - p - 1$

$df_2 = \frac{\delta_1^2}{\delta_2}, dimana\ \delta_1 = \boldsymbol{tr}\left(\left[(\boldsymbol{I} - \boldsymbol{L})^T(\boldsymbol{I} - \boldsymbol{L})\right]^i\right), \quad i = 1,2$

I yaitu matriks identitas berukuran $nt \times nt$ serta L yaitu matriks proyeksi dari model GWPR. Misal $x_i^T = (1, x_{i1}, x_{i2}, \dots, x_{ip})$ yaitu baris ke-*i* dari matriks X, dan $\hat{\beta}(i)$ yaitu vektor estimasi parameter di lokasi ke-*i*, maka estimasi nilai y pada lokasi ke-*i* dapat diperoleh sebagai berikut:

$$\hat{\boldsymbol{y}}_t = \boldsymbol{x}_i^T\hat{\boldsymbol{\beta}}(i)$$
$$\hat{\boldsymbol{y}}_t = \boldsymbol{x}_i^T[\boldsymbol{X}^T\boldsymbol{W}(it)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(it)\boldsymbol{y} \tag{2.40}$$

Dimana $\boldsymbol{x}_i^T[\boldsymbol{X}^T\boldsymbol{W}(it)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(it)\boldsymbol{y}$ dinamakan matriks proyeksi, matriks yang memproyeksikan nilai y menjadi $\hat{y}$ pada lokasi ke-*i*.

$$\hat{\boldsymbol{y}} = \boldsymbol{L}\boldsymbol{y} \tag{2.41}$$

$$\boldsymbol{L} = \begin{bmatrix} x_{11}^T(X^TW(11)X)^{-1}X^TW(11) \\ x_{21}^T(X^TW(21)X)^{-1}X^TW(21) \\ \vdots \\ x_{N1}^T(X^TW(N1)X)^{-1}X^TW(N1) \\ x_{12}^T(X^TW(12)X)^{-1}X^TW(12) \\ x_{22}^T(X^TW(22)X)^{-1}X^TW(22) \\ \vdots \\ x_{N2}^T(X^TW(N2)X)^{-1}X^TW(N2) \\ \vdots \\ x_{1T}^T(X^TW(1T)X)^{-1}X^TW(1T) \\ x_{2T}^T(X^TW(2T)X)^{-1}X^TW(2T) \\ \vdots \\ x_{NT}^T(X^TW(NT)X)^{-1}X^TW(NT) \end{bmatrix}$$

Kriteria pengujian tolak H_0 jika nilai $F_{hitung} > F_{\alpha,df_1,df_2}$ atau $p\text{-}value < \alpha$, sehingga dapat disimpulkan terdapat perbedaan yang signifikan antara model regresi global dan GWPR.

2.8.3. Pengujian Signifikansi Parameter Model GWPR

Parameter yang mempunyai pengaruh signifikan terhadap variabel respon di lokasi ke-i dapat dilihat dengan menggunakan uji signifikansi parameter. Berikut ini hipotesis untuk uji signifikansi parameter (Qur'ani, 2014):

H_0: $\beta_k(u_i, v_i) = 0$ untuk $k = 1, 2, \ldots, p$ dan $i = 1,2, \ldots, n$ (tidak terdapat perbedaan pengaruh yang signifikan dari variabel prediktor terhadap model)

H_1: minimal terdapat satu $\beta_k(u_i, v_i) \neq 0$ untuk setiap $k = 1, 2, \ldots, p$ dan $i = 1,2, \ldots, n$ (terdapat perbedaan pengaruh yang signifikan dari variabel prediktor terhadap model)

Estimasi parameter $\boldsymbol{\hat{\beta}}_k(u_i, v_i)$ akan mengikuti distribusi normal dengan rata-rata $\boldsymbol{\beta}_k(u_i, v_i)$ dan matriks varian kovarian $\boldsymbol{C_i C^T \sigma^2}$, dengan $\boldsymbol{C_i} = (\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)$ sehingga didapat:

$$\frac{\hat{\beta}_k(u_i,v_i) - \beta_k(u_i,v_i)}{\sigma\sqrt{\boldsymbol{C_{kk}}}} \sim \mathrm{N}(0,1) \tag{2.42}$$

dengan:

$\boldsymbol{C_{kk}}$: Elemen diagonal ke-*k* dari matriks $\boldsymbol{C_i C^T}$ dan $\hat{\sigma} = \sqrt{\frac{RSS(H_1)}{\delta_1}}$

$RSS(H_0) = \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{y}$: *Residual Sum of Square* model regresi data panel dimana $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$

$RSS(H_1) = \boldsymbol{y^T}(\boldsymbol{I} - \boldsymbol{L})^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{y}$: *Residual Sum of Square* model regresi GWPR

Sehingga statistik uji yang digunakan yaitu:

$$T_{hit} = \frac{\hat{\beta}_k(u_i,v_i)}{\sigma\sqrt{\boldsymbol{C_{kk}}}} \tag{2.43}$$

Kriteria uji yaitu tolak H_0 jika nilai $|T_{hit}| > t_{\frac{\alpha}{2};df}$ ata *p-value* $< \alpha$ dimana $df = \frac{\delta_1^2}{\delta_2}$.

2.9. Pemilihan Model Terbaik

Nilai-nilai koefisien determinasi (R^2) dan AIC dapat digunakan untuk menentukan model terbaik. .

2.9.1. Koefisien Determinasi

Koefisien determinasi (R^2) digunakan untuk menentukan *goodness of fit* garis regresi (Gujarati, 2004). Menurut Putri & Salamah (2013) untuk menentukan seberapa besar pengaruh yang menjelaskan variabel prediktor terhadap respon dapat menggunakan R^2. Menurut Qur'ani (2014) nilai R^2 dalam model *Fixed Effect* GWPR pada matriks pembobot W ukuran $(nT \times nT)$ dan matriks Y berukuran $(nT \times 1)$ dinyatakan sebagai berikut:

$$R^2(u_i, v_i) = \frac{TSS - RSS}{TSS} \tag{2.44}$$

dengan:

TSS: *total sum of squares*

RSS: *residual sum of squares*.

$$TSS = \sum_{j=1}^{N} w_{ij}(y_i - \bar{y}_i)^2$$

$$RSS = \sum_{i=1}^{N} w_{ij}(y_i - \hat{y}_{it})^2$$

dengan:

y_i : Variabel respon lokasi ke-*i*

$\bar{y}_i$: Rata-rata variabel respon

$\hat{y}_{it}$: Nilai prediksi variabel respon

2.9.2. *Akaike Information Criterion* (AIC)

Salah satu kriteria pemilihan model terbaik yaitu menggunakan *Akaike Information Criterion* (AIC) berdasarkan nilai terkecil (Fotheringham, dkk., 2002). Adapun rumus perhitungan nilai AIC yaitu:

$$AIC = 2n\ \ln(\hat{\sigma}) + n\ln(2\pi) + n + tr(L) \qquad (2.45)$$

dengan:

$\hat{\sigma}$: Nilai estimator standar deviasi dari residual, yaitu $\hat{\sigma} = \frac{RSS}{n}$

L : Matriks proyeksi dimana $\hat{y} = Ly$

2.10. Indeks Pembangunan Manusia

Menurut Baeti (2013), Indeks Pembangunan Manusia (IPM) adalah suatu alat ukur yang dapat digunakan untuk menghitung seberapa besar peningkatan kemampuan dasar sumber daya manusia yang dihasilkan dari usaha-usaha peningkatan kapabilitas sumber daya manusia. IPM mengukur tingkat kemajuan pembangunan melalui penilaian aspek pendidikan, kesehatan dan daya beli masyarakat. Perhitungan IPM dilakukan menggunakan rumus sebagai berikut:

$$IPM = \sqrt{I_{kesehatan} \times I_{pendidikan} \times I_{pengeluaran}} \times 100 \qquad (2.46)$$

Menurut BPS (2015) nilai IPM berkisar antara $0 - 100$, Capaian IPM di suatu daerah dapat dikategorikan menjadi 4 kategori, yaitu:

Tabel 2.3. Kategori IPM

Kategori	Nilai IPM
Rendah	$\text{IPM} < 60$
Sedang	$60 \leq \text{IPM} < 70$
Tinggi	$70 \leq \text{IPM} < 80$
Sangat Tinggi	$\text{IPM} \geq 80$

METODOLOGI PENELITIAN

3.1. Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun ajaran 2023/2024 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2. Data Penelitian

Data yang digunakan pada penelitian adalah data sekunder yang didapat dari website https://www.bps.go.id/. Data yang digunakan berupa data Indeks Pembangunan Manusia dari 34 Provinsi di Indonesia tahun 2017-2022. Jenis data yang digunakan yaitu data panel, yang menggabungkan data *cross section* dan *time series*. Data *cross section* berupa data 34 Provinsi Indonesia, sedangkan data *time series* berupa data banyaknya Indeks Pembangunan Manusia dan variabel yang berpengaruh terhadap Indeks Pembangunan Manusia, dengan jumlah data 204 observasi. Tabel 3.1 menyajikan variabel-variabel yang digunakan dalam penelitian ini.

Tabel 3.1. Variabel Penelitian

Variabel		Nama Variabel	Keterangan
Variabel Respon	IPM	Indeks Pembangunan Manusia	Ukuran keberhasilan upaya meningkatkan kemampuan model dasar manusia (BPS, 2022)
Variabel Prediktor	TPT	Tingkat Pengangguran Terbuka	Persentase angkatan pengangguran terhadap total angkatan kerja (BPS, 2022)
	PP	Pengeluaran per Kapita	Konsumsi bulanan seluruh anggota rumah tangga, baik dari pembelian, produksi sendiri dan pemberian yang dibagi dengan jumlah anggota rumah tangga(BPS, 2022).
	JPM	Jumlah Penduduk Miskin	Penduduk yang memiliki rata-rata pengeluaran per kapita per bulan berada dibawah garis kemiskinan (BPS, 2022)
	IKK	Indeks Kemahalan Konstruksi	Indeks harga yang menunjukkan tingkat kemahalan konstruksi suatu kabupaten atau kota dibandingkan dengan kota acuan (BPS, 2022)

Tabel 3. 1. Lanjutan

Variabel		NamaVariabel	Keterangan
Variabel Prediktor	APS	Angka Partisipasi Sekolah	Perbandingan anak sekolah di usia jenjang pendidikan tertentu pada kelompok usia yang sesuai dengan jenjang pendidikan tersebut (BPS, 2022)
	RRLS	Rata-Rata Lama Sekolah	Jumlah tahun yang digunakan oleh penduduk dalam menjalani pendidikan formal (BPS, 2022)
	PSAM	Persediaan Sumber Air Minum	Sumber air yang digunakan untuk minum sehari-hari (BPS, 2022)
	GR	Gini Ratio	Suatu ukuran ketimpangan pengeluaran yang digunakan (BPS, 2023)

3.3. Metode Penelitian

Adapun tahapan analisis yang dilakukan pada penelitian ini yaitu:

1. Melakukan ekplorasi data untuk mengetahui karakteristik data secara umum menggunakan statistik deskriptif dan peta tematik untuk melihat persebaran variabel penelitian di setiap provinsi.
2. Melakukan analisis data panel, seperti model pengaruh gabungan (*common effect model*), model pengaruh tetap (*fixed effect model*), dan model pengaruh acak (*random effect model*).

3. Melakukan uji chow dan uji hausman untuk memilih model panel terbaik antara CEM, FEM, dan REM.
   a. Uji Chow yaitu uji yang digunakan untuk menentukan model terbaik antara model CEM dan FEM dengan hipotesis sebagai berikut:

      H_0: Model CEM

      H_1: Model FEM

      Pengujian hipotesis menggunakan nilai statistik uji Chow didapatkan menggunakan persamaan (2.19). Keputusan statistik ujinya yaitu jika nilai $F_{hitung} > F_{\alpha,N-1,N(T-1)-k}$ atau nilai *p-value* $< \alpha$ maka tolak H_0, dapat disimpulkan model terbaik yaitu model FEM.
   b. Uji Hausman yaitu uji yang digunakan untuk menentukan model terbaik antara model REM dan FEM dengan hipotesis sebagai berikut:

      H_0: Model FEM

      H_1: Model REM

      Pengujian hipotesis menggunakan nilai statistik uji Hausman didapatkan menggunakan persamaan (2.20). Keputusan statistik ujinya yaitu jika nilai $W > {\chi^2}_{(k-1;\alpha)}$ atau nilai *p-value* $< \alpha$ maka tolak H_0, artinya model yang tepat adalah FEM.
4. Melakukan uji asumsi pada hasil pemodelan regresi panel terbaik.
   a. Uji normalitas dilakukan dengan uji *Jarque-Bera*, dan hipotesisnya yaitu:

      H_0: Sisaan menyebar normal

      H_1: Sisaan tidak menyebar normal

      Pengujian hipotesis ini menggunakan uji *Jarque-Bera* yang didapatkan menggunakan persamaan (2.13). Keputusan nilai statistik uji yaitu tolak H_0 jika $JB > \chi^2$ atau nilai *p-value* $< \alpha$ artinya sisaan tidak menyebar secara normal.
   b. Uji multikolinearitas menggunakan nilai *Variance Inflation Factors* (VIF). Nilai VIF > 10 menunjukkan bahwa data terjadi multikolinearitas. Nilai VIF didapatkan dengan perhitungan seperti pada persamaan (2.8).

c. Uji autokorelasi dilakukan dengan uji Run, dan hipotesisnya yaitu:

H_0: Tidak terjadi autokorelasi

H_1: Terjadi autokorelasi

Pengujian hipotesis dilakukan dengan uji Run yang didapatkan menggunakan persamaan (2.11). Keputusan nilai statistik uji yaitu tolak H_0 jika nilai $|Z| > Z_{\frac{\alpha}{2}}$ atau nilai *p-value* $< \alpha$ artinya data terjadi autokorelasi.

5. Melakukan transformasi data penelitian dengan *within transformation*, menggunakan persamaan (2.17).
6. Melakukan uji heterogenitas spasial menggunakaan uji *Breusch Pagan* (BP-Test) pada model terpilih, dengan hipotesisnya yaitu:

   H_0: Tidak terdapat heterogenitas spasial

   H_1: Terdapa heterogenitas spasial

   Pengujian hipotesis ini didapatkan menggunakan persamaan (2.24), dengan keputusan statistik uji yaitu tolak H_0 bila $BP > \chi^2_{(p)}$ atau jika *p-value* $< \alpha$, artinya terdapat heterogenitas spasial.
7. Menghitung jarak *Euclidean* setiap provinsi menggunakan rumus:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$$

8. Menentukan fungsi pembobot spasial terbaik berdasarkan nilai R^2, $adjusted\ R^2$ terbesar, dan AIC terkecil.
9. Menghitung matriks pembobot spasial menggunakan fungsi pembobot spasial terbaik.
10. Mengestimasi nilai parameter pemodelan GWPR mengguanakn pendekatan *Weighted Least SquareI* (WLS).
11. Melakukan uji kecocokan model GWPR untuk menentukan model terbaik antara model regresi data panel dan model GWPR dengan hipotesis sebagai berikut:

    H_0: Tidak terdapat perbedaan yang signifikan antara model regresi global dan GWPR

    H_1: Terdapat perbedaan yang signifikan antara model regresi global dan GWPR

Pengujian hipotesis menggunakan nilai statistik uji F yang didapatkan menggunakan persamaan (2.39). Keputusan statistik uji yaitu tolak H_0 apabila $F_{hitung} > F_{\alpha,df_1,df_2}$ atau p-$value < \alpha$, sehingga dapat disimpulkan terdapat perbedaan yang signifikan antara model regresi global dan model GWPR.

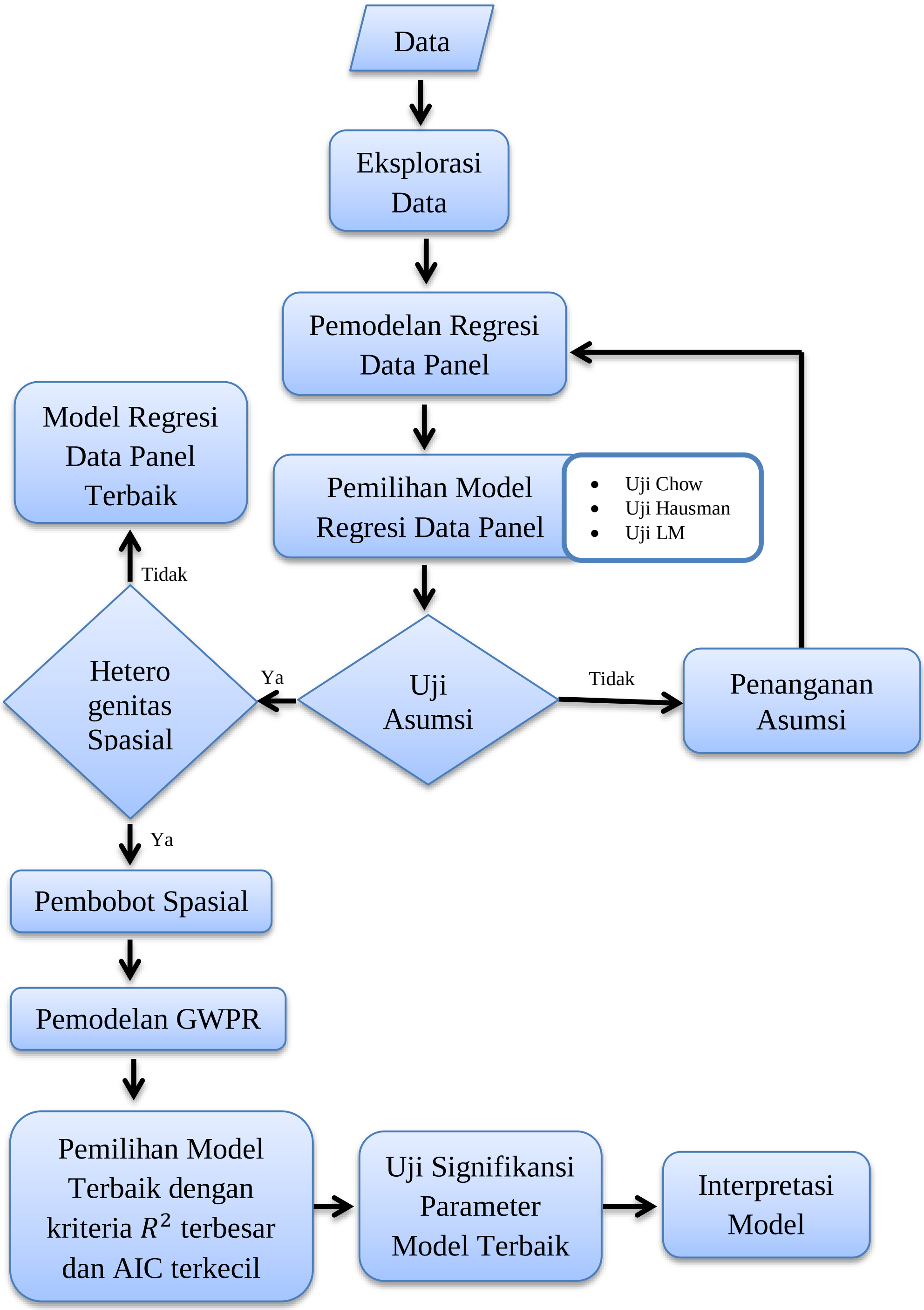
12. Melakukan perbandingan model antara regresi data panel (model regresi global) dan GWPR berdasarkan kriteria R^2 terbesar dan AIC terkecil.
13. Melakukan uji signifikansi parameter model terbaik untuk melihat parameter yang berpengaruh signifikan terhadap variabel respon, dengan hipotesis sebagai berikut:

    H_0: Tidak terdapat perbedaan pengaruh yang signifikan dari variabel prediktor terhadap model

    H_1: Terdapat perbedaan pengaruh yang signifikan dari variabel prediktor terhadap model

    Pengujian hipotesis menggunakan nilai statistik uji t yang didapatkan menggunakan persamaan (2.43) Keputusan statistik uji yaitu tolak H_0 jika nilai $|T_{hit}| > t_{\frac{\alpha}{2};df}$ ata p-$value < \alpha$ artinya minimal terdapat satu $\beta_k(u_i, v_i)$ yang signifikan terhadap model.
14. Interpretasi hasil model yang terpilih

Berikut ini merupakan merupakan diagram alir metode penelitian:



Gambar 3. 1. Diagram Alir Metode Penelitian

KESIMPULAN

Model GWPR dengan fungsi pembobot *adaptive kernel bisquare* merupakan model terbaik untuk memodelkan data IPM di Indonesia, karena memiliki nilai R^2 terbesar dan AIC terkecil dibandingkan model regresi global. Persamaan model dan variabel berpengaruh signifikan yang dihasilkan dalam pemodelan GWPR berbeda untuk setiap provinsi. Pengaruh arah variabel yang signifikan terhadap IPM bervariasi disetiap provinsi sesuai dengan hasil model GWPR yang dihasilkan. Secara keseluruhan, semua variabel prediktor yang digunakan dalam penelitian berpengaruh signifikan terhadap IPM pada taraf signifikansi $\alpha = 0{,}05$. Untuk variabel pengeluaran per kapita dan rata–rata lama sekolah berpengaruh signifikan di seluruh provinsi di Indonesia. Berdasarkan kesamaan variabel yang mempengaruhi IPM di provinsi yang letaknya berdekatan, model GWPR dengan fungsi pembobot *adaptive kernel bisquare* membentuk 8 kelompok.

DAFTAR PUSTAKA

Amaluddin, A., Payapo, R.W., Laitupa, A.A., & Serang, M.R. 2018. A Modified Human Development Index and Poverty in the Villages of West Seram Regency, Maluku Province, Indonesia. *International Journal of Economics and Financial Issues*. **8**(2): 325-330,

Anselin, L. 1988. *Spatial Econometrics*: *Methods and Models*. Kluwer Academic Publisher.

Asmawi & Pangidoan, E. 2021. Pengaruh Angka Harapan Hidup, Rata-Rata Lama Sekolah, Pertumbuhan Ekonomi Dan Pengeluaran Perkapita Terhadap Indeks Pembangunan Manusia di Provinsi Sumatera Utara. *Jurnal Sains Ekonomi*. **2**(1): 96-109.

Astuti, M. 2018. Analisis Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia Di Provinsi Derah Istimewa Yogyakarta Tahun 2010-2016 (Skripsi). Universitas Islam Indonesia Fakultas Ekonomi, Yogyakarta.

Badan Pusat Statistik (BPS). 2015. Indeks Pembangunan Manusia. https://www.bps.go.id/pressrelease/2016/06/15/1278/indeks-pembangunan-manusia-2015.html. Diakses pada tanggal 26 September 2023.

Baltagi, B.H. 2005. *Econometric Analysis of Panel Data*. 3th Edition. John Wiley & Sons Ltd, England.

BPS Republik Indonesia. 2022. *Indeks Pembangunan Manusia*. BPS Republik Indonesia, Jakarta.

BPS Republik Indonesia. 2022. *Indeks Kemahalan Konstruksi Provinsi dan Kabupaten/Kota*. BPS Republik Indonesia, Jakarta.

BPS Republik Indonesia. 2022. *Statistik Perumahan dan Permukiman*. BPS Republik Indonesia, Jakarta.

BPS Republik Indonesia. 2022. Kemiskinan dan Ketimpangan. https://www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html. Diakses pada tanggal 18 Oktober 2023.

BPS Republik Indonesia. 2023. *Tingkat Ketimpangan Pengeluaran Penduduk Indonesia*. BPS Republik Indonesia, Jakarta.

Baeti, N. 2013. Pengaruh Pengangguran, Pertumbuhan Ekonomi dan Pengeluaran Pemerintah Terhadap Pembangunan Manusia Kabupaten/Kota di Provinsi Jawa Tengah Tahun 2007- 2011. *Economics Development Analysis Journal*. **2**(3): 85-98.

Bidanset, P.E. & Lombard, J.R. 2014. The Effect of Kernel and Bandwidth Specification in Geographically Weighted Regression Models on the Accuracy and Uniformity of Mass Real Estate Appraisal. *Journal of Property Tax Assessment & Administration*. **10**(3): 5-14.

Brunsdon, C., Fotheringham, A.S., & Charlton, M. 1996. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*. **28**(4): 281-298.

Cai, R., Yu, D., & Oppenheimer, M. 2014. Estimating the Spatially Varying Responses of Corn Yields to Weather Variations using Geographically Weighted Panel Regression. *Journal of Agricultural and Resource Economics*. **39**(2): 230-252.

Caraka, R.E. & Yasin, H. 2017. *Geographically Weighted Regression* (*GWR*): *Sebuah Pendekatan Regresi Geografis*. Mobius, Yogyakarta.

Croissant, Y. & Millo, G. 2008. Panel Data Econometrics in R: The plm Package. *Journal of Statistical Software*. **27**(2): 1-43.

Fotheringham, A.S., Brunsdon, C., & Charlton, M. 2002. *Geographically Weighted Regression: the analysis of spatially varying relationships*. John Wiley & Sons, England.

Greene, W.H. 2000, *Econometric Analysis*. 4th Edition. Prentice-Hall, Inc., New Jersey.

Gujarati, D. 2004. *Basic Econometrics*. 4th Edition. McGraw–Hill, New York.

Gwarda, K.L. 2018. Geographically Weighted Regression in the Analysis of Unemployment in Poland. *ISPRS International Journal of Geo-Information*. **7**(1): 1-16.

Huang, B., Wu, B., & Barry, M. 2010, Geographically and Temporally Weighted Regression for Modeling Spatio-Temporal Variation in House Prices. *International Journal of Geographical Information Science*. **24**(3): 383-401.

Hsiao, C. 2003. *Analysis of Panel Data*. 2nd Edition. Cambridge University Pres, New York.

Isbiyantoro, K., Wilandari, Y., & Sugito. 2014. Perbandingan Model Pertumbuhan Ekonomi di Jawa Tengah dengan Metode Regresi Linier Berganda dan Metode Geographically Weighted Regression. *Jurnal Gaussian*. **3**(3): 461-469.

Jaya, I.G.N.M., & Sunengsih, N. 2009. Kajian Analisis Regresi dengan Data Panel, hlm. 51-58. Prosiding Seminar Nasional Penelitian Universitas Negeri Yogyakarta, Yogyakarta.

Lesage, J.P. 1999. *The Theory and Practice of Spatial Econometrics*. Department of Economics University of Toledo, Toledo.

Leung, Y., Mei, C., & Zhang, W. 2000, Statistical Tests for Spatial Non Stationarity Based on The Geographically Weighted Regression Model Environment and Planning. *Environment and Planning*. **32**(1): 9-32.

Melliana, A. & Zain, I. 2013. Analisis Statistika Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Timur dengan Menggunakan Regresi Panel. *Jurnal Sains Dan Seni Pomits*. **2**(2): 237-242.

Munikah, T., Pramoedyo, H., & Fitriani, R. 2014. Pemodelan Geographically Weighted Regression dengan Pembobot Fixed Gaussian Kernel pada Data Spasial (Studi Kasus Ketahanan Pangan di Kabupaten Tanah Laut Kalimantan Selatan). *Natural*. **2**(3): 296-302.

Ningrum, J.W., Khairunnisa, A.H., & Huda, N. 2020, Pengaruh Kemiskinan, Tingkat Pengangguran, Pertumbuhan Ekonomi dan Pengeluaran Pemerintah Terhadap Indeks Pembangunan Manusia (IPM) di Indonesia Tahun 2014-2018 dalam Perspektif Islam. *Jurnal Ilmiah Ekonomi Islam*. **6**(2): 212-222.

Pamungkas, R.A., Yasin, H., & Rahmawati, R. 2016. Perbandingan Model GWR Dengan Fixed dan Adaptive Bandwith Untuk Presentase Penduduk Miskin Di Jawa Tengah. *Jurnal Gaussian*. **5**(3): 535-544.

Pangestika, S. 2015. Analisis Estimasi Model Regresi Data Panel dengan Pendekatan Common Effect Model (CEM), Fixed Effect Model (FEM), dan Random Effect Model (REM) (Skripsi). Jurusan Statistika FMIPAUniversitas Negeri Semarang, Semarang.

Park, H.M. 2005. *Linear Regression Models for Panel Data*. Indiana University, Indiana.

Pijnenburg, K. 2013. *Spatial Dependence and Spatial Heterogeneity in the Analysis of Regional Economic Performance and House Price Developments*. Freie Universitat Berlin.

Putri, A. & Salamah, M. 2013. Pemodelan Kasus Balita Gizi Buruk di Kabupaten Bojonegoro dengan Geographically Weighted Regression. *Jurnal Sains dan Pomits*. **2**(1): 106-111.

Qur'ani, A.Y. 2014. Pemodelan Geographically Weighted Regression Panel (GWR-Panel) sebagai Pendekatan Model Geographically Weighted Regression (GWR) dengan Menggunakan Fixed Effect Model Time Trend. *Jurnal Mahasiswa Statistik*. **2**(3): 181-184.

Rahmadeni & Wulandari, N. 2017. Analisis Faktor-Faktor yang Mempengaruhi Inflasi pada Kota Metropolitan di Indonesia dengan Menggunakan Analisis Data Panel. *Jurnal Sains Matematika dan Statistika*. **3**(2): 34-42.

Rencher, A.C. & Schaalje, B.G. 2008. *Linear Models in Statistics*. 2nd Edition. John Willey & Sons, inc., New Jersey.

Susanti, D.S., Lestia, A.S., & Sukmawaty, Y. 2016. Pemodelan Tingkat Kesejahteraan Penduduk Propinsi Kalimantan Selatan dengan Pendekatan Geographically Weighted Regression (GWR), hlm. 184-191. Prosiding Seminar Nasional MIPA UNLAM, Banjarbaru.

Walpole, R.E., Myers, R.H., Myers, S.L., & Ye, K. 2007. *Probability & Statistics for Engineers & Scientists*. 9th Edition. Prentice Hall, New Jersey.

Wang, P. 2006. *Exploring Spatial Effect on Housing Price : the case study of the city of Calgary*. University of Calgary, Canada.

Warsito, B., Yasin, H., Ispriyanti, D., & Hakim, A.R. 2018. The Step Construction of Geographically Weighted Panel Regression in Air Polluter Standard Index (APSI) Data Budi. *E3S Web of Conferences*. **73**(6): 71-74.

Wibisono, D. 2005. *Metode Penelitian dan Analisis Data*. Salemba Medika, Jakarta.

Winarno, W.W. (2015). *Analisis Ekonometrika dan Statistika dengan EViews*. Ed. ke-4. UPP STIM YKPN, Yogyakarta.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press Cambridge, London.

Yolanda, Y. 2017. Analysis of Factors Affecting Inflation and its Impact on Human Development Index and Poverty in Indonesia. *European Research Studies Journal*. **20**(4): 38-56.

Yu, D. 2010, Exploring Spatiotemporally Varying Regressed Relationships: The Geographically Weighted Panel Regression Analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. **38**(2): 134-149.