
METODE REGRESI GULUD UNTUK MENGATASI MASALAH MULTIKOLINEARITAS PADA KASUS INDEKS KUALITAS LINGKUNGAN HIDUP DI INDONESIA TAHUN 2021

Awwaliatul Habibah^{1*}, Fuad Muhajirin Farid², Selvi Annisa³

^{1,2,3}Program Studi Statistika Fakultas MIPA Universitas Lambung Mangkurat
Jl. A. Yani KM. 36, Banjarbaru 70714, Kalimantan Selatan

*e-mail corresponding author: awaliatulhabibah@gmail.com

Abstract

The environmental quality index is an indicator of environmental quality in Indonesia. EQI shows an increasing trend from 2018 to 2022, although Indonesia's environmental quality is ranked 116 out of 180 countries in the world. Therefore, the purpose of this study is to describe the characteristics of the EQI and the factors that can affect EQI, estimate the parameters of the ridge regression model and test the significance of the ridge regression model parameters to overcome the multicollinearity problem in the EQI case. This study uses 4 independent variables, namely population density, traffic, waste, and sanitation in Indonesia in 2021 as secondary data. In this study, ridge regression analysis is used because there is an almost linear relationship between the factors that are thought to affect the EQI. Ridge Regression is a technique that imposes limits (penalties) on parameter estimates in the regression model so that a reduction in the estimated coefficient value can overcome the multicollinearity problem. The results show that the optimal lambda value is 4.737078 when using the 5-fold cross-validation method. Variables that partially affect EQI are population density, waste, and environmental sanitation. The independent variables in this study that can explain the variability of EQI in Indonesia in 2021 is 59.52%, while 40.48% is explained by other variables that are not included in the model. Indonesia's EQI score in 2021 is 71.45 with the predicate "Good".

Keywords: Environmental Quality Index, Ridge Regression, Cross Validation.

1. PENDAHULUAN

Di era globalisasi saat ini, lingkungan menjadi salah satu isu global yang sangat diperhatikan oleh masyarakat. Lingkungan merupakan suatu kawasan serta yang ada didalamnya saling memengaruhi [1]. Oleh karenanya, meningkatkan kualitas lingkungan hidup sangat penting untuk dilakukan agar menjadikan lingkungan yang baik dan sehat. Sedangkan penurunan kualitas lingkungan perlu dicegah agar tidak terjadinya kerusakan dan pencemaran lingkungan hidup. Beberapa akibat dari kerusakan dan pencemaran lingkungan diantaranya yaitu terjadinya banjir, tanah longsor dan munculnya penyakit seperti sesak nafas dan infeksi saluran pernapasan atas (ISPA) [2]. Oleh karena itu, kemerosotan kualitas lingkungan hidup harus ditanggulangi agar kerugian akibat kerusakan dan pencemaran lingkungan.

Indeks kualitas lingkungan hidup atau disingkat dengan IKLH merupakan parameter untuk mengukur kinerja perbaikan kualitas lingkungan di Indonesia [3]. Pada tahun 2018 sampai 2022, nilai IKLH menunjukkan tren peningkatan. Walaupun ada peningkatan nilai IKLH. Menurut laporan *Environmental Performance Index (EPI)*

pada tahun 2020, Indonesia menempati peringkat ke 116 dari 180 negara di dunia dengan memperoleh skor EPI sebesar 37.8. EPI merupakan indikator kualitas lingkungan yang dikembangkan oleh Universitas Yale dan Universitas Columbia untuk memberikan peringkat kepada 180 negara di dunia [4]

IKLH diduga dipengaruhi oleh beberapa faktor yaitu kepadatan penduduk, transportasi, sampah dan sanitasi. Oleh karena itu perlu dilakukan sesuatu analisis untuk memodelkan hubungan antara kepadatan penduduk, transportasi, sampah dan sanitasi terhadap IKLH yaitu analisis regresi. Namun dalam menganalisis regresi ada suatu masalah serius yang mungkin dapat terjadi yaitu multikolinieritas. Diduga antar variabel Kepadatan Penduduk, Transportasi, Sampah dan Sanitasi terdapat korelasi (hubungan yang kuat), sehingga perlu diatasi. Diantara metode analisis yang dapat mengatasi multikolinieritas adalah regresi gulud. Pada tahun 1970, Hoerl dan Kennard memperkenalkan metode analisis yaitu analisis regresi gulud [5].

2. TINJAUAN PUSTAKA

2.1. Indeks Kualitas Lingkungan Hidup (IKLH)

IKLH sebagai ukuran kualitas lingkungan dapat membantu untuk mengembangkan kebijakan dan memfasilitasi komunikasi dengan masyarakat tentang kondisi lingkungan. IKLH meliputi beberapa indikator yaitu indeks kualitas air (IKA), indeks kualitas udara (IKU), indeks kualitas lahan (IKL) dan indeks kualitas air laut (IKAL).

2.2. Kepadatan Penduduk

Perbandingan antara banyaknya populasi dengan luas kawasan yang ditempati disebut kepadatan penduduk [6]. Satuan Jiwa/Km² merupakan satuan dari kepadatan penduduk.

2.3. Transportasi

Menurut Badan Pusat Statistik, salah satu alat transportasi yang penting dalam industri angkutan jalan adalah kendaraan bermotor. Semua jenis kendaraan terdaftar sebagai kendaraan bermotor kecuali kendaraan milik TNI/Polri dan Korps Diplomatik [7]. Satuan data jumlah transportasi yang digunakan dalam satuan unit.

2.4. Sampah

Timbulan sampah adalah jumlah sampah yang dimiliki oleh masyarakat dalam satuan volume dan berat per orang per hari, atau perpanjang bangunan, atau perpanjang jalan [8]. Satuan untuk banyaknya sampah dalam satuan ton.

2.5. Sanitasi

Sanitasi lingkungan adalah upaya untuk mengelola faktor lingkungan fisik manusia yang mungkin dapat menyebabkan sesuatu yang berbahaya bagi kesehatan manusia [9]. Satuan angka sanitasi yang digunakan dalam bentuk persentase.

2.6. Analisis Deskriptif

Statistika deskriptif merupakan suatu analisis dalam yang mengumpulkan, mengolah dan menyajikan serta menginterpretasi data secara kuantitatif yang dapat disajikan dalam bentuk tabel atau grafik baik berupa angka rata-rata, median, minimal dan maksimal [10].

2.7. Analisis Regresi Linier Berganda

Hubungan variabel bebas dan variabel tak bebas dapat dibentuk ke dalam model regresi linear berganda sebagai berikut [5]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

2.8. Uji Asumsi Klasik terhadap Residual

a. Uji Normalitas

Tujuannya untuk mengetahui apakah nilai residual terdistribusi normal atau sebaliknya. Salah satu metode statistik yang uji normalitas yaitu uji *Shapiro Wilk*. Berikut ini rumus yang digunakan dalam pengujian *Shapiro Wilk* [11].

Hipotesis:

H_0 : Nilai residual berdistribusi normal

H_1 : Nilai residual tidak berdistribusi normal

Statistik Uji:

$$W = \frac{(\sum_i (\hat{e}_i - \bar{\hat{e}}) Z_{(i)})^2}{\sum_i (\hat{e}_i - \bar{\hat{e}})^2 \sum_i z_{(i)}^2} \quad (2)$$

Keputusan: H_0 ditolak jika $Sig. < \alpha = 0.05$

b. Uji Heteroskedastisitas

Tujuan dari melakukan uji ini yaitu untuk mengetahui apakah suatu model regresi linear terjadi ketidaksamaan varians dari residual atau tidak. Salah satu metode statistik yang digunakan untuk menguji heteroskedastisitas pada suatu data adalah uji *Glejser*. Berikut ini hipotesa yang digunakan dalam pengujian *Glejser* [12].

Hipotesis:

H_0 : tidak terjadi gejala heteroskedastisitas

H_1 : terjadi gejala heteroskedastisitas

Statistik Uji:

$$|e| = \{X_1, X_2, X_3, X_4, X_5\} \quad (3)$$

Keputusan: H_0 ditolak jika $Sig. < \alpha = 0.05$.

2.9. Uji Multikolinearitas

Uji multikolinieritas dilakukan untuk mengetahui apakah terdapat dua atau lebih variabel bebas dalam suatu model regresi yang saling berkorelasi [13]. Ada beberapa teknik untuk mendeteksi multikolinearitas pada data pengamatan, salah satunya menghitung nilai *Variance Inflation Factors* (VIF). Berikut ini hipotesa yang digunakan dalam menghitung nilai VIF.

Hipotesis:

H_0 : Tidak terjadi multikolinearitas

H_1 : Terjadi Multikolinearitas

VIF dapat dihitung dengan rumus sebagai berikut.

$$VIF_j = \frac{1}{1 - R_j^2}, \quad j = 1, 2, \dots, p \quad (4)$$

Keputusan dalam uji multikolinearitas yaitu tolak H_0 jika $VIF > 10$.

2.10. Regresi Gulud (*Ridge Regression*)

Pada tahun 1962, Hoerl pertama kali menemukan regresi gulud yang digunakan untuk mengendalikan ketidakstabilan penduga Metode Kuadrat Terkecil (MKT). Regresi gulud dan Metode Kuadrat Terkecil memiliki tujuan yang sama yaitu untuk meminimumkan jumlah kuadrat error (JKE) pada pendugaan koefisien regresi. Persamaan JKE adalah sebagai berikut:

$$\begin{aligned} S(\beta_0, \beta_1, \dots, \beta_k) &= \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2 \\ &= \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right)^2 \end{aligned} \quad (5)$$

Perbedaan regresi gulud dengan MKT yaitu regresi gulud menambahkan penalti penyusutan ($\lambda \sum_{j=1}^k \beta_j^2$) dalam meminimumkan JKE. Pendugaan koefisien regresi gulud dilakukan dengan meminimumkan persamaan:

$$\hat{\beta}_{Gulud} = \operatorname{argmin} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right)^2 + \lambda \sum_{j=1}^k \beta_j^2 \right\} \quad (6)$$

Didapatkan nilai $\hat{\beta}_{gulud}$ kedalam persamaan matriks sebagai berikut:

$$\hat{\beta}_{gulud} = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})^{-1}(\mathbf{X}'\mathbf{y}) \quad (7)$$

2.11. Cross Validation

Validasi silang (*cross validation*) adalah metode yang paling sederhana dan paling umum untuk memperkirakan kesalahan. Salah satu jenis validasi silang yaitu validasi silang lipat-k (*K-fold cross validation*).

2.12. Akurasi Model

MSE (*Mean Squared Error*) merupakan metode yang paling umum untuk mengukur kemampuan prediksi model. *Mean Squared Error* (MSE) dihitung dengan mengkuadratkan residual, menjumlahkannya, dan membaginya dengan jumlah sampel [14].

$$MSE = \frac{1}{n} \sum_{i=1}^n \varepsilon^2 \quad (8)$$

2.13. Koefisien Determinasi

Ukuran kebaikan model diperoleh dari nilai koefisien determinasi (R^2) pada data, semakin besar nilai R^2 maka kemampuan model dalam menjelaskan variabel tak bebas akan semakin baik. Perhitungan nilai R^2 yaitu [5].

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

2.14. Ukuran Pemusatan dan Penskalaan

Transformasi data ini umumnya digunakan untuk meningkatkan stabilitas numerik dari beberapa perhitungan [14]. Penskalaan dan pemusatan dilakukan dengan cara mentransformasikan variabel tak bebas dan variabel bebas ke suatu persamaan sebagai berikut [5].

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_{X_j}}, \quad i = 1, 2, 3, \dots, n \text{ dan } j = 1, 2, \dots, k \quad (10)$$

dan

$$Y_i^* = Y_i - \bar{Y}, \quad i = 1, 2, 3, \dots, n \quad (11)$$

2.15. Uji Signifikansi Parameter

Diantaranya adalah uji F yang digunakan untuk mengetahui keseluruhan pengaruh variabel bebas terhadap variabel tak bebas. Berikut ini hipotesa menggunakan uji F [5].

Hipotesis:

H_0 : model regresi tidak signifikan

H_1 : model regresi signifikan

Statistik Uji:

$$F_{hitung} = \frac{\frac{SSR}{k}}{\frac{SSE}{(n - k - 1)}} = \frac{MSR}{MSE} \quad (12)$$

Keputusan: H_0 ditolak jika $F_{hitung} > F_{tabel}$ atau $p_{value} < \alpha$

Selanjutnya ada uji t yang digunakan untuk mengetahui pengaruh pada masing-masing variabel secara parsial (individu), maka hipotesa yang digunakan sebagai berikut [5].

Hipotesis:

H_0 : parameter β_j tidak signifikan

H_1 : parameter β_j signifikan

Statistik Uji:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (13)$$

Keputusan: H_0 ditolak jika $|t_{hitung}| > t_{tabel}$ atau $p_{value} < \alpha$

3. METODE PENELITIAN

Berikut langkah-langkah dalam penelitian ini.

- a) Memperoleh data yang tersedia di situs resmi Kementerian Lingkungan Hidup dan Kehutanan, Badan Pusat Statistik dan Sistem Informasi Pengelolaan Sampah nasional: SIPSN.
- b) Melakukan analisis deskriptif untuk melihat karakteristik data IKLH dan variabel-variabel yang diduga memengaruhi IKLH di provinsi-provinsi yang ada di Indonesia.
- c) Membangun model regresi linier berganda pada kasus IKLH untuk mendapatkan residual (error). Setelah itu, dilakukan uji asumsi klasik pada residual (error) dari model regresi linier berganda yaitu uji normalitas dan uji varians.
- d) Melakukan uji multikolinearitas pada model regresi linier berganda pada masing-masing kasus variabel bebas.
- e) Melakukan pemusatan dan penskalaan (*Centering* dan *Scaling*) pada data dengan mentransformasikan variabel bebas dan tak bebas kedalam persamaan (10) dan persamaan (11). Hasil data tersebut digunakan untuk memilih nilai lambda (λ) yang optimal dengan pertimbangan nilai MSE terkecil. Nilai lambda yang dicobakan dengan teknik validasi silang lipat-5.
- f) Memasukkan nilai lambda yang optimal kedalam Persamaan (7) untuk mengestimasi parameter $\hat{\beta}_{Gulud}$.
- g) Melakukan uji multikolinearitas pada model regresi gulud pada masing-masing kasus variabel bebas.
- h) Mengevaluasi performa model regresi gulud dengan mengukur nilai R^2 dan melakukan uji tingkat signifikansi pada parameter β_{Gulud} untuk mengetahui hubungan parameter dalam model regresi gulud.
- i) Menarik kesimpulan dari hasil penelitian.

4. HASIL DAN PEMBAHASAN

4.1 Statistika Deskriptif

Pada bagian ini akan dibahas tentang deskripsi variabel IKLH yang merupakan variabel tak bebas sedangkan kepadatan penduduk, transportasi, sampah dan sanitasi adalah variabel bebas. Hasil statistik deskriptif dapat dilihat pada Tabel 1 di bawah ini:

Tabel 1 Statistika Deskriptif

Variabel	Min	Median	Mean	Maks
Indeks Kualitas Lingkungan Hidup (Y)	54.43	72.64	72.24	81.80
Kepadatan Penduduk (X_1)	9	103.5	744.26	15978
Transportasi (X_2)	171,936	2,393,585	4,176,252	22,688,339
Sampah (X_3)	19,612	436,842	915,810	5,494,227
Sanitasi (X_4)	40.81	80.67	80.97	97.12

Pada Tabel 1 mengenai deskripsi variabel-variabel diperoleh informasi bahwa provinsi di Indonesia tahun 2021 yaitu Provinsi Papua Barat memperoleh nilai IKLH tertinggi dengan predikat “Baik” dan Provinsi DKI Jakarta memperoleh nilai IKLH terendah dengan predikat “Sedang”. Provinsi DKI Jakarta memperoleh angka kepadatan penduduk tertinggi dan Provinsi Kalimantan Utara memperoleh angka kepadatan penduduk terendah. Provinsi Jawa Timur memperoleh jumlah transportasi tertinggi dan Provinsi Kalimantan Utara memperoleh jumlah transportasi terendah. Provinsi Jawa Tengah memperoleh jumlah sampah tertinggi dan Provinsi Maluku memperoleh jumlah sampah terendah. Sedangkan untuk Provinsi DI Yogyakarta memperoleh angka sanitasi tertinggi dan Provinsi Papua memperoleh angka sanitasi terendah.

4.2 Uji Asumsi Klasik terhadap Residual

Hubungan variabel tak bebas dan bebas dibentuk menjadi model regresi linier berganda. Model regresi linier berganda dalam penelitian ini menggunakan data Kepadatan Penduduk, Transportasi, Sampah dan Sanitasi terhadap IKLH di Indonesia pada tahun 2021 adalah sebagai berikut:

$$\hat{Y} = 87.49 - (87.81 \times 10^{-5})X_1 + (13.01 \times 10^{-8})X_2 - (23.95 \times 10^{-7})X_3 - 0.16X_4$$

Dari hasil model diatas didapatkan hasil nilai residual. Selanjutnya dilakukan pengecekan asumsi menggunakan nilai residual tersebut dengan uji normal dan uji varians.

4.2.1 Uji Normalitas

Dari hasil uji normalitas menggunakan uji *Shapiro-Wilk* didapatkan nilai *p-value* pada Tabel 2 dibawah ini.

Tabel 2 *Shapiro-Wilk normality test*

Data: Residual	
$W = 0.9891$	$p - value = 0.9779$

Hipotesis:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Berdasarkan Tabel 2 hasil uji *Shapiro-Wilk* terlihat bahwa nilai $p - value = 0.9779 > 0.05$, artinya gagal tolak H_0 dengan kata lain nilai residual dari data model regresi linier berganda berdistribusi normal artinya asumsi normalitas terpenuhi.

4.2.2 Uji Heteroskedastisitas

Dari hasil uji varians dengan uji *Glejser* didapatkan nilai signifikansi ditunjukkan pada Tabel 3 berikut.

Tabel 3 Uji *Glejser*

Coefficients:	Estimate	Std.Error	t-value	Sig.
(Intercept)	-0.0264	2.85	-0.009	0.993
Kepadatan Penduduk	(63.85×10^{-6})	(17.44×10^{-5})	0.366	0.717
Transportasi	(-30.14×10^{-8})	(19.25×10^{-8})	-1.566	0.128
Sampah	(10.76×10^{-7})	(75.1×10^{-8})	1.432	0.163
Sanitasi	0.03281	0.0352	0.932	0.359

Hipotesis:

H_0 : tidak terjadi gejala heteroskedastisitas

H_1 : terjadi gejala heteroskedastisitas

Berdasarkan Tabel 3 hasil dari uji *Glejser* terlihat bahwa nilai *Sig.* masing-masing variabel adalah $X_1 = 0.717, X_2 = 0.128, X_3 = 0.163$ dan $X_4 = 0.359 > 0.05$, artinya gagal tolak H_0 dengan kata lain tidak terjadi gejala heteroskedastisitas.

4.3 Uji Multikolinieritas

Dari hasil uji multikolinieritas pada Lampiran 3. didapatkan hasil nilai VIF pada Tabel 4 dibawah ini.

Tabel 4 Uji Multikolinieritas

Variabel	VIF
Kepadatan Penduduk (X_1)	2.023148
Transportasi (X_2)	12.188062
Sampah (X_3)	9.581741
Sanitasi (X_4)	1.096423

Hipotesis:

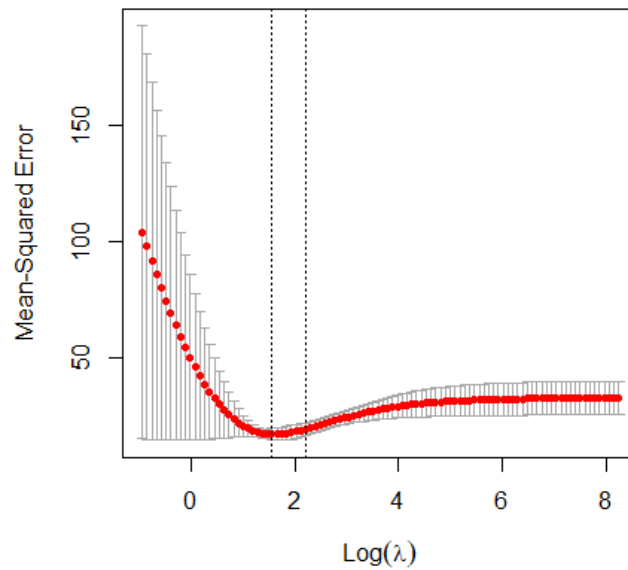
H_0 : Tidak terjadi masalah Multikolinieritas

H_1 : Terjadi masalah multikolinieritas

Berdasarkan Tabel 4 hasil perhitungan nilai VIF masing-masing variabel adalah $X_1 = 2.023148 < 10, X_2 = 12.188062 > 10, X_3 = 9.581741 < 10$ dan $X_4 = 1.096423 < 10$ maka tolak H_0 artinya terjadi gejala multikolinieritas dikarenakan nilai VIF untuk variabel Transportasi (X_2) lebih dari 10.

4.4 Pemilihan Lambda (*Tuning Parameter*) Optimal

Sebelum melakukan pemilihan lambda optimal, perlu dilakukan transformasi data dengan pemusatan dan penskalaan yang digunakan untuk meminimumkan kesalahan dalam pembulatan data. Pemilihan nilai lambda (λ) yang optimal dalam penelitian ini menggunakan validasi silang lipat-5 mempertimbangkan nilai MSE (*Mean Square Error*) terkecil. Dipilih nilai lambda ($\lambda = 4.737078$) yang optimal karena memiliki nilai MSE terkecil dibandingkan dari nilai lambda yang dicobakan lainnya seperti yang diilustrasikan pada Gambar 1 berikut.



Gambar 1 Hasil *Tuning Parameter* Lambda

4.5 Estimasi Parameter Regresi gulud

Setelah didapatkan nilai lambda ($\lambda = 4.737078$), selanjutnya memasukkan nilai lambda yang optimal kedalam Persamaan (7) dan mengestimasi parameter $\hat{\beta}_{gulud}$. Maka didapatkan persamaan regresi gulud adalah sebagai berikut.

$$\hat{Y}_{gulud} = 85.93595 - (66.41 \times 10^{-5})X_1 - (15.25 \times 10^{-9})X_2 - (12.35 \times 10^{-7})X_3 - 0.1412X_4$$

Berdasarkan model regresi gulud diatas dapat disimpulkan bahwa konstanta (*intercept*) berpengaruh positif terhadap variabel IIKLH. Sedangkan parameter (*slope*) dari variabel Kepadatan Penduduk, Transportasi, Sampah dan Sanitasi berpengaruh negatif terhadap variabel IKLH. Adapun penjelasan mengenai interpretasi dari model regresi gulud diatas yaitu untuk konstanta sebesar 85.93595 menunjukkan ketika semua variabel bebas bernilai 0 maka nilai dari IKLH sebesar 85.93595 poin.

Sedangkan dengan asumsi semua variabel bebas lainnya dalam keadaan tetap (konstan), untuk parameter regresi dari variabel kepadatan penduduk sebesar -66.41×10^{-5} menunjukkan setiap kenaikan satu Jiwa/Km² maka IKLH akan menurun sebesar 66.41×10^{-5} poin, untuk parameter regresi dari variabel transportasi sebesar -15.25×10^{-9} menunjukkan setiap kenaikan 1 unit kendaraan bermotor maka IKLH akan menurun sebesar 15.25×10^{-9} poin, untuk parameter regresi dari variabel sampah sebesar -12.35×10^{-7} menunjukkan setiap kenaikan 1 ton timbulan sampah maka IKLH akan menurun sebesar 12.35×10^{-7} poin, dengan asumsi bahwa variabel bebas lainnya tetap atau konstan dan untuk parameter regresi dari variabel Sanitasi sebesar -0.1412 yang menunjukkan setiap kenaikan 1 % sanitasi layak maka IKLH akan turun sebesar 0.1412 poin, dengan asumsi bahwa variabel bebas lainnya konstan.

4.6 Uji Multikolinieritas pada Model Regresi Gulud

Dari hasil uji multikolinieritas pada model regresi gulud didapatkan hasil nilai VIF pada Tabel 5 dibawah ini.

Tabel 5 Uji Multikolinieritas pada Model Regresi Gulud

Variabel	VIF
Kepadatan Penduduk (X_1)	0.0277
Transportasi (X_2)	0.02727
Sampah (X_3)	0.02706
Sanitasi (X_4)	0.02428

Hipotesis:

H_0 : Tidak terjadi multikolinieritas

H_1 : Terjadi Multikolinieritas

Berdasarkan Tabel 5 perhitungan nilai VIF masing-masing variabel adalah $X_1 = 0.0277 < 10$, $X_2 = 0.02727 < 10$, $X_3 = 0.02706 < 10$ dan $X_4 = 0.02428 < 10$ maka gagal tolak H_0 artinya tidak terjadi gejala multikolinieritas pada model regresi gulud dikarenakan nilai VIF semua variabel bebas kurang dari 10. Disimpulkan bahwa metode analisis regresi gulud dapat mengatasi multikolinieritas pada kasus IKLH di Indonesia pada tahun 2021.

4.7 Koefisien Determinasi

Berdasarkan hasil perhitungan nilai R^2 diperoleh nilai sebesar 0.5952 atau 59.52% yang berarti bahwa variabel bebas dalam penelitian ini yaitu variabel kepadatan penduduk, transportasi, sampah dan sanitasi dapat menjelaskan variasi IKLH di Indonesia tahun 2021 sebesar 59.52%, sedangkan sisanya sebesar 40.48% dijelaskan oleh variabel lain yang tidak dimasukkan dalam model.

4.8 Uji Signifikansi Parameter

4.8.1 Uji F

Hasil pengujian signifikansi parameter secara keseluruhan ditunjukkan pada Tabel 6 adalah sebagai berikut:

Tabel 6 Uji Signifikansi secara Keseluruhan (Uji F)

F_{tabel}	F_{hitung}	p_{value}	Kesimpulan
2.701	16.43914	19.61×10^{-7}	Signifikan

Berdasarkan Tabel 6 dapat diketahui bahwa tolak H_0 karena $F_{hitung} > F_{tabel}$ dan $p_{value} < \alpha$, maka disimpulkan variabel kepadatan penduduk, transportasi, sampah dan sanitasi secara keseluruhan signifikan didalam model regresi gulud artinya variabel kepadatan penduduk, transportasi, sampah dan sanitasi secara keseluruhan berpengaruh terhadap IKLH di Indonesia tahun 2021.

4.8.2 Uji t

Hasil pengujian signifikansi parameter secara parsial ditunjukkan pada Tabel 7 adalah sebagai berikut:

Tabel 7 Uji Signifikansi secara Parsial (Uji t)

Variabel	Koefisien	$ t_{hitung} $	p_{value}	Kesimpulan
Kepadatan Penduduk	-66.41×10^{-5}	3.3199	0.0024	Signifikan
Transportasi	-15.25×10^{-9}	1.7088	0.0978	Tidak Signifikan
Sampah	-12.35×10^{-7}	3.158	0.0036	Signifikan
Sanitasi	-0.1412	2.7508	0.0100	Signifikan

Berdasarkan Tabel 7 dapat dilihat bahwa tolak H_0 karena $p_{value} < \alpha = 0.05$ dan $t_{hitung} > t_{tabel}$ dengan nilai t_{tabel} sebesar 1.694 maka disimpulkan bahwa variabel kepadatan penduduk, sampah dan sanitasi secara parsial signifikan terhadap IKLH artinya variabel kepadatan penduduk, sampah dan sanitasi berpengaruh signifikan secara parsial terhadap IKLH di Indonesia pada tahun 2021. Sedangkan variabel transportasi tidak berpengaruh secara signifikan terhadap IKLH di Indonesia pada tahun 2021.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan pada bab sebelumnya, dapat disimpulkan bahwa provinsi di Indonesia tahun 2021 yaitu Provinsi Papua Barat memperoleh nilai IKLH tertinggi dan Provinsi DKI Jakarta memperoleh nilai IKLH terendah. Provinsi DKI Jakarta memperoleh angka kepadatan penduduk tertinggi dan Provinsi Kalimantan Utara memperoleh angka kepadatan penduduk terendah. Provinsi Jawa Timur memperoleh jumlah transportasi tertinggi dan Provinsi Kalimantan Utara memperoleh jumlah transportasi terendah. Provinsi Jawa Tengah memperoleh jumlah sampah tertinggi dan Provinsi Maluku memperoleh jumlah sampah terendah. Sedangkan untuk Provinsi DI Yogyakarta memperoleh angka sanitasi tertinggi dan Provinsi Papua memperoleh angka sanitasi terendah. Nilai IKLH Indonesia pada tahun 2021 mendapatkan predikat "Baik" dengan nilai IKLH sebesar 71.45 poin. Didapatkan model regresi gulud dalam kasus ini adalah:

$$\hat{Y}_{gulud} = 85.93595 - (66.41 \times 10^{-5})X_1 - (15.25 \times 10^{-9})X_2 - (12.35 \times 10^{-7})X_3 - 0.1412X_4$$

Diperoleh juga nilai koefisien determinasi sebesar 59.52% artinya variabel bebas dalam penelitian ini dapat menjelaskan variabilitas IKLH di Indonesia pada tahun 2021 sebesar 59.52% dan 40.48% dijelaskan oleh variabel lain yang tidak dimasukkan dalam model. Serta hasil uji tingkat signifikansi dengan uji F dan uji t adalah variabel kepadatan penduduk, transportasi, sampah dan sanitasi juga dimasukkan ke dalam model regresi gulud secara keseluruhan berpengaruh terhadap IKLH di Indonesia tahun 2021, sedangkan variabel kepadatan penduduk, sampah dan sanitasi berpengaruh signifikan secara parsial terhadap IKLH di Indonesia pada tahun 2021 hanya variabel transportasi yang tidak berdampak parsial terhadap IKLH di Indonesia pada tahun 2021.

DAFTAR PUSTAKA

- [1] D. Djamin, *Pengawasan dan Pelaksanaan Undang-undang Lingkungan Hidup*. Yayasan Obor Indonesia, 2007.

- [2] Dinas Lingkungan Hidup Kota Semarang, "5 Bencana Kota Semarang Akibat Kerusakan Lingkungan," *Dinas Lingkungan Hidup Kota Semarang*, 2020. <https://dlh.semarangkota.go.id/5-bencana-kota-semarang-akibat-kerusakan-lingkungan/>
- [3] Kementerian Lingkungan Hidup dan Kehutanan, *Indeks Kualitas Lingkungan Hidup 2019*. Jakarta, 2020.
- [4] Z. A. Wendling, J. W. Emerson, A. de Sherbinin, D. C. Esty, and E. Al., *2020 Environmental Performance Index*. New Haven, CT: Yale Center for Environmental Law & Policy, 2020.
- [5] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, Fifth. New Jersey, 2012.
- [6] I. B. Mantra, *Demografi Umum*. Yogyakarta: Pustaka Pelajar, 2000.
- [7] Badan Pusat Statistik, *Statistik Transportasi Darat 2021 Land Transportation Statistics 2021*. Jakarta: Badan Pusat Statistik, 2022.
- [8] Badan Standardisasi Nasional, "SNI 19-2454-2002 :Tata Cara Teknik Operasional Pengelolaan Sampah Perkotaan." pp. 1-31, 2002.
- [9] Kesehatan Lingkungan Politeknik Kesehatan Makassar, "Sanitasi," *Kesehatan Lingkungan Potekkes Makassar*, 2016. <https://kesling.poltekkes-mks.ac.id/271/>
- [10] R. E. Walpole, R. H. Myers, S. L. Myres, and K. Ye, *Probability & Statistics for Engineers & Scientists*, Ninth. United States: Pearson Education, Inc., 2012.
- [11] E. Suárez, C. M. Pérez, R. Rivera, and M. N. Martínez, *Applications of Regression Models in Epidemiology*, vol. 6, no. August. New Jersey: John Wiley & Sons, Inc, 2017.
- [12] C. Apriliantyza, "Pengaruh Content Marketing dan Brand Love terhadap Keputusan Pembelian Produk Scarlett Whitening Melalui Media Sosial Tiktok," Sekolah Tinggi Ilmu Ekonomi (STIE) PGRI Dewantara Jombang, 2022. [Online]. Available: <http://repository.stiedewantara.ac.id/id/eprint/3668>
- [13] W. Mendenhall and T. Sincich, *A Second Course in Statistics Regression Analysis*, Seventh. United States, 2012.
- [14] M. Kuhn and K. Johnson, *Applied predictive modeling*. New York: Springer, 2013. doi: 10.1007/978-1-4614-6849-3.