

# **Application of Robust Regression and Weighted Least Squares in Handling Outlier and Heteroscedasticity in Modeling Rice Production in West Java in 2023**

**Penerapan Regresi Kekar dan Kuadrat Terkecil Terboboti dalam Penanganan Pencilan dan Heteroskedastisitas Pada Pemodelan Produksi Padi di Jawa Barat Tahun 2023**

**Muhammad Jodi At-Takbir<sup>1</sup>, Siti Arbaynah<sup>1</sup>, Fauzan Fajari<sup>1</sup>,  
Muhammad Fahrezi Maulana<sup>1</sup>, Sachnaz Desta Oktarina<sup>1</sup>, Akbar Rizki<sup>1</sup>, Akmal Riza Wibisono<sup>1</sup>,**

Department of Statistics, IPB University, Indonesia

corresponding author: jodiattakbir@apps.ipb.ac.id

## **ABSTRAK**

Produksi padi berperan penting dalam menjaga ketahanan pangan di beberapa negara asia, salah satunya indonesia. Mengingat pentingnya ketahanan pangan dan ketergantungan Indonesia pada impor beras, proyeksi produksi padi yang akurat menjadi sangat krusial. Jawa Barat menjadi salah satu provinsi di Indonesia yang mengalami penurunan produksi padi pada tahun 2023, sehingga menjadi fokus penelitian. Penelitian ini bertujuan untuk mengevaluasi model terbaik berdasarkan peubah-peubah yang memengaruhi nilai produksi padi. Metode analisis yang digunakan adalah regresi robust dan metode Weighted Least Squares (WLS) untuk mengatasi masalah outlier dan heteroskedastisitas dalam pemodelan produksi padi. Faktor-faktor yang diteliti meliputi luas tanam (ha), persentase jumlah petani (%), jumlah pupuk urea (ton/ha), jumlah pupuk NPK (ton/ha), jumlah usaha tani padi (unit), jumlah teknologi (unit), dan harga beras (Rp/kg). Data yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik (BPS) dan Open Data Jawa Barat. Hasil analisis menunjukkan bahwa luas tanam, persentase jumlah petani, jumlah teknologi berpengaruh signifikan terhadap produksi padi pada tahun 2023. Dengan model akhir  $Y = -456.433 + 5.347X_1 - 24500X_2 + 895.5X_6$  dan nilai *adjusted R squared* 0,9277

**Kata kunci :** ketahanan pangan, produksi padi, heteroskedastisitas, pencilan, regresi kekar

## **1. Pendahuluan**

### **1.1. Latar Belakang**

Produksi merupakan suatu kegiatan ekonomi yang bertujuan untuk menciptakan atau menghasilkan barang atau jasa dengan tujuan untuk memenuhi kebutuhan individu atau organisasi (Gunawan, F. 2018). Proses produksi ini dilakukan dengan cara menambahkan nilai guna terhadap suatu barang atau jasa melalui berbagai tahapan, mulai dari pengolahan bahan baku hingga menjadi produk jadi yang siap untuk digunakan atau dikonsumsi.

Dalam konteks usahatani padi, produksi memegang peran yang sangat penting sebagai fondasi utama dalam mencapai tujuan petani untuk memenuhi kebutuhan mereka akan hasil panen. Namun, harapan petani akan tingginya produksi seringkali tidak selalu sejalan dengan kenyataan lapangan. Faktor-faktor seperti kondisi cuaca yang tidak terduga, serangan hama dan penyakit tanaman, serta kendala-kendala lainnya dapat menghambat produktivitas dan menyebabkan hasil produksi tidak sesuai dengan yang diinginkan.

Menurut data dari Badan Pusat Statistik Provinsi Jawa Barat (2023), produksi padi di Jawa Barat mengalami penurunan sebesar 3,11% pada tahun 2023. Penurunan ini merupakan fenomena yang memerlukan analisis mendalam mengingat pentingnya padi sebagai komoditas strategis dalam ketahanan pangan nasional.

Faktor-faktor seperti luas tanam, jumlah petani, penggunaan pupuk urea, keberadaan usaha tani, penerapan teknologi pertanian, fluktuasi harga beras, ketersediaan beras, serta data produksi beras dari setiap kota dan kabupaten di Jawa Barat, semuanya memiliki potensi pengaruh terhadap hasil produksi padi (Badan Pusat Statistik Provinsi Jawa Barat, 2023). Penelitian ini bertujuan untuk menganalisis peubah-peubah tersebut guna mengidentifikasi faktor-faktor yang paling signifikan dalam mempengaruhi produksi padi di Jawa Barat. Dengan demikian, dapat diambil langkah-langkah strategis untuk meningkatkan produksi padi di masa yang akan datang.

Dengan mengidentifikasi faktor-faktor yang paling signifikan, dapat dirumuskan strategi dan kebijakan yang tepat sasaran untuk meningkatkan produktivitas sektor pertanian padi di Jawa Barat. Hal ini penting untuk menjamin ketahanan pangan di masa depan dan mencegah terjadinya krisis pangan akibat penurunan produksi padi yang berkelanjutan.

### **1.2. Tujuan Penelitian**

- a. Mengetahui peubah-peubah yang memengaruhi produksi padi secara signifikan di Jawa Barat Tahun 2023
- b. Mencari model terbaik dalam pemodelan produksi padi di Jawa Barat Tahun 2023 dengan metode penanganan Robust Regression dan Weighted Least squares (WLS)

## **2. Metodologi**

### **2.1. Bahan dan Data**

Penelitian ini menggunakan data sekunder yang berasal dari Badan Pusat Statistik (BPS) dan Open Data Jawa Barat terhadap faktor-faktor yang memengaruhi produksi padi. Data terdiri dari 27 amatan terdiri dari kabupaten dan kota di Jawa

Barat. Peubah yang digunakan berjumlah delapan terdiri dari satu peubah respon dan tujuh peubah penjelas yang disajikan pada Tabel 1.

Tabel 1: Peubah yang digunakan

Keterangan	Satuan	Sumber Data
Produksi Padi (Y)	Ton	BPS Jawa Barat
Luas Lahan Padi (X1)	Hektar	Open Data JABAR
Persentase Jumlah Petani (X2)	Persen	BPS Jawa Barat
Jumlah Pupuk Urea (X3)	Ton/Hektar	Open Data JABAR
Jumlah Pupuk NPK (X4)	Ton/Hektar	Open Data JABAR
Jumlah Wirausaha Petani (X5)	Unit	BPS Jawa Barat
Jumlah Teknologi Pertanian (X6)	Unit	BPS Jawa Barat
Harga Beras (X7)	Rp/Kilogram	BPS Jawa Barat

Peubah bebas pertama, yaitu luas lahan (X1) dalam konteks ini mengacu pada ukuran bidang geografis yang digunakan dalam praktek budidaya pertanian, menjadi panggung bagi berbagai tahap proses produksi yang terjalin dalam struktur sosial dan ekonomi masyarakat agraris (Haryanto 2014 dalam Barkah dan Masdari 2020).

Peubah bebas berikutnya adalah petani (X2) merupakan individu yang menggantungkan kehidupannya pada kegiatan pertanian sebagai sumber penghasilan, dengan mencakup berbagai jenis usaha seperti pertanian tanaman pangan, hortikultura, perkebunan, peternakan, dan perikanan (Sahri *et al.*, 2022).

Pupuk merupakan komponen esensial dalam pertumbuhan dan keberlangsungan hidup tanaman. Secara substansial, pupuk menyediakan unsur hara yang krusial bagi tanaman. Pendekatan dalam pemupukan melibatkan penggunaan dua kategori utama, yakni pupuk sintetis dan organik (Raksun 2019). Peubah bebas lainnya yaitu pupuk sintetis, seperti urea (X3) dan NPK (X4), umumnya digunakan oleh petani untuk memperkaya tanaman dengan unsur hara yang spesifik.

Peubah bebas berikutnya yaitu usahatani (X5) merupakan suatu domain di mana individu atau kelompok mengelola faktor-faktor produksi seperti sumber daya alam, tenaga kerja, modal, dan keterampilan, dengan fokus pada kegiatan produksi di bidang pertanian. Ini melibatkan perencanaan, penerapan teknik pertanian, manajemen risiko, dan interaksi dengan dinamika lingkungan dan pasar. Usahatani juga mencakup aspek sosial, ekonomi, dan teknologi, serta membutuhkan kolaborasi yang kuat antara berbagai pemangku kepentingan untuk mencapai tujuan produksi yang berkelanjutan (Shinta 2011 dalam Amili *et al.*, 2020).

Peubah bebas berikutnya adalah penggunaan teknologi pertanian modern (X6), seperti sistem pemantauan dan pengendalian berbasis sensor, penggunaan drone untuk pemetaan lahan, serta penerapan mesin-mesin pertanian yang efisien, memiliki potensi signifikan dalam meningkatkan produktivitas tanaman padi dan mengoptimalkan pemanfaatan. Sistem pemantauan berbasis sensor memungkinkan pengumpulan data secara real-time terkait kondisi tanah, kelembaban, dan

kebutuhan nutrisi tanaman, sehingga memungkinkan intervensi yang lebih tepat waktu dan akurat (Siregar, 2023).

## 2.2. Metode Analisis

Tahapan-tahapan analisis data yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan eksplorasi terhadap data dengan memeriksa distribusi variabel respon dan variabel penjelas.
2. Membuat model regresi awal untuk memahami hubungan antara variabel penjelas dan respon. Selain itu, peneliti juga melakukan pendeteksian multikolinieritas untuk memastikan tidak ada korelasi yang kuat antara variabel penjelas.
3. Menentukan Model Regresi Terbaik

Model regresi terbaik (*best subset regression*) adalah pendekatan yang mempertimbangkan semua kombinasi variabel penjelas dan memilih model terbaik berdasarkan kriteria statistik. Salah satu kriterianya adalah *Akaike Information Criteria* (AIC) (Kassambara 2018). AIC mengukur kualitas relatif dari model statistika dan membantu untuk memilih model terbaik dari berbagai kombinasi model lainnya (Emawati dan Somayasa 2023). AIC didasarkan pada metode maximum likelihood estimation.

$$AIC = n \log(\hat{\sigma})^2 + n \log[(n - p - 1)/n] + 2p + 4 \quad (3)$$

Keterangan:

$\hat{\sigma}$  = galat baku

$n$  = jumlah observasi

$p$  = jumlah parameter

Adapun formulasi AIC yang digunakan untuk membandingkan model satu dengan lainnya dengan menggunakan rumus sebagai berikut:

$$AIC_c = AIC + \frac{2(p+2)(p+3)}{(n-p-3)} \quad (4)$$

Menurut Fathurrahman(2016), model regresi terbaik memiliki nilai AIC terkecil.

4. Melakukan pemeriksaan data

### a. Pencilan(outlier)

Menurut Montgomery and Pack (1992), pencilan adalah suatu pengamatan yang ekstrim. Pencilan (outlier) adalah pengamatan yang memiliki nilai yang jauh lebih besar atau lebih kecil daripada nilai-nilai lainnya. Untuk mendeteksi pencilan, peneliti menggunakan rumus berikut:

$$r_i = \frac{e_i}{s\sqrt{1-h_{ii}}} \quad (1)$$

Dengan  $e_i$  merupakan residual ke- $i$ ,  $s$  merupakan standar error sisaan,  $h_{ii}$  adalah elemen-elemen diagonal dari matriks  $H_{n \times n} = X(X'X)^{-1}X'$ . Amatan dapat dikatakan pecilan jika nilai  $|r_i| > 2$  (Chatterjee dan Simonoff 2013).

#### b. Titik Leverage

Titik Leverage adalah pengamatan yang memiliki nilai ( $x$ ) yang jauh dari nilai-nilai ( $x$ ) lainnya. Dalam analisis regresi, suatu pengamatan memiliki leverage yang tinggi jika nilainya pada variabel penjelas ( $x$ ) dianggap “ekstrim”. Untuk mendeteksi titik leverage, kita dapat menggunakan rumus berikut:

$$h_{ii} = x_i'(X'X)^{-1}x_i \quad (2)$$

Dengan  $h_{ii}$  merupakan unsur diagonal ke- $i$  pada matriks  $H$ , dan  $x_i$  adalah vektor baris ke- $i$  pada matriks  $X$ . Jika nilai unsur diagonal  $h_{ii} > \frac{2p}{n}$ , maka observasi tersebut dianggap sebagai titik leverage dalam analisis dengan  $p$  banyak parameter dan  $n$  banyaknya observasi (Kannan dan Manoj 2015).

#### c. Amatan Berpengaruh

Amatan berpengaruh adalah pengamatan yang dapat menarik model regresi menuju arahnya dan memengaruhi koefien model. Penghapusan suatu pengamatan dapat menyebabkan perubahan besar dalam model regresi. Untuk menganalisis amatan berpengaruh, dapat menggunakan metode seperti DFBETAS, DFFITS, COVRATIO, atau *Cook's Distance* (Chatterjee dan Simonoff 2013).

### 5. Penerapan *Robust Regression*

Regresi Kekar (*Robust Regression*) merupakan metode yang digunakan untuk menangani masalah pencilan pada sebuah data. Dalam regresi kekar ini digunakan estimator-M pembobot Huber untuk meminimumkan *Mean Squared error* (MSE) dari Metode Kuadrat Terkecil sehingga menghasilkan model yang tidak sensitif terhadap pencilan (Rather *et al.*, 2022).

### 6. Uji Asumsi Model Regresi Kekar

#### a. Nilai Harapan Sisaan sama dengan Nol

Dalam memeriksa asumsi ini digunakan *t-test* dengan hipotesis alternatif sebagai berikut.

$$H_1 = \text{Nilai Harapan Sisaan Tidak sama dengan Nol}$$

Keputusan yang diambil berdasarkan *p-value* dengan taraf nyata sebesar 5%.

## b. Autokorelasi

Asumsi ini diperiksa dengan *Durbin-Watson test* dengan hipotesis alternative sebagai berikut.

$$H_1 = \text{Terjadi Autokorelasi}$$

Keputusan yang diambil berdasarkan *p-value* dengan taraf nyata sebesar 5%.

## c. Homoskedastisitas

Dalam memeriksa asumsi ini digunakan *Breusch-Pagan test* dengan hipotesis alternatif sebagai berikut.

$$H_1 = \text{Terjadi Heteroskedastisitas}$$

Keputusan yang diambil berdasarkan *p-value* dengan taraf nyata sebesar 5%.

## d. Normalitas (Sisaan Menyebar Normal)

Asumsi ini diperiksa dengan *Shapiro-Wilk Normality test* dengan hipotesis alternatif sebagai berikut (Mardiatmoko 2020).

$$H_1 = \text{Sisaan Tidak Menyebar Normal}$$

Keputusan yang diambil berdasarkan *p-value* dengan taraf nyata sebesar 5%.

## 7. Penerapan WLS dalam Penanganan Heteroskedastisitas

Dalam analisis regresi, metode Ordinary Least Square (OLS) digunakan untuk mendapat taksiran bersifat Best Linear Unbiased Estimator (BLUE) (Rihaksa et al. 2023). Dalam metode tersebut, terdapat asumsi yang harus dipenuhi oleh OLS, salah satunya adalah asumsi homoskedastisitas (Setyaningsih dan Noeryanti 2019). Metode WLS memiliki prinsip yang sama dengan OLS, perbedaan mendasar antara keduanya terletak pada penambahan variabel baru (*w*) pada metode WLS. Estimasi parameter  $\beta_j$  ditentukan oleh formulasi sebagai berikut (Fox 2016):

$$\hat{\beta} = (XTWX)^{-1}XTWY \quad (11)$$

Dengan *X* adalah matriks prediktor berukuran  $n \times p$  dengan  $n$  jumlah observasi dan  $p$  banyaknya parameter. *W* adalah matriks bobot diagonal berukuran  $n \times n$ , dimana setiap elemen diagonalnya ( $W_{ii}$ ) merupakan nilai bobot pada observasi ke- $i$  dan *Y* adalah vektor variabel respons berukuran  $n \times 1$ .

## 8. Uji Asumsi Model Regresi Kekar yang Sudah Diterapkan WLS

## 9. Interpretasi Model Akhir

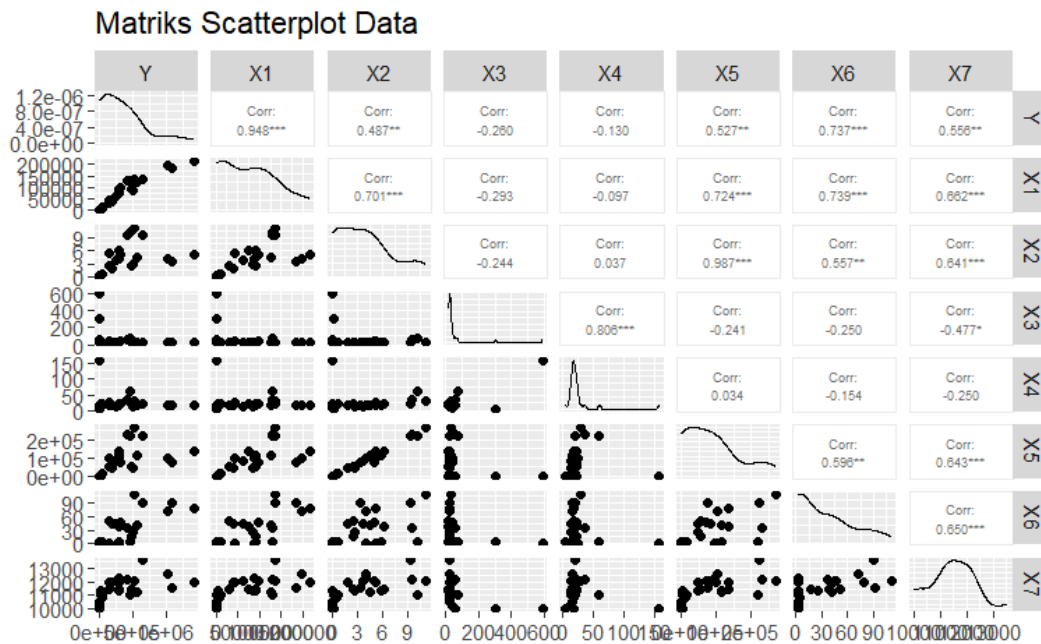
## 3. Hasil dan Pembahasan

## 3.1. Eksplorasi Data

Tabel 2: Ringkasan Data Produksi Padi Jawa Barat Tahun 2023

Minimum	Q1	Median	Q3	Maksimum
33	16229	281153	507361	1419736

Berdasarkan Tabel 2, dapat diketahui bahwa produksi padi minimum di Jawa Barat pada Tahun 2023 terjadi di Kota Depok, yaitu sebesar 33 Ton, sedangkan produksi padi maksimum terjadi di Kabupaten Indramayu, yaitu lebih dari 1 juta Ton. Selain itu, dapat diketahui pula bahwa sebaran produksi padi menyebar dari 16 ribu Ton hingga lebih dari 500 ribu Ton dengan nilai tengah berkisar di 280 ribu Ton.



Gambar 1: Matriks Korelasi Antar Peubah

Berdasarkan Gambar 1, dapat terlihat matriks korelasi yang menunjukkan besar koefisien antarpeubah. Terlihat bahwa terdapat korelasi terkuat antarpeubah, yaitu antara X2 (Persentase Jumlah Petani) dan X5 (Jumlah Wirausaha Petani) sebesar 0.987. Sedangkan yang memiliki nilai koefisien korelasi terlemah yaitu antara X3 (Jumlah Pupuk Urea) dan X7 (Harga Beras) yaitu sebesar -0.477. Selain itu, terlihat bahwa setiap peubah X memiliki korelasi yang kuat terhadap peubah Y, kecuali pada peubah X3 dan X4.

### 3.2. Pemodelan Regresi Awal dan Deteksi Multikolinearitas

Pemodelan data sementara atau sebelum dilakukan pemeriksaan lebih lanjut dapat dituliskan dengan model sebagai berikut.

$$\hat{Y} = 264812 + 6.21X_1 - 61740X_2 - 66.06X_3 + 164.3X_4 + 1.13X_5 + 1296X_6 - 25.14X_7$$

Tabel 3: Nilai VIF Masing-Masing Peubah

X1	X2	X3	X4	X5	X6	X7
3.147	43.855	4.071	3.563	46.397	2.751	2.693

Peubah penjelas dideteksi sebagai multikolinearitas jika nilai VIF > 10. Berdasarkan Tabel 3, terdapat dua peubah yang memenuhi syarat multikolinearitas, yaitu X2 dan X5. Salah satu metode yang digunakan untuk menangani multikolinearitas adalah menghapus peubah yang memiliki nilai VIF > 10 terbesar. Dapat terlihat bahwa X5 memiliki nilai VIF > 10 terbesar, sehingga dapat dilakukan penghapusan pada peubah tersebut.

### 3.3. Best Subset Regression

Tabel 4: Best Subset Regression

No	Peubah Penjelas	$R^2$	$Adj R^2$	AIC	C(p)
1	X1	0.8589	0.8448	711.802	39.1099
2	X1 X2	0.9307	0.9274	688.316	3.1667
3	<b>X1 X2 X6</b>	<b>0.9226</b>	<b>0.9158</b>	<b>685.199</b>	<b>1.7487</b>
4	X1 X2 X5 X6	0.9244	0.9109	686.913	2.7839
5	X1 X2 X5 X6 X7	0.9278	0.9116	687.929	4.0754
6	X1 X2 X3 X5 X6 X7	0.9182	0.9097	689.860	6.0271
7	X1 X2 X3 X4 X5 X6 X7	0.9207	0.9177	691.822	8.0000

*Best Subset Regression* merupakan salah satu metode yang digunakan untuk menduga atau memilih model terbaik. Metode ini akan menyeleksi peubah dari yang paling sederhana atau disebut sebagai model satu peubah. Selanjutnya, setiap peubah akan diseleksi sampai ditemukannya model terbaik. Kriteria model terbaik didasarkan pada peningkatan  $R^2$  dan  $Adj R^2$ , penurunan AIC dan C(p) Mallow (Hanum, 2021). Berdasarkan Tabel 3, diperoleh model terbaik dengan 3 peubah penjelas, diantaranya X1, X2, dan X3 yang didasarkan pada perolehan nilai AIC dan C(p) terkecil.

### 3.4. Deteksi Multikolinearitas Model Terbaik

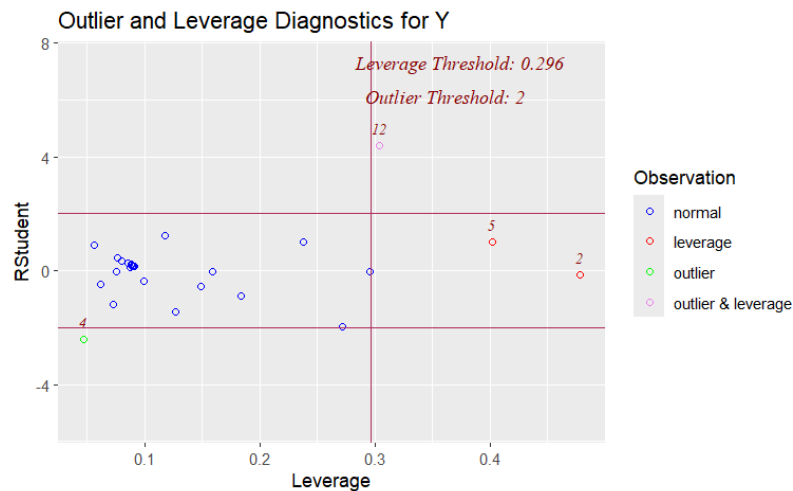
Tabel 5: Nilai VIF Model Terbaik

X1	X2	X6
3.837	2.007	2.876



Peubah penjelas terdeteksi memiliki multikolinearitas jika memiliki nilai VIF > 10. Berdasarkan Tabel 5, tidak terdapat multikolinearitas pada peubah penjelas.

### 3.5. Pemeriksaan Data Model Terbaik



Gambar 2: Grafik Diagnostik Titik Pencilan dan Leverage

Berdasarkan Gambar 2, dapat terlihat pengamatan dengan nilai *RStudent* dan Leverage yang besar. Pengamatan-pengamatan tersebut berpotensi menjadi pencilan (*outlier*) dan leverage.

a. Pencilan

Amatan yang terdeteksi sebagai pencilan adalah amatan ke-4 dan 12, yaitu Kabupaten Cianjur dan Kabupaten Sumedang.

b. Leverage

Amatan yang terdeteksi sebagai leverage adalah amatan ke-2, 5, dan 12, yaitu Kabupaten Bogor, Kabupaten Bandung, dan Kabupaten Sumedang.

c. Amatan Berpengaruh

Amatan yang terdeteksi sebagai amatan berpengaruh adalah amatan ke-2 dan 12, yaitu Kabupaten Bogor dan Kabupaten Sumedang.

### 3.6. Penanganan Pencilan dengan Regresi Kekar (*Robust Regression*)

*Robust Regression* merupakan salah satu metode yang digunakan untuk mengatasi masalah pencilan pada sebuah data. Pencilan tersebut akan memengaruhi metode kuadrat terkecil sehingga berpotensi menghasilkan koefisien regresi yang bias (Deria *et al.*, 2019). Dalam menangani masalah tersebut, digunakan metode *Robust Regression* dengan Estimasi-M pembobot Huber. Estimasi-M pembobot Huber akan meminimumkan MSE (*Mean Squared Error*) pada data pencilan, sehingga MSE Estimasi-M akan lebih kecil dibandingkan dengan MKT (Metode Kuadrat Terkecil). Berdasarkan analisis tersebut, diperoleh hasil persamaan sebagai berikut.

$$\hat{Y} = -7117.162 + 5.626X_1 - 36347.349X_2 + 1896.885X_6$$

Dengan  $R^2$  sebesar 0.9223 dan  $Adj R^2$  sebesar 0.9146.

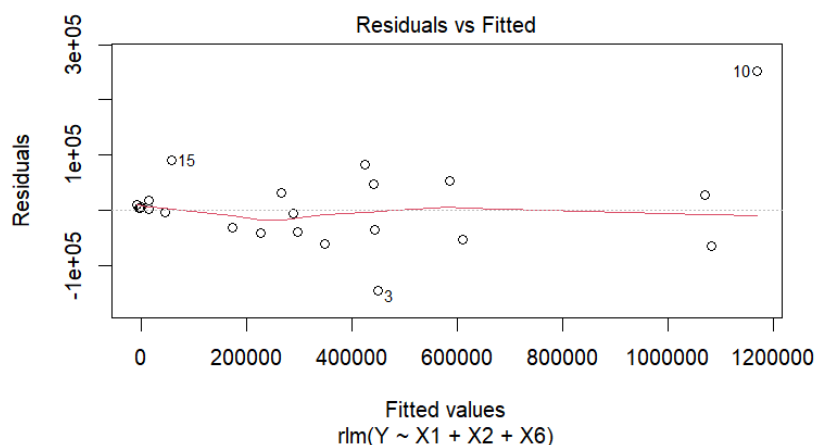
### 3.7. Uji Asumsi Model Regresi Kekar

Uji asumsi formal dilakukan pada model Robust Regression dengan tiga peubah penjelas yang disajikan pada table berikut.

Tabel 6: Hasil Uji Asumsi Formal

Asumsi	Uji	<i>p-value</i>	Keputusan
$E(\varepsilon) = 0$	<i>t-test</i>	0.674	Tak Tolak $H_0$
Autokorelasi	<i>Durbin-Watson test</i>	0.901	Tak Tolak $H_0$
Homoskedastisitas	<i>Breusch-Pagan test</i>	0.021	Tolak $H_0$
Normalitas	<i>Shapiro-Wilk normality test</i>	0.103	Tak Tolak $H_0$

Berdasarkan Tabel 6, terdapat satu asumsi yang dilanggar, yaitu homoskedastisitas dengan keputusan Tolak  $H_0$  (tidak memenuhi). Hal ini mengindikasikan bahwa ragam dari sisaan pada data tidak homogen, atau terjadi heteroskedastisitas



Gambar 3: Plot Sisaan vs Y-Duga Model Regresi Kekar

Berdasarkan Gambar 3, dapat terlihat pula bahwa lebar pita Plot Sisaan dengan Y-Duga cenderung tidak sama, dan membentuk pola corong yang mengindikasikan ragam sisaan tidak homogen.

### 3.8. Penanganan Kondisi Tak Standar dengan WLS

Metode *Weightened Least Squares* (WLS) adalah suatu pendekatan yang memberikan bobot pada setiap titik data berdasarkan tingkat ragamnya. Dalam metode ini, titik data yang memiliki ragam yang besar diberikan bobot yang kecil, sedangkan titik data dengan ragam yang lebih kecil diberikan bobot yang besar. Berikut adalah persamaan dugaan regresi dengan metode WLS.

$$\hat{Y} = -456.443 + 5.347 X_1 - 24500 X_2 + 895.512 X_6$$

Dengan  $R^2$  sebesar 0.9341 dan  $Adj R^2$  sebesar 0.9277.

### 3.9. Uji Asumsi Model yang Sudah Ditangani

Uji asumsi formal dilakukan kembali pada model terbaru dengan taraf nyata sebesar 5%. Berikut hasil uji asumsi formal.

Tabel 7: Hasil Uji Asumsi Formal dengan Metode *Weighted Least Squares*

Asumsi	Uji	p-value	Keputusan
$E(\varepsilon) = 0$	t-test	0.572	Tak Tolak $H_0$
Autokorelasi	Durbin-Watson test	0.831	Tak Tolak $H_0$
Homoskedastisitas	Breusch-Pagan test	1.000	Tak Tolak $H_0$
Normalitas	Shapiro-Wilk normality test	0.082	Tak Tolak $H_0$

Berdasarkan Tabel 7, semua nilai *p-value* di atas 0.05 yang memiliki makna Tak Tolak  $H_0$  atau semua asumsi terpenuhi, sehingga tidak perlu untuk melakukan penanganan asumsi terlanggar kembali.

### 3.10. Interpretasi Model Akhir

Tabel 8: Pendugaan Parameter Model Akhir

Peubah	Koefisien	Galat Baku	t-stat	p-value	
Intercept	-456.443	5886.022	-0.078	0.939	
X1	5.347	0.452	11.816	0.000	**
X2	-24500	6102.172	-4.014	0.000	*
X3	895.512	562.145	1.593	0.038	*

\*)

nyata pada  $\alpha = 5\%$ , \*\*) nyata pada  $\alpha = 1\%$ , \*\*\*) nyata pada  $\alpha = 0.1\%$

$$\text{Produksi P.} = -456.443 + 5.347 \text{ Luas Tanam Padi} - 24500 \text{ Persentase Jumlah Petani} + 895.512 \text{ Jumlah Teknologi Pertanian}$$

Hasil pendugaan parameter model regresi dapat dilihat melalui Tabel 7 dengan tiga peubah penjelas. Peubah yang nyata pada taraf nyata 0.1% adalah peubah Luas Tanam Padi dan Persentase Jumlah Petani, sedangkan Jumlah Teknologi Pertanian nyata pada taraf nyata 5%. Hal ini menunjukkan bahwa ketiga peubah tersebut secara nyata memengaruhi produksi padi di Jawa Barat pada Tahun 2023. Model regresi tersebut memiliki nilai koefisien determinasi ( $R^2$ ) yang

cukup tinggi, yakni sebesar 93,41% dengan  $R^2$  terkoreksi sebesar 92,77%. Dikarenakan nilai koefisien determinasi tersebut terbilang cukup tinggi maka dapat dikatakan bahwa model regresi tersebut memiliki pengaruh yang cukup baik dalam memodelkan data produksi padi di Jawa Barat.

Berdasarkan model akhir, terdapat tiga peubah utama yang memengaruhi produksi padi di Jawa Barat pada Tahun 2023. Peubah-peubah tersebut meliputi luas tanam padi, persentase jumlah petani, dan jumlah teknologi pertanian. Peubah pertama yaitu luas tanam padi yang memiliki peran besar dalam menentukan produksi padi di suatu wilayah. Menurut Nazaruddin (2019), luas tanam padi sangat terkait dengan produksi padi, yang mana semakin luas area tanam padi, maka semakin banyak bibit padi yang bisa tumbuh, sehingga menghasilkan produksi padi yang melimpah.

Peubah kedua yaitu persentase jumlah petani yang merupakan proporsi jumlah petani padi di wilayah tertentu terhadap seluruh petani padi di lingkup seluruh wilayah yang diamati. Model akhir memiliki koefisien negatif terhadap peubah persentase jumlah petani, hal ini menimbulkan kontradiksi terhadap pemahaman bahwa penambahan tenaga kerja akan meningkatkan produksi. Menurut Dhakal (2019), jika proporsi tenaga kerja melewati batas optimal, maka terdapat kemungkinan untuk merugikan atau mengurangi produksi. Hal ini didukung dengan data bahwa petani di Jawa Barat sulit untuk beregenerasi dan jumlah petani yang terlatih tidak banyak, sehingga produktivitas petani semakin menurun yang berujung pada penurunan hasil produksi padi sawah (Oktafiani et al., 2021). Oleh karena itu, dibutuhkan strategi yang tepat dalam penyerapan tenaga kerja petani padi yang akan menguntungkan hasil produksi.

Peubah selanjutnya yaitu jumlah teknologi pertanian yang dapat mempengaruhi produksi padi. Teknologi pertanian akan meningkatkan efektivitas dan efisiensi dalam menghasilkan bibit-bibit pertanian, sehingga produksi akan semakin meningkat (Febrianti et al., 2021).

## **4. Kesimpulan dan Saran**

### **4.1. Kesimpulan**

Metode *Robust Regression* dan *Weightened Least Squares* (WLS) efektif dalam menangani permasalahan pencilan dan heteroskedastisitas pada pemodelan produksi padi di Jawa Barat Tahun 2023. Dibuktikan dengan model akhir yang mampu menjelaskan keragaman sebesar 94.72% dengan tiga peubah utama yang signifikan dalam memengaruhi produksi padi di Jawa Barat Tahun 2023, yaitu luas tanam padi, persentase jumlah petani padi, dan jumlah teknologi pertanian.

### **4.2. Saran**

Metodologi analisis data sebaiknya dimulai dengan eksplorasi data dan mendeteksi pencilan, leverage, dan amatan berpengaruh. Selain itu, jika *Robust Regression* dan *Weightened Least Squares* (WLS) ingin dilakukan dalam menangani

masalah pencilan dan heteroskedastisitas, pastikan memilih urutan metode dengan  $R^2$  terbesar.

### Daftar Pustaka

- Amili F, Rauf A, Saleh Y. 2020. Analisis usahatani padi sawah (*Oryza sativa*, L) serta kelayakannya di Kecamatan Mootilango Kabupaten Gorontalo. *AGRINESIA: Jurnal Ilmiah Agribisnis*. 4(2): 89-94.
- Badan Pusat Statistik. *Sensus Pertanian 2023*. Jakarta: Badan Pusat Statistik.
- Barkah S. 2020. Pengaruh luas lahan dan modal terhadap pendapatan petani padi di Kampung Buyung-buyung Kecamatan Tabalar. *Eco-Build; Economy Bring Ultimate Information All About Development Journal* 4(2); 55-63.
- Chatterjee S, Simonoff JS. 2013. *Handbook of regression analysis*. New Jersey: John Wiley & Sons.
- Dhaka CP. 2019. Multiple regression model fitted for rice production forecasting in nepal: a case of time series data. *Nepalese Journal of Statistics*. 12(2): 89-98.
- Dewayanti AA, Utami H. 2021. Estimasi robust pada model regresi untuk menangani outlier dan heteroskedastisitas. *Jurnal Matematika Thales*. 3(1): 2-12.
- Draper NR & Smith H. 1992. Analisis Regresi Terapan, Edisi kedua. Ir. Bambang Sumantri, penerjemah. Jakarta: Gramedia Pustaka Utama.
- Terjemahan dari: *Applied Regression Analysis*, 2<sup>nd</sup> edition

- Ernawati I, Somayasa W. 2023. Pemilihan model regresi linier berganda dengan kriteria AIC. *Jurnal Matematika Komputasi dan Statistika*. 3(1): 233-237.
- Febrianti VP, Permata TA, Humairoh M, Putri OM, Amelia L, Fatimah S, Khastini RO. 2021. Analisis pengaruh perkebembangan teknologi pertanian di era revolusi industry 4.0 terhadap hasil produksi padi. *Jurnal Pengolahan Pangan*. 6(2): 54-60.
- Fox J. 2016. *Applied regression analysis and generalized linear models*. London (UK): SAGE Publications.
- Gunawan F. 2018. Pengaruh penggunaan faktor produksi terhadap produksi padi di Desa Barugae Kabupaten Bone (*Doctoral dissertation*. UNIVERSITAS NEGERI MAKASSAR). *Jurnal Agribisnis Indonesia*. 12(3): 20-22.
- Hanum H. 2021. Perbandingan metode stepwise, best subset regression, dan fraksi dalam pemilihan model regresi terbaik. *Jurnal Penelitian Sains*. 14(2): 1-6.
- Kannan KS, Manoj K. 2015. Outlier detection in multivariate data. *Applied mathematical sciences*. 47(9): 2317-2324
- Kassambara A. 2018. *Machine learning essentials: practical guide in R*. New York (US): STHDA.
- Mardiatmoko G. 2020. Pentingnya uji asumsi klasik pada analisis regresi linier berganda (studi kasus penyusunan persamaan allometrik kenari muda). *Jurnal Ilmu Matematika dan Terapan*. 14(3): 333-342.
- Nazaruddin. 2019. Luas tanam dan luas panen padi di jawa barat. *Jurnal Triton*. 10(1): 59-68.
- Oktafiani I, Sitohang MY, Saleh R. 2021. Sulitnya regenerasi petani pada kelompok generasi muda. *Jurnal Studi Pemuda*. 10(1): 1-13.
- Open Data Jabar. *Dinas Tanaman Pangan dan Hortikultura*. Jawa Barat: Open Data Jabar
- Pebralia J. 2022. Analisis curah hujan menggunakan machine learning linier berganda berbasis python dan jupyter notebook. *Jurnal Ilmu Pembelajaran*. 6(2): 23-30.
- Raksun A, Japa L, Mertha IG. 2019. Aplikasi pupuk organik dan NPK untuk meningkatkan pertumbuhan vegetative melon (*Cucumis melo* L.). *Jurnal Biologi Tropis*. 19(1): 19-24.
- Rather KU, Kocyigit EG, Onyango R, Kadilar C. 2022. Improved regression in ratio type estimators based on robust M-estimation. *Plos One Journal*. 17(12): 1-16.
- Rihaksa TA, Susanti H. 2023. Penyuluhan pentingnya infrastruktur dalam permintaan pariwisata internasional Indonesia. *Jurnal Abdimas Bina Bangsa*. 4(1): 731-744.

- Sahri RJ, Hidayah N, Fadhillah N, Fuadi A, Abidin I, Hannifa W, Wulandari S. 2022. Tanaman pangan sebagai sumber pendapatan petani di kabupaten karo. *Jurnal Inovasi Pertanian*. 2(10): 3223-3230.
- Saputro WA & Sariningsih W. 2020. Kontribusi pendapatan usahatani kakao terhadap pendapatan rumah tangga petani di taman teknologi pertanian nglangeran, kecamatan pathuk, kabupaten gunungkidul. *Jurnal Sosial Ekonomi Petanian dan Agribisnis*. 16(2): 208-217.