

ESTIMASI HUBUNGAN ANTARA PRODUKSI PADI TERHADAP LUAS PANEN DAN PRODUKTIVITAS PADI DI INDONESIA DENGAN MENGGUNAKAN METODE REGRESI ROBUST ESTIMASI-M

Febriani Astuti¹

¹ Jurusan Statistika, Fakultas Sains Terapan, Institut Sains & Teknologi AKPRIND
Email: febriani@akprind.ac.id

Abstract. *Indonesia is an agricultural country with national rice production reaching 32.07 million tons in 2022. Two factors are thought to influence rice production, namely harvested area and rice productivity. The latest data for 2022 shows that an increase in rice production is directly proportional to an increase in harvested area, while productivity tends to fluctuate every year. One method that can be used to analyze rice production is linear regression. Regression analysis is a statistical technique used to determine the relationship between the dependent and independent variables. In this case, rice production is the dependent variable, while the harvested area and productivity are the independent variables. However, if there are outlier elements that violate the rules of normality in the data in this variable, it can cause inaccurate results. Therefore other methods are used to produce better results. The method in question is the M estimation method of robust regression. From the regression model, the coefficient of determination R-Sq(adj) is 99.87%. This shows that rice production in Indonesia in 2022 is influenced by harvested area and rice productivity, other variables influence the remaining 0.13%.*

Keywords: *Rice Production, Robust Regression, Estimation of M.*

Abstrak. Indonesia adalah negara agraris dengan hasil produksi padi nasional mencapai 32,07 juta ton pada tahun 2022. Dua faktor yang dianggap berpengaruh terhadap produksi padi yaitu luas panen dan produktivitas padi. Data terbaru tahun 2022, kenaikan produksi padi berbanding lurus dengan kenaikan luas panennya, sedangkan produktivitas cenderung mengalami fluktuasi setiap tahunnya. Salah satu metode yang dapat digunakan untuk menganalisis produksi padi adalah regresi linear. Analisis regresi merupakan teknik statistika yang digunakan untuk mengetahui hubungan antara variabel dependen dan variabel independen. Dalam hal ini produksi padi sebagai variabel dependen, sedangkan luas panen dan produktivitas sebagai variabel independen. Akan tetapi apabila ada unsur pencilon dan melanggar aturan normalitas dalam data di variabel tersebut dapat menyebabkan hasil yang tidak akurat. Oleh karena itu digunakan metode lain untuk menghasilkan hasil yang lebih baik. Metode yang dimaksud adalah metode regresi robust estimasi M. Dari model regresi tersebut diperoleh koefisien determinasi R-Sq(adj) 99,87%. Hal ini menunjukkan bahwa produksi padi di Indonesia tahun 2022 dipengaruhi oleh luas panen dan produktivitas padi, sisanya sebesar 0,13% dipengaruhi oleh variabel lain.

Kata kunci: *Produksi Padi, Regresi Robust, Estimasi M.*

1. Pendahuluan

Indonesia menjadi negara dengan jumlah penduduk terbesar keempat di dunia. Worldometers memperkirakan jumlah penduduk di dalam negeri sebanyak 279,36 juta orang hingga 24 Juni 2022. Selain sebagai negara penyumbang penduduk terbanyak urutan keempat dunia, Indonesia juga dikenal dengan negara agraris. Sebagian besar penduduknya bekerja di bidang pertanian dengan komoditas utama padi. Hal ini disebabkan karena makanan pokok yang dikonsumsi di Indonesia adalah nasi. Dari hasil penghitungan BPS, produksi beras nasional 2022 diperkirakan mencapai 32,07 juta ton. Jumlah ini naik 0,72 juta ton dari 2021, yang 31,36 juta ton. Kenaikan produksi beras ini berbanding lurus dengan kenaikan luas panen dan produktivitas padi. Pada tahun 2022 luas panen padi di Indonesia mencapai sekitar 10,45 juta

hektar, ini artinya terdapat kenaikan sebanyak 40,87 ribu hektar atau 0,39 persen dibandingkan luas panen padi di 2021 yaitu sebesar 10,41 juta hektar. Di sisi lain, Badan Pusat Statistik (BPS) mencatat, produktivitas padi di Indonesia mencapai 52,38 kuintal/hektare (ku/ha) pada 2022. Jumlahnya naik 0,23% dibandingkan tahun sebelumnya yang sebesar 52,26 ku/ha. Namun secara tren, produktivitas padi secara nasional cenderung berfluktuasi dalam sedekade terakhir. Kondisi kenaikan produksi padi dan beberapa faktor pendukungnya ini membawa lembaga penelitian beras dunia, International Rice Research Institute (IRRI) memberikan penghargaan Indonesia, sebagai negara yang mampu mencapai swasembada beras. IRRI menilai, Indonesia mencapai swasembada karena mampu memenuhi kebutuhan masyarakat lebih dari 20 persen.

Data produksi padi dengan dua faktor utamanya yaitu luas panen dan produktivitas menjadi bahan yang menarik untuk dikaji lebih dalam, khususnya data terkini tahun 2022. Salah satu metode yang dapat digunakan untuk menganalisis produksi padi adalah regresi linear. Analisis regresi merupakan teknik statistika yang digunakan untuk mengetahui hubungan antara variabel dependen dan variabel independen. Jika Y variabel dependen dan X_1, X_2, \dots, X_k variabel independen, maka model regresi linear secara umum dapat dinyatakan sebagai $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i$, dengan $\beta_0, \beta_1, \dots, \beta_k$ adalah parameter-parameter regresi dan ε_i adalah sisaan yang berdistribusi normal dengan mean nol dan variansi konstan [8]. Akan tetapi apabila ada unsur pencilan dan melanggar aturan normalitas dalam data tersebut dapat menyebabkan hasil yang tidak akurat. Oleh karena itu digunakan metode lain untuk menghasilkan analisis yang lebih baik. Metode yang dimaksud adalah metode regresi robust. Regresi robust adalah regresi yang digunakan ketika distribusi dari residu tidak normal atau adanya beberapa pencilan yang berpengaruh pada model [1]. Ada beberapa metode dalam regresi robust yang dapat digunakan untuk menangani pencilan, diantaranya adalah estimasi M, estimasi S, estimasi MM dan estimasi GM. Estimasi M merupakan estimasi pada regresi robust yang paling sederhana dan sering digunakan. Estimasi M merupakan salah satu metode dalam regresi robust untuk mengestimasi parameter yang disebabkan oleh data pencilan. Estimasi ini dikatakan “tipe maksimum likelihood” karena mempunyai sifat tak bias dan varians minimum [3]. Estimasi M dipandang dengan baik untuk mengestimasi parameter yang disebabkan oleh suatu pencilan dan memiliki breakdown point $1/n$. Breakdown point adalah salah satu cara untuk mengukur ke-robust-an suatu estimator [9]. Breakdown point merupakan proporsi minimal dari banyaknya pencilan dibandingkan keseluruhan data yang dapat ditangani sebelum pengamatan tersebut mempengaruhi model.

Penelitian sebelumnya dilakukan oleh [7] dengan menggunakan regresi robust estimasi S, M dan MM diperoleh hasil faktor yang berpengaruh terhadap produksi padi adalah luas panen dan produksi benih. Lain hanya dengan [5] yang melakukan pemodelan produksi padi menggunakan regresi robust estimasi GM sebagai alternatif lain selain estimasi M. Dari penelitian sebelumnya, belum ada penelitian terbaru yang menggunakan metode robust estimasi M untuk mengestimasi hubungan antara antara produksi padi terhadap luas panen dan produktivitas padi di Indonesia tahun 2022. Dengan demikian, diharapkan penelitian ini dapat memberikan kontribusi yang baik untuk pemerintah Indonesia, peneliti, akademisi dan berbagai pihak terkait lainnya untuk kepentingan yang lebih luas lagi.

2. Metode Penelitian

Penelitian ini berupa penelitian kuantitatif dengan menggunakan metode regresi robust estimasi M dengan fungsi pembobot Tukey Bisquare untuk mengestimasi hubungan antara produksi padi terhadap luas panen dan produktivitas padi di Indonesia tahun 2022. Data yang digunakan merupakan data sekunder dari website BPS berupa data produksi padi serta luas panen dan produktivitas padi dari 34 provinsi di Indonesia.

2.1 Langkah Analisis Data

Langkah-langkah yang dilakukan dalam penelitian ini mencakup beberapa hal, diantaranya diberikan sebagai berikut:

1. menentukan data banyaknya produksi padi pada tiap provinsi di Indonesia sebagai variabel dependen (Y),
2. menentukan variabel independen (X) sebagai berikut
 - a. luas panen (X1),
 - b. produktivitas (X2),
3. menentukan model regresi menggunakan analisis regresi berganda,
4. menguji asumsi klasik,
5. menguji adanya data pencilan,
6. melakukan analisis regresi menggunakan regresi robust estimasi-M dengan pembobot Huber

2.2 Estimasi M

Tahap-tahap analisis pada penelitian ini adalah :

1. mengestimasi β dengan Metode Kuadrat Terkecil,
2. menghitung nilai sisaan, $e_i = y_i - \hat{y}_i$,
3. menghitung nilai skala regresi robust estimasi M, $\hat{\sigma}$

$$MAD = \text{median} | e_i - \text{median}(e_i) |$$

$$\hat{\sigma} = 1,5 \times MAD,$$

4. menghitung nilai residual (Ui)
- $$U_i = \frac{e_i}{\hat{\sigma}},$$
5. mendefinisikan fungsi pembobot
- $$w_i = w(U_i),$$

$$\text{dengan metode Tukey Bisquare } w_B = \begin{cases} (1 - (\frac{u_i}{4,685})^2)^2 & , |u_i| \leq 4,685 \\ 0 & , |u_i| > 4,685 \end{cases}$$

6. memperbaiki estimasi $\hat{\beta}$ berdasarkan *Weighted Least Squares*(WLS) dengan pembobot w_i sehingga diperoleh $\hat{\beta}$ baru pada iterasi ke 1,
7. mengulangi tahap 2-6 sehingga diperoleh $\hat{\beta}$ yang konvergen,
8. mencari model terbaik berdasarkan iterasi paling sedikit dari kedua.

3. Hasil dan Pembahasan

3.1 Statistika Deskriptif

Dari data sekunder yang sudah disampaikan sebelumnya, akan ditampilkan statistika deskriptif dalam tabel berikut.

Tabel 1. Statistika Deskriptif Variabel Penelitian

Variabel	Minimum	Maksimum	Rata-rata	Jumlah Data
Produksi Padi (ton)	507	9526516	1610264	34
Luas Panen (ha)	179	1693211	307432	34
Produktivitas Padi (ku/ha)	28,24	60,59	45,82	34

3.2 Model Regresi Linier Berganda

Pada penelitian ini, hanya diambil data produksi padi, luas panen dan produktivitas padi di Indonesia tahun 2022. Tujuan dari penelitian ini adalah untuk mengetahui estimasi hubungan antara produksi padi terhadap luas panen dan produktivitas padi di Indonesia dengan menggunakan metode regresi *robust* Estimasi-M. Untuk melihat hubungan ketiganya, penulis menggunakan model regresi linier berganda.

Untuk menentukan estimasi model regresi terhadap data dilakukan dengan menduga parameter-parameter regresinya sehingga diperoleh estimasi model regresi sebagai berikut,

$$y = -480789 + 5,5317 x_1 + 8521 x_2$$

hasil analisis regresi di atas mengartikan bahwa:

- setiap kenaikan 1 hektar luas panen maka produksi padi akan bertambah sebesar 5,5317 ton,
- setiap kenaikan 1 kuintal per hektar produktivitas padi maka akan menambah produksi padi sawah sebesar 8521 ton.

3.3 Uji Asumsi Klasik

Uji asumsi klasik merupakan syarat statistik yang harus dipenuhi analisis regresi linear berganda yang berbasis *ordinary least square* (OLS). Selain itu, untuk mendapatkan model regresi linear berganda yang baik harus memenuhi kriteria *Best Linear Unbiased Estimator* (BLUE). BLUE dapat dicapai jika memenuhi asumsi klasik. Adapun uji klasik yang digunakan antara lain uji normalitas residu, uji asumsi non multikolinearitas, homoskedastisitas dan non autokorelasi. Hasil dai uji asumsi klasik diberikan pada Tabel 2.

Tabel 2. Hasil Uji Asumsi Klasik

Asumsi	Hasil Keputusan
Normalitas Residu	Terpenuhi
Non Multikolinearitas	Terpenuhi
Homoskedastisitas	Tidak Terpenuhi
Non Autokorelasi	Tidak Terpenuhi

3.4 Deteksi Pencilan

Berdasarkan statistik uji untuk mengetahui pencilan terhadap Y yaitu $TRES$ dengan menarik kesimpulan menolak H_0 apabila nilai $|TRES| > t_{n-k, \alpha} = 2.042$ maka diperoleh kesimpulan bahwa terdapat pencilan data ke-20 dan 27.

Berdasarkan statistik uji untuk mengetahui pencilan terhadap X yaitu hii yang dengan menarik kesimpulan bahwa pengamatan menolak H_0 apabila nilai $hii > \frac{2p}{n} = 0.117647$ maka diperoleh kesimpulan bahwa terdapat pencilan data ke-10, 12, 13, 15, 17, 20 dan 21.

Tabel 3. TRES dan Hii

No.	TRES	Hii	6	-0,31917	0,059197
1	0,099224	0,089534	7	0,188584	0,056902
2	-1,000308	0,040348	8	-1,007832	0,045484
3	-0,570609	0,044529	9	0,803094	0,044426
4	0,393189	0,039718	10	1,934179	0,172155
5	0,231438	0,040485	11	0,453065	0,062556

12	2,032611	0,271057	24	1,18891	0,075815
13	0,195223	0,284293	25	0,323728	0,039273
14	-0,018698	0,060218	26	-0,601503	0,032089
15	1,361679	0,283614	27	-2,797731	0,099739
16	-0,343752	0,058211	28	-0,27987	0,041885
17	0,183889	0,196463	29	0,179175	0,074539
18	-0,134146	0,070536	30	0,117327	0,06856
19	-0,914754	0,038355	31	0,803872	0,05291
20	-3,399786	0,162624	32	1,049292	0,055869
21	-0,331658	0,123701	33	0,717392	0,041732
22	-1,585958	0,058464	34	0,480315	0,050218
23	0,33585	0,064501			

3.5 Model Regresi Robust dengan Estimasi-M (Tukey Bisquare)

Proses perhitungan estimasi-M yang iteratif dimulai dengan menentukan estimasi awal koefisien regresi, yang diperoleh dari MKT yaitu $\hat{\beta}^0 = (-13608; 6.15; 222)$ kemudian berdasarkan algoritma Estimasi-M, dihitung nilai \hat{y}_i^0 dan sisa $\hat{e}_i^0 = y_i - \hat{y}_i^0$. Proses iterasi menggunakan MKT terboboti dilanjutkan dengan menghitung sisaan dan pembobot $w(u_i)$ yang baru dan dilakukan pendugaan parameter secara berulang-ulang sampai konvergen. Kekonvergenan tercapai jika koefisien regresi sudah sama dengan koefisien regresi sebelumnya [9].

Tabel 4. Iterasi Estimasi M

Iterasi	b_0	b_1	b_2
iterasi 0	- 480789	5,5317	8521
iterasi 1	-353633	5,5641	5869
iterasi 2	-286870	5,5828	4471
iterasi 3	-256293	5,5931	3846
iterasi 4	-244883	5,5979	3615
iterasi 5	-232937	5,5990	3569
iterasi 6	-242482	5,5993	3559
Iterasi 7	-242372	5,5994	3557
Iterasi 8	-242348	5,5994	3556
Iterasi 9	-242344	5,5994	3556
Iterasi 10	-242344	5,5994	3556
Iterasi 11	-242345	5,5994	3556

Iterasi 12	-242345	5,5994	3556
Iterasi 13	-242345	5,5994	3556
Konvergen Iterasi 12			
dengan $\sigma = -96820,2841$			

Berdasarkan Tabel 4. Iterasi Estimasi M di atas, terlihat koefisien regresi sudah konvergen pada iterasi ke-12, dengan model:

$$y = -242345 + 5,5994 x_1 + 3556 x_2$$

Hasil R-Sq(adj) sebesar 99,87%, sehingga interpretasi model yaitu sebesar 99,87% produksi padi di Indonesia tahun 2022 dapat diterangkan oleh variabel luas panen dan produktivitas. Setiap kenaikan satu hektar luas panen, akan menambah produksi padi di Indonesia sebesar 5,5994 ton. Setiap kenaikan satu kuintal per hektar produktivitas, akan menambah produksi padi di Indonesia sebesar 3556 ton.

3.6 Uji Signifikansi

3.6.1 Uji Signifikansi Serempak

Berdasarkan hasil regresi diperoleh $p\text{-value} = 0.000 < 0.05$ itu artinya paling tidak ada salah satu diantara luas panen atau produktivitas padi yang berpengaruh secara signifikan terhadap produksi padi.

Selanjutnya dilakukan uji parsial untuk mengetahui signifikansi atau pengaruh masing-masing variabel independen terhadap model regresi yang dihasilkan.

3.6.2 Uji Signifikansi untuk Masing-Masing Variabel

Akan diuji apakah setiap variabel signifikan masuk ke dalam model. Tabel 5 menunjukkan hasil uji t pada Estimasi-M untuk dua variabel independen.

Tabel 5. Uji signifikansi masing-masing variabel

Variabel Independen	p-value	Keterangan
X_1 (Luas Panen)	$0.000 < 0.05$	Koefisien regresi signifikan
X_2 (Produktivitas Padi)	$0.160 > 0.05$	Koefisien regresi tidak signifikan

Tabel 5 di atas menunjukkan bahwa luas panen mempengaruhi jumlah produksi padi sedangkan produktivitas padi tidak berpengaruh signifikan terhadap produksi padi di Indonesia tahun 2022.

Berdasarkan analisis di atas, berarti upaya pemerintah dalam meningkatkan produksi padi di Indonesia adalah dengan menambah luas panen. Selain itu, perlu pula meningkatkan produktivitas padi, misalnya dengan cara penggunaan kualitas benih padi yang baik, pemupukan serta perawatan yang teratur.

4. Kesimpulan

Berdasarkan pembahasan di atas maka dapat diambil kesimpulan sebagai berikut.

1. Model regresi *robust* dengan metode Estimasi-M adalah sebagai berikut :

$$y = -242345 + 5,5994 x_1 + 3556 x_2$$

artinya setiap kenaikan satu hektar luas panen, akan menambah produksi padi di Indonesia sebesar 5,5994 ton. Setiap kenaikan satu kuintal per hektar produktivitas, akan menambah produksi padi di Indonesia sebesar 3556 ton.

2. Berdasarkan analisis yang dilakukan diperoleh bahwa faktor yang berpengaruh terhadap besarnya produksi padi sawah di Kabupaten Klaten adalah luas panen.
3. Dari model regresi tersebut diperoleh koefisien determinasi $R\text{-Sq}(\text{adj}) = 99,87\%$, Hal ini menyatakan bahwa produksi padi di Indonesia tahun 2022 dipengaruhi oleh luas panen dan produktivitas padi, sisanya sebesar 0,13% dipengaruhi oleh variable lain.

Daftar Pustaka

- [1] Chen, Colin. 2002. Robust Regression and Outlier Detection With The ROBUSTREG Procedure, Sugi Paper 265-27, SAS Institute Inc., Cary, NC.
- [2] Draper, N.R. and Smith, H. 1998. *Applied Regression Analysis Third Edition*. Willey Interscience Publication, United States.
- [3] Gasul, K., K., Atti, A., dan Kleden, M., A. 2023. Model Regresi Robust dengan Metode Estimasi M, Estimasi S dan Estimasi MM untuk Produksi Beras di Nusa Tenggara Timur. *Estimasi*, Vol. 4, No. 1, Januari, 2023, Hal. 20-32
- [4] Gujarati, Damodar. 1978. *Ekonometrika Dasar*. Penerbit Erlangga, Jakarta.
- [5] Kurniawan, A., Susanti, Y., dan Pratiwi, H. 2023. Pemodelan Produksi Padi di Indonesia Menggunakan Regresi Robust Estimasi Generalized M. *Prosiding Pendidikan Matematika, Matematika dan Statistika Volume 7*.
- [6] Montgomery, D. C and Peck, E. A. 1992. *Introduction to Linear Regression Analysis Second Edition*. John Willey & Sons, Inc, New York.
- [7] Rasantaka, M., P., R., Ashshidiqi, M., F., Yulianti, R., Zeinawaqi, Z. dan Widodo, E. 2022. Implementasi Regresi Robust untuk Mengetahui Faktor-Faktor yang Memengaruhi Produksi Padi di Indonesia. *Jurnal Statistika dan Aplikasinya*. No. 2, Vol. 6, 234-242.
- [8] Sembiring, R.K. 2003. *Analisis Regresi*. Penerbit ITB, Bandung.
- [9] Yohai, V.J. 1987. High Breakdown Point and High Efficiency Robust Estimates for Regression, *Annals of Statistics*, Vol. 15, No. 20, pp. 642-656.