

Pemodelan Produksi Padi di Indonesia Menggunakan Regresi *Robust* Estimasi *Generalized M*

Angelaretha Kurniawan¹, Yuliana Susanti², Hasih Pratiwi³

Universitas Sebelas Maret^{1,2,3}

angelaretha72@student.uns.ac.id

Abstrak—Indonesia sebagai negara agraris yang memiliki populasi jumlah penduduk tergolong cukup banyak, dimana menempati urutan ke-4 dunia setelah negara Tiongkok, India, dan Amerika Serikat. Mayoritas penduduk Indonesia mengolah padi dan menjadikan olahan padi sebagai makanan pokok. Produksi padi di Indonesia dapat mengalami kenaikan atau penurunan setiap tahunnya. Faktor-faktor yang dapat mempengaruhi data produksi padi yaitu luas tanam, luas lahan terkena serangan organisme pengganggu tanaman (OPT), benih padi inbrida, dan jumlah penduduk di Indonesia. Faktor-faktor tersebut memiliki pencilan, dan residualnya tidak berdistribusi normal, sehingga jumlah produksi padi tidak dapat diperkirakan dengan menggunakan analisis regresi metode kuadrat terkecil (MKT). Regresi *robust* dengan estimasi *Generalized M* (GM) dapat digunakan untuk menyelesaikan masalah pencilan dan pelanggaran asumsi normalitas yang terjadi. Tujuan dari penelitian ini adalah untuk mengetahui bagaimana model regresi *robust* dengan estimasi GM untuk mengestimasi produksi padi di Indonesia dan menentukan faktor-faktor yang berpengaruh secara signifikan. Model regresi *robust* estimasi GM pada data produksi padi di Indonesia menunjukkan bahwa peningkatan luas tanam (X_1), benih padi inbrida (X_3), dan jumlah penduduk (X_4) akan meningkatkan jumlah produksi padi, sedangkan peningkatan luas lahan terkena serangan OPT (X_2) akan menurunkan jumlah produksi padi. Hasil dari penelitian ini menunjukkan bahwa seluruh variabel independen berpengaruh secara signifikan terhadap jumlah produksi padi dengan nilai R_{adj}^2 sebesar 99,99%. Pada hasil model regresi *robust* yang diperoleh, diharapkan adanya kebijakan yang diambil dengan mempertimbangkan faktor-faktor yang mempengaruhi jumlah produksi padi guna meningkatkan jumlah produksi padi di Indonesia.

Kata kunci: *Produksi Padi, Analisis Regresi, Regresi Robust, Estimasi GM*

I. PENDAHULUAN

Sektor pertanian merupakan letak kekuatan ekonomi Indonesia sebagai negara agraris dan sebagai salah satu sektor penting dalam kehidupan manusia. Memiliki populasi jumlah penduduk yang tergolong cukup banyak, sebagian besar dari jumlah penduduk Indonesia memanfaatkan sumber daya alam yang ada untuk memenuhi kebutuhan hidupnya, salah satunya dengan bertani [1]. Di Indonesia, padi merupakan salah satu tanaman yang paling banyak ditanam oleh petani. Padi menjadi salah satu komoditas yang sangat menjanjikan karena padi dijadikan sumber bahan makanan pokok bagi masyarakat Indonesia. Perhitungan rata-rata konsumsi beras per kapita per tahun di Indonesia mencapai 111,58 kg. Pada tahun 2021, produksi padi di Indonesia mencapai 54,42 juta ton, dimana mengalami penurunan sebanyak 233,91 ribu ton atau sebesar 0,43% dibandingkan produksi padi tahun sebelumnya [2]. Produksi padi setiap tahunnya dapat mengalami kenaikan atau penurunan, hal ini disebabkan oleh adanya berbagai faktor.

Faktor yang memengaruhi jumlah produksi padi adalah luas tanam, luas lahan terkena serangan organisme pengganggu tanaman (OPT), benih padi inbrida, dan jumlah penduduk. Luas tanam menjadi salah satu faktor jumlah produksi panen. Banyaknya padi yang dapat diproduksi dipengaruhi oleh banyaknya lahan tanam padi, akan tetapi padi tidak dapat diproduksi ketika lahan tanam mengalami gagal panen. Adanya serangan OPT merupakan salah satu penyebab lahan mengalami gagal panen. OPT adalah seluruh organisme yang dapat menyebabkan kerusakan fisik, gangguan fisiologi dan biokimia atau kompetisi hara terhadap tanaman [3]. OPT sebagai salah satu faktor kendala dalam kegiatan budidaya

tanaman, apabila tidak ditangani akan menyebabkan berbagai kerugian. Luas lahan yang terkena serangan OPT dapat mengakibatkan luas panen padi berkurang, sehingga produksi padi juga akan berkurang. Jumlah benih padi inbrida yang digunakan juga akan mempengaruhi jumlah produksi padi. Penggunaan benih padi dengan mutu yang baik akan meningkatkan jumlah produksi padi. Selain faktor yang dapat memengaruhi tanaman padi, faktor luar dimungkinkan untuk memengaruhi hasil produksi padi, salah satunya jumlah penduduk. Bertambahnya jumlah penduduk dapat menyebabkan kenaikan produksi padi dikarenakan semakin meningkat jumlah penduduk, maka akan meningkatkan besarnya kebutuhan dan permintaan bahan makanan pokok. Pemerintah, masyarakat, dan instansi yang terkait telah melakukan berbagai upaya guna mengatasi masalah produksi padi di Indonesia yang mengalami naik turun. Permasalahan produksi padi yang terjadi di Indonesia dapat di analisis menggunakan regresi sebagai metode pemecahan masalahnya.

Data produksi padi di Indonesia pada tahun 2021 mengandung beberapa pencilan dan melanggar asumsi normalitas, sehingga apabila dianalisis dengan menggunakan analisis regresi hasilnya tidak akan optimal. Regresi *robust* dapat digunakan untuk mengatasi masalah pencilan dan pelanggaran asumsi normalitas. Regresi *robust* memiliki beberapa metode, antara lain: estimasi M, estimasi MM, estimasi S, estimasi *Least Median Square* (LMS), estimasi *Least Trimmed Square* (LTS), estimasi *Generalized S* (GS), dan estimasi *Generalized M* (GM).

Artikel ini membahas pemodelan dengan menggunakan regresi *robust* estimasi GM pada data produksi padi di Indonesia tahun 2021 dengan menggunakan variabel independen luas tanam, luas lahan terkena serangan OPT, benih padi inbrida, dan jumlah penduduk. Prediksi produksi padi dan upaya untuk meningkatkan jumlah produksi padi di Indonesia pada tahun berikutnya dapat dilakukan dengan menggunakan model yang didapatkan. Hasil dari penelitian ini ialah sebagai bahan evaluasi, masukan, dan pertimbangan bagi instansi terkait beserta pemerintah dalam upaya mengambil kebijakan, aksi dan program peningkatan produksi padi di Indonesia sehingga dapat meningkatkan produksi padi di Indonesia dengan mempertimbangkan faktor-faktor yang memengaruhinya dan meningkatkan kesejahteraan masyarakat khususnya para petani. Selain itu penelitian ini dapat digunakan sebagai referensi bagi peneliti dan akademisi guna penelitian lanjutan.

II. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini memuat subjek penelitian, teknik analisis data, prosedur penelitian, dan instrumen penelitian.

A. Subjek Penelitian

Subjek penelitian ini adalah produksi padi tiap provinsi di Indonesia pada tahun 2021. Variabel yang digunakan yakni produksi padi (Y), luas tanam (X_1), luas lahan terkena serangan OPT (X_2), benih padi inbrida (X_3), dan jumlah penduduk (X_4) yang bersumber dari *website* resmi Badan Pusat Statistik (BPS) Indonesia dan Kementerian Pertanian Indonesia.

B. Teknik Analisis Data

Hubungan variabel dependen dan independen dapat diketahui dengan menggunakan analisis regresi. Menurut Freund [4] model regresi merupakan aplikasi dari model linier dengan variabel dependen yang dipengaruhi oleh satu atau lebih variabel kuantitatif sebagai faktor atau variabel independen. Hubungan atau pengaruh antara dua atau lebih variabel dependen maka model regresi yang digunakan adalah model regresi linear berganda. Model regresi linier berganda secara umum dapat ditulis sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_j X_{ij} + \varepsilon_i \sim N(0, \sigma^2) \quad (1.1)$$

dengan : Y_i : variabel dependen pengamatan ke- i

β_0 : *intercept* model regresi

β_j : parameter koefisien regresi ke- j , $j = 1, 2, \dots, k$

X_{ij} : variabel independen ke- j pengamatan ke- i

ε_i : residu ke- i

Metode Kuadrat Terkecil (MKT) merupakan estimasi parameter yang biasa digunakan dalam analisis regresi. Model regresi melalui MKT diuji dengan uji asumsi klasik untuk mengetahui adakah penyimpangan pada data. Uji asumsi klasik yang digunakan yakni asumsi normalitas, asumsi homoskedastisitas, asumsi non autokorelasi, asumsi multikolinearitas.

Data yang mengandung pencilan apabila dianalisis dengan menggunakan analisis regresi hasilnya kurang tepat. Pencilan merupakan data yang jauh lebih besar atau jauh lebih kecil dari data lainnya. Salah

satu cara mendeteksi adanya pencilan yaitu menggunakan *DFFITS*. Nilai *DFFITS* dirumuskan sebagai berikut:

$$DFFITS_i = t_i \sqrt{\frac{h_{ii}}{1 - h_{ii}}} \quad (1.2)$$

dengan t_i *studentized deleted* residu untuk kasus ke- i dan h_{ii} nilai *leverage* untuk kasus ke- i . Apabila nilai $|DFFITS| > 2\sqrt{(k + 1/n)}$, maka pencilan dapat diketahui.

Salah satu solusi untuk menganalisa data yang mengandung pencilan berpengaruh sehingga model dapat kekar terhadap pencilan tersebut yaitu dengan menggunakan regresi *robust* [4]. Regresi *robust* adalah regresi yang digunakan ketika distribusi dari residu tidak normal atau adanya beberapa pencilan yang berpengaruh pada model [5]. Penyimpangan-penyimpangan yang ada dalam MKT dapat diatasi dengan regresi *robust*. Pada saat model regresi ditetapkan oleh peneliti dan diketahui melanggar asumsi saat menguji, sedangkan pengaruh pencilan yang menjadikan prediksi bias tidak dapat dihilangkan atau dilemahkan dengan transformasi, maka dalam keadaan ini regresi *robust* adalah metode terbaik [6].

Salah satu metode estimasi regresi *robust* yang sering digunakan adalah estimasi M. Huruf M menunjukkan bahwa estimasi M adalah estimasi ‘tipe maksimum likelihood’. Pada prinsipnya, estimasi M merupakan estimasi yang meminimumkan suatu fungsi residu ρ

$$\hat{\beta}_m = \min_{\beta} \sum_{i=1}^n \rho\left(\frac{e_i}{\sigma}\right) = \min_{\beta} \sum_{i=1}^n \rho\left(\frac{y_i - x_i' \beta}{\sigma}\right) \quad (1.3)$$

dengan $\sigma = \frac{\text{median}|e_i - \text{median}(e_i)|}{0.6745}$, dimana residu $e_i = y_i - \hat{y}_i$

Fungsi ρ yang digunakan adalah fungsi pembobot Tukey *bisquare*.

$$\rho(u_i) = \begin{cases} \frac{u_i^2}{2} - \frac{u_i^4}{2(4,685)^2} + \frac{u_i^6}{6(4,685)^4}, & |u_i| \leq 4,685 \\ \frac{(4,685)^2}{6}, & |u_i| > 4,685 \end{cases}$$

Seiring berjalannya waktu, terdapat metode estimasi yang lain yang dapat digunakan dalam analisis regresi *robust*, yakni metode estimasi GM. Regresi *robust* dengan metode estimasi GM merupakan pengembangan dari metode estimasi M, ketika estimasi M kurang sensitif terhadap pencilan. Metode estimasi GM memiliki *high breakdown point* sebesar 50%. Oleh karena itu, regresi *robust* dengan metode estimasi GM ini dinilai mampu memberikan model prediksi yang baik pada data produksi padi di Indonesia. Estimasi GM menggunakan fungsi pembobot w_i yang bergantung pada variabel x_i . Menurut Wilcox [7], secara umum, estimasi GM didefinisikan sebagai berikut:

$$S(\beta_j) = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n w_i \rho\left(\frac{e_i}{v(x_i)}\right) \quad (1.4)$$

dengan $v(x_i)$ merupakan fungsi pembobot untuk variabel x_i . Karena estimator $\hat{\beta}$ yang diperoleh bukan merupakan skala *invariant*, maka digunakan nilai $\frac{e_i}{\hat{\sigma}}$ sebagai pengganti e_i , dengan $\hat{\sigma}$ adalah faktor skala yang juga perlu diestimasi, sehingga persamaannya menjadi,

$$S(\beta_j) = \sum_{i=1}^n w_i \rho\left(\frac{y_i - \sum_{j=0}^k x_{ij} \beta_j}{v(x_i) \hat{\sigma}}\right) \quad (1.5)$$

Persamaan (1.5) dapat diselesaikan dengan cara menurunkan terhadap β , dan disamakan dengan nol, diperoleh sebagai berikut:

$$\sum_{i=1}^n w_i \rho\left(\frac{y_i - \sum_{j=0}^k x_{ij} \beta_j}{v(x_i) \hat{\sigma}}\right) = 0 \quad ; j = 0, 1, \dots, k \quad (1.6)$$

$$\sum_{i=1}^n w_i \psi\left(\frac{y_i - \sum_{j=0}^k x_{ij} \beta_j}{v(x_i) \hat{\sigma}}\right) = 0 \quad ; j = 0, 1, \dots, k \quad (1.7)$$

Menggunakan pembobot *Schweppe* dengan $w_i = \sqrt{1 - h_{ii}}$ dan $(v(x_i)) = w_i$ diperoleh

$$w_i = w_i \frac{\psi\left(\frac{u_i}{v(x_i)}\right)}{(u_i)}$$

$$\begin{aligned}
 &= \frac{w_i}{u_i} \psi\left(\frac{u_i}{v(x_i)}\right) \\
 &= \frac{\sqrt{1-h_{ii}}}{u_i} \psi\left(\frac{u_i}{\sqrt{1-h_{ii}}}\right)
 \end{aligned} \tag{1.8}$$

dimana $u_i = \frac{e_i}{\hat{\sigma}}$, dengan $\hat{\sigma} = 1,48M_l$, dan M_l merupakan median dari $(n - k)$ terbesar dari residu e_i dan $\psi(x) = \max\{-K, \min(K, x)\}$ adalah fungsi pengaruh *Huber* dengan $K = 1,5\hat{\sigma}$. sehingga menjadi,

$$\sum_{i=1}^n x_{ij} w_i \left(y_i - \sum_{j=0}^k x_{ij} \beta_j \right) = 0 \quad j = 0, 1, \dots, k \tag{1.9}$$

Residu awal yang digunakan pada estimasi GM adalah residu yang diperoleh dari MKT. Persamaan (1.9) dapat diselesaikan dengan MKT terboboti secara iterasi yang disebut *Iteratively Reweighted Least Square* (IRLS) hingga diperoleh nilai $\hat{\beta}$ yang konvergen.

Pada penelitian sebelumnya, penelitian yang dilakukan oleh Ishaq [8] tentang faktor-faktor yang memengaruhi produksi padi di Provinsi Jawa Timur menggunakan regresi semiparametrik *spline*. Hasil penelitian menunjukkan bahwa faktor-faktor yang berpengaruh signifikan terhadap produksi padi, yaitu luas panen padi dan curah hujan, sedangkan faktor-faktor yang tidak berpengaruh signifikan adalah luas puso padi, realisasi pupuk bersubsidi, dan ketinggian rata-rata dari permukaan laut dengan nilai R_{adj}^2 yang didapatkan adalah sebesar 99,17%. Yulisti'annah [9] telah meneliti mengenai estimasi parameter model *Mixed Geographically Weighted Poisson Regression* (MGWPR) yang mengandung pencilan dengan metode estimasi GM, hasilnya didapatkan bahwa metode estimasi GM dapat digunakan untuk menangani data yang mengandung pencilan yang dapat menyebabkan analisis menjadi tidak tepat. Perbandingan metode regresi *robust* estimasi GM dengan menggunakan berbagai pembobot telah dilakukan oleh Subzar [10] dalam studi kasus metode regresi kuat untuk penduga hingga rerata populasi berdasarkan pengambilan sampel acak sederhana tanpa penggantian dalam kasus *outlier* atau pencilan. Didapatkan hasil bahwa metode estimasi GM dengan pembobot *Mallows*, metode estimasi GM dengan pembobot *Schweppe*, dan metode estimasi GM dengan pembobot *SIS*, lebih baik daripada metode MKT dan metode estimasi M dengan pembobot *Huber*.

C. Prosedur Penelitian

Prosedur yang dilakukan pada penelitian ini adalah

1. Melakukan estimasi koefisien regresi pada data menggunakan MKT.
2. Menguji asumsi klasik dari model regresi.
3. Mendeteksi adanya pencilan pada data.
4. Melakukan estimasi koefisien regresi *robust* menggunakan estimasi GM.
 - a. Menentukan estimasi parameter $\hat{\beta}^0$ dengan MKT.
 - b. Menentukan residu e_i .
 - c. Menentukan nilai estimasi $\hat{\sigma}$.
 - d. Menentukan nilai u_i .
 - e. Menentukan pembobot w_i .
 - f. Mengestimasi parameter $\hat{\beta}_{GM}$ menggunakan metode WLS (*Weighted Least Squares*) dengan pembobot w_i .
 - g. Melakukan pengulangan langkah b-f hingga mendapatkan nilai $\hat{\beta}_{GM}$ yang konvergen.
 - h. Melakukan uji hipotesis untuk mengetahui apakah variabel independen berpengaruh signifikan terhadap variabel dependen.

D. Instrumen Penelitian

Analisis dilakukan dengan menggunakan *software Rstudio*.

III. HASIL DAN PEMBAHASAN

Penelitian ini dilakukan dengan mengestimasi koefisien regresi dengan MKT, melakukan uji asumsi klasik, mendeteksi pencilan, dan melakukan analisis regresi *robust* dengan estimasi GM.

A. Estimasi MKT

Model regresi yang didapatkan untuk produksi padi di Indonesia pada tahun 2021 dengan menggunakan MKT adalah sebagai berikut:

$$\hat{Y} = -111700 + 5,187 X_1 - 16,78 X_2 + 57,05 X_3 + 13,38 X_4 \quad (3.1)$$

dengan nilai R_{adj}^2 sebesar 99,85% yang berarti setidaknya variabel independen dapat menjelaskan sebesar 99,85% variansi variabel dependen dan sisanya sebesar 0,15% dipengaruhi oleh faktor lain yang tidak dimasukkan ke dalam model pada penelitian ini.

Dari model persamaan regresi (3.1) menunjukkan bahwa peningkatan 1 hektar luas tanam (X_1) akan meningkatkan 5,187 ton produksi padi (Y), peningkatan 1 hektar luas lahan terkena serangan OPT (X_2) akan menurunkan 16,78 ton produksi padi (Y), peningkatan 1 ton benih padi inbrida (X_3) akan meningkatkan 57,05 ton produksi padi (Y), dan peningkatan 1000 jiwa penduduk akan meningkatkan 13,38 ton produksi padi (Y).

B. Uji Asumsi Klasik

Uji Asumsi Normalitas

- 1) H_0 : Residu berdistribusi normal
 H_1 : Residu tidak berdistribusi normal
- 2) Taraf signifikansi
 $\alpha = 0,05$
- 3) Daerah kritis
 H_0 ditolak jika nilai $D > D_{(\alpha,n)}$ atau $p - value < \alpha$, dimana $D_{(\alpha,n)}$ adalah nilai pada tabel *Kolmogolov-Smirnov*
- 4) Statistik Uji
Berdasarkan *output Rstudio* diperoleh nilai $D = 0,55882$ dan $p - value = 1,573 \times 10^{-10}$
- 5) Kesimpulan
 H_0 ditolak karena $D = 0,55882 > D_{(0,05,34)} = 0,150$ atau $p - value = 1,573 \times 10^{-10} < \alpha = 0,05$ sehingga dapat disimpulkan bahwa residu tidak berdistribusi normal.

Uji Asumsi Homoskedastisitas

- 1) H_0 : $var(e_i) = \sigma^2$ (Variansi residu bersifat homogen)
 H_1 : $var(e_i) \neq \sigma^2$ (Variansi residu bersifat tidak homogen)
- 2) Taraf signifikansi
 $\alpha = 0,05$
- 3) Daerah kritis
 H_0 ditolak jika nilai $\phi_{hitung} > \chi_{(\alpha;k)}^2$ atau $p - value < \alpha$
- 4) Statistik Uji
Berdasarkan *output Rstudio* diperoleh nilai $\phi_{hitung} = 0,55154$ dan $p - value = 0,9683$
- 5) Kesimpulan
 H_0 gagal ditolak karena nilai $\phi_{hitung} = 0,55154 < \chi_{(0,05;4)}^2 = 9,4877$ atau $p - value = 0,9683 > \alpha = 0,05$, sehingga dapat disimpulkan bahwa variansi residu bersifat homogen.

Uji Asumsi Non Autokorelasi

- 1) H_0 : Tidak terdapat autokorelasi antar residu
 H_1 : Terdapat autokorelasi antar residu
- 2) Taraf signifikansi
 $\alpha = 0,05$
- 3) Daerah kritis
 H_0 ditolak jika nilai $0 < Durbin-Watson < dL$ atau $4-dL < Durbin-Watson < 4$ atau $p - value < \alpha$, dimana dU adalah batas atas nilai tabel *Durbin-Watson* dan dL adalah batas bawah nilai tabel *Durbin-Watson*.
- 4) Statistik Uji
Berdasarkan *output Rstudio* diperoleh nilai *Durbin-Watson* = 1,4897 dan $p - value = 0,05612$
- 5) Kesimpulan
 H_0 gagal ditolak karena nilai $0 < 1,4897 < 1,7277$ atau $p - value = 0,051612 > \alpha = 0,05$, sehingga dapat disimpulkan bahwa tidak terdapat autokorelasi antar residu.

Uji Asumsi Non Multikolinearitas

- 1) H_0 : Tidak terdapat multikolinearitas antar variabel independen
 H_1 : Terdapat multikolinearitas antar variabel independen
- 2) Taraf Signifikansi
 $\alpha = 0,05$
- 3) Daerah kritis
 H_0 ditolak jika terdapat nilai *VIF* yang melebihi 10
- 4) Statistik Uji
 Berdasarkan *output Rstudio* diperoleh nilai *VIF* sebagai berikut

TABEL 1. HASIL PERHITUNGAN NILAI *VIF*

Variabel Independen	VIF	Kesimpulan
X_1	9,200383	Tidak terdapat multikolinearitas
X_2	5,684700	Tidak terdapat multikolinearitas
X_3	1,932918	Tidak terdapat multikolinearitas
X_4	9,953463	Tidak terdapat multikolinearitas

- 5) Kesimpulan
 Berdasarkan hasil pada Tabel 1, H_0 gagal ditolak karena keempat variabel independen memiliki nilai *VIF* kurang dari 10, sehingga dapat disimpulkan bahwa tidak terdapat multikolinearitas antar variabel independen.

Hasil dari uji *VIF* menunjukkan bahwa pengujian asumsi non multikolinearitas pada data produksi padi di Indonesia tahun 2021 dapat terpenuhi.

Berdasarkan uji asumsi klasik yang telah dilakukan menunjukkan bahwa pengujian asumsi homoskedastisitas, asumsi non autokorelasi, dan asumsi non multikolinearitas pada data produksi padi di Indonesia terpenuhi, sedangkan pengujian asumsi normalitas tidak terpenuhi.

C. Deteksi Pencilan

Nilai *DFFITs* digunakan untuk mengetahui ada atau tidaknya pencilan pada data. Berdasarkan data yang dimiliki nilai $2\sqrt{(k + 1/n)} = 0,766965$. Apabila nilai $|DFFITs|$ lebih besar dari 0,766965 maka bisa dikatakan data tersebut merupakan pencilan. Pencilan dari data produksi padi di Indonesia tahun 2021 adalah sebagai berikut :

TABEL 2. IDENTIFIKASI PENCILAN

Observasi ke	DFFITs	$2\sqrt{(k + 1/n)}$
8	0,90424759	0,766965
13	0,79388108	
20	-0,87376983	

Berdasarkan Tabel 2, observasi ke 8, 13, dan 20 merupakan pencilan yaitu Provinsi Lampung, Provinsi Jawa Tengah, dan Provinsi Kalimantan Barat. MKT kurang tepat digunakan pada data yang mengandung pencilan, sehingga perlu dilakukan analisis regresi *robust* dengan metode estimasi GM untuk mengatasi masalah pencilan pada penelitian ini.

D. Estimasi GM

Model regresi *robust* dengan metode estimasi GM pada data produksi padi di Indonesia tahun 2021, modelnya adalah sebagai berikut:

$$\hat{Y} = -61743,511 + 5,268 X_1 - 23,301 X_2 + 91,235 X_3 + 8,918 X_4 \quad (3.2)$$

Nilai R_{adj}^2 sebesar 99,99% yang berarti setidaknya variabel independen dapat menjelaskan sebesar 99,99% variansi variabel dependen dan sisanya sebesar 0,01% dipengaruhi oleh faktor lain yang tidak dimasukkan ke dalam model pada penelitian ini.

Model tersebut didapatkan setelah iterasi ke-77. Dari model persamaan regresi (3.2) menunjukkan bahwa peningkatan 1 hektar luas tanam (X_1) akan meningkatkan 5,268 ton produksi padi (Y), peningkatan 1 hektar luas lahan terkena serangan OPT (X_2) akan menurunkan 23,301 ton produksi padi (Y), peningkatan 1 ton benih padi inbrida (X_3) akan meningkatkan 91,235 ton produksi padi (Y), dan peningkatan 1000 jiwa penduduk akan meningkatkan 8,918 ton produksi padi (Y). Selanjutnya dilakukan uji signifikansi secara serentak dengan menggunakan uji F untuk mengetahui ada atau tidaknya pengaruh variabel independen secara keseluruhan terhadap variabel dependen pada data produksi padi di Indonesia tahun 2021.

- 1) $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (semua variabel independen tidak berpengaruh secara signifikan terhadap produksi padi)
 $H_1 : \text{minimal ada satu } \beta_i \neq 0, i = 1, 2, 3, 4$ (paling tidak ada satu variabel independen yang berpengaruh secara signifikan terhadap produksi padi)
- 2) Taraf signifikansi
 $\alpha = 0,05$
- 3) Daerah kritis
 H_0 ditolak jika nilai $F_{hitung} > F_{(\alpha, k, n-k-1)}$ atau $p - \text{value} < \alpha$
- 4) Statistik Uji
 Berdasarkan *output Rstudio* diperoleh nilai $F_{hitung} = 1,041 \times 10^5$ dan $p - \text{value} = < 2,2 \times 10^{-16}$
- 5) Kesimpulan
 H_0 ditolak karena $F_{hitung} = 1,041 \times 10^5 > F_{(0,05;4;29)} = 2,70$ atau $p - \text{value} = < 2,2 \times 10^{-16} < \alpha = 0,05$, sehingga dapat disimpulkan bahwa paling tidak ada satu variabel independen yang berpengaruh secara signifikan terhadap produksi padi.

Selanjutnya dilakukan uji signifikansi secara parsial dengan menggunakan uji t untuk mengetahui variabel independen mana yang secara individu berpengaruh signifikan terhadap variabel dependen pada data produksi padi di Indonesia.

- 1) $H_0 : \beta_i = 0, i = 1, 2, 3, 4$ (variabel independen ke- i tidak berpengaruh secara signifikan terhadap produksi padi)
 $H_1 : \beta_i \neq 0, i = 1, 2, 3, 4$ (variabel independen ke- i berpengaruh secara signifikan terhadap produksi padi)
- 2) Taraf signifikansi
 $\alpha = 0,05$
- 3) Daerah kritis
 H_0 ditolak jika nilai $|t_{hitung}| > t_{(\frac{\alpha}{2}, n-k-1)}$ atau $p - \text{value} < \alpha$
- 4) Statistik Uji
 Berdasarkan *output Rstudio* diperoleh nilai $|t_{hitung}|$ sebagai berikut.

TABEL 3. HASIL UJI T ESTIMASI GM

Variabel Independen	$ t_{hitung} $	$p - \text{value}$
X_1	170,928	$< 2 \times 10^{-16}$
X_2	-11,470	$2,69 \times 10^{-12}$
X_3	16,601	$2,39 \times 10^{-16}$
X_4	5,794	$2,81 \times 10^{-6}$

- 5) Kesimpulan
 Berdasarkan Tabel 3, dapat diketahui bahwa semua variabel independen memiliki nilai $|t_{hitung}| > t_{(0,025;29)} = 2,0452$ atau nilai $p - \text{value} < \alpha = 0,05$ sehingga H_0 ditolak dapat disimpulkan bahwa seluruh variabel independen yaitu luas tanam (X_1), luas lahan terkena serangan OPT (X_2), benih padi inbrida (X_3), dan jumlah penduduk (X_4) berpengaruh secara signifikan terhadap jumlah produksi padi (Y).

IV. SIMPULAN DAN SARAN

A. Simpulan

Regresi *robust* digunakan untuk menganalisa data yang mengandung pencilan berpengaruh sehingga model dapat kekar terhadap pencilan. Metode estimasi GM sebagai salah satu metode regresi *robust* merupakan pengembangan dari estimasi M, ketika estimasi M kurang sensitif terhadap pencilan. Metode estimasi GM memiliki *high breakdown point* sebesar 50%. Model regresi *robust* dengan metode estimasi GM pada data produksi padi di Indonesia yang dapat digunakan untuk memprediksi jumlah produksi padi adalah $\hat{Y} = -61743,511 + 5,268 X_1 - 23,301 X_2 + 91,235 X_3 + 8,918 X_4$ dengan R_{adj}^2 sebesar 99,99% yang berarti variabel independen dapat menjelaskan 99,99% variabel dependen sisanya 0,01% dipengaruhi variabel lain yang tidak dimasukkan dalam model pada penelitian ini. Model dari regresi *robust* estimasi GM pada produksi padi di Indonesia menunjukkan bahwa seluruh variabel independen berpengaruh secara signifikan.

B. Saran

Bagi peneliti yang berminat untuk melanjutkan penelitian dengan topik yang sama dapat menggunakan variabel independen lain seperti jumlah pupuk, jumlah petani, luas lahan terkena banjir, atau luas lahan terkena kekeringan sebagai faktor-faktor yang dapat mempengaruhi produksi padi di Indonesia. Penelitian lebih lanjut mengenai karakteristik dan kinerja metode estimasi GM perlu dilakukan dengan menggunakan data yang lebih besar untuk mendapatkan hasil estimasi yang lebih optimal.

DAFTAR PUSTAKA

- [1] Ruauw, E., 2010. Nilai Tukar Petani Sebagai Indikator Kesejahteraan Petani. *Jurnal penelitian ASE*.
- [2] BPS. 2022. *Luas Panen, Produksi, dan Produktivitas Padi Menurut Provinsi*. [online] Terdapat pada : <<https://www.bps.go.id/indicator/53/1498/1/luas-panen-produksi-dan-produktivitas-padi-menurut-provinsi.html>> [Diakses 29 Agustus 2022].
- [3] Pakpahan, A. V. & Doni, 2019. Implementasi Metode Forward Chaining untuk Mendiagnosis Organisme Pengganggu Tanaman (OPT) Kopi. *Jurnal Simetris*, pp. 117-126.
- [4] Freund, R. J., Wilson, W. J. & Sa, P., 2006. *Regression Analysis Statistical Modeling of a Response Variable Second Edition*. Boston: Academic Press.
- [5] Chen, C. (2002). *Robust Regression and Outlier Detection with the ROBUSTREG Procedure*. North Carolina: SAS Institute Inc.
- [6] Susanti, Y., Pratiwi, H., Handajani, S. S. & Liana, T., 2014. M Estimation, S Estimation, And MM Estimation. *International Journal of Pure and Applied Mathematics*, Volume 91.
- [7] Wilcox, R., 2005. *Introduction to Robust estimation and Hypothesis*. San Diego: Academic Press.
- [8] Ishaq, M., Rumiati, A. T. & Permatasari, E. O., 2017. Analisis faktor-Faktor yang Memengaruhi Produksi Padi di Provinsi Jawa Timur Menggunakan Regresi Semiparametrik Spline. *Jurnal Sains dan Seni ITS*, pp. 94-100.
- [9] Yulisti'anah, 2018. *Estimasi Parameter Model Mixed Geographically Weighted Poisson Regression (MGWPR) yang Mengandung Outlier dengan Metode GM-Estimator*, Malang : Universitas Islam Negeri Maulana Malik Ibrahim.
- [10] Subzar, M., Al-Omari, A.I. & Alanzi, A.R.A, 2020. The Robust Regression Methods for Estimating of Finite Population Mean Based on SRSWOR in Case of Outliers. *Journal of Computers, Materials & Continua*, Volume 65, pp. 125-138.