

PERBANDINGAN REGRESI ROBUST METODE *LEAST TRIMMED SQUARE* (LTS) DAN METODE ESTIMASI-S PADA PRODUKSI PADI DI KABUPATEN BLITAR

ENDAH SETYOWATI*, RACHMADANIA AKBARITA, RIZKA RIZQI ROBBY

*Program Studi S1 Matematika,
Fakultas Ilmu Eksakta, Universitas Nahdlatul Ulama Blitar, Indonesia
email : setyowatiendah5798@gmail.com, dania.barita@gmail.com, rizka.ertiga@gmail.com*

Diterima 12 Juni 2021 Direvisi 22 Juni 2021 Dipublikasikan 26 Juli 2021

Abstrak. Produksi padi di Kabupaten Blitar mengalami peningkatan dan penurunan, hal ini dipengaruhi oleh beberapa faktor, diantaranya jumlah petani, alokasi pupuk, rata-rata curah hujan, luas panen, luas tanam, produktivitas, dan alat pengolah padi. Oleh karena itu, untuk mengetahui faktor-faktor yang lebih signifikan tersebut, guna mencapai produksi padi yang optimal dapat digunakan analisis regresi. Namun, adanya data pencilan pada suatu data penelitian dapat mengganggu proses analisis data. Regresi robust merupakan metode yang efisien untuk menganalisis data yang mengandung pencilan. Regresi robust memiliki beberapa metode estimasi, dua diantaranya adalah *Least Trimmed Square* (LTS) dan Estimasi S yang memiliki persamaan karakteristik pada efisiensi dan *breakdown point*. Penelitian ini bertujuan untuk membandingkan kedua metode tersebut pada data produksi padi di Kabupaten Blitar tahun 2018 dengan tujuh variabel bebas (jumlah petani, alokasi pupuk, rata-rata curah hujan, luas panen, luas tanam, produktivitas, dan alat pengolah padi). Pengambilan data pada tahun 2018 didasarkan pada kelengkapan dokumen serta adanya kekhawatiran pandemi Covid-19 mempengaruhi data. Estimasi regresi robust menggunakan metode *Least Trimmed Square* (LTS) pada produksi padi di Kabupaten Blitar diperoleh model: $Y = -11262,756 - 0,01x_1 + 0,031x_2 - 14,304x_3 + 2,292x_4 + 3,741x_5 + 188,274x_6 - 0,419x_7$ dan estimasi regresi robust menggunakan metode Estimasi S pada produksi padi di Kabupaten Blitar diperoleh model: $Y = -9698,949 - 0,14x_1 - 0,49x_2 - 19,531x_3 + 0,133x_4 + 5,714x_5 + 175,018x_6 - 0,507x_7$. Hasil penelitian menunjukkan regresi robust metode *Least Trimmed Square* (LTS) merupakan metode yang menghasilkan model terbaik, karena metode *Least Trimmed Square* (LTS) memiliki nilai koefisien determinasi (R^2) sebesar 0,99999 yang lebih besar dibandingkan nilai koefisien determinasi (R^2) metode Estimasi S sebesar 0,99882, dan metode *Least Trimmed Square* (LTS) memiliki nilai *Mean Square Error* (MSE) sebesar 0,62105 yang lebih kecil dibandingkan nilai *Mean Square Error* (MSE) metode Estimasi S sebesar 9,04800.

Kata Kunci: Data Pencilan (*outlier*), Produksi Padi, Regresi Robust

*Penulis Korespondensi

1. Pendahuluan

Kabupaten Blitar merupakan salah satu wilayah di provinsi Jawa Timur yang terdiri dari 22 kecamatan dan memiliki potensi sumber daya alam yang beraneka ragam, terutama di sektor pertanian. Berdasarkan data yang dihimpun dari [1], produksi padi sawah dan padi ladang di Kabupaten Blitar pada tahun 2017 sebesar 423.485 ton berupa gabah kering panen, sedangkan pada tahun 2018 sebesar 380.687 ton (GKP). Produksi padi yang dihasilkan mengalami penurunan pada tahun 2013 ke 2014 sekitar 4,26 persen seiring menurunnya luas lahan yang turun sebesar 0,24 persen. Meningkat dan menurunnya produksi padi dipengaruhi oleh beberapa faktor. Menurut beberapa penelitian sebelumnya beberapa faktor yang mempengaruhi meningkat dan menurunnya produksi padi adalah luas tanam, jumlah pupuk, jumlah pompa air, kondisi lahan, produktivitas, kondisi perekonomian petani, curah hujan, jumlah petani dan alat pengolah padi. Oleh karena itu, untuk mengetahui faktor-faktor yang lebih signifikan tersebut, guna mencapai produksi padi yang optimal dapat digunakan analisis regresi.

Analisis regresi merupakan metode yang digunakan untuk menentukan hubungan antara variabel satu dengan satu atau lebih variabel bebas [2]. Untuk memperkirakan nilai variabel respon yang dipengaruhi oleh variabel bebas dilakukan dengan metode estimasi. Metode estimasi yang sering digunakan adalah metode *Ordinary Least Squares* (OLS) atau sering disebut kuadrat terkecil. Model terbaik yang dihasilkan pada metode ini harus memenuhi beberapa asumsi klasik yaitu, tidak terdapat multikolinearitas, tidak terdapat heteroskedastisitas, dan tidak terdapat autokorelasi. Selain itu, adanya **pencilan** (*outlier*) pada data akan mempengaruhi hasil dari metode *Ordinary Least Squares* (OLS) ini menjadi kurang efisien. Untuk mengatasi masalah pencilan pada data digunakan metode regresi robust. Menurut [2] ada beberapa metode yang digunakan untuk menaksir parameter regresi robust, diantaranya : *M-Estimator* , *MM-Estimator*, *S-Estimator*, *Least Median of Squares* (LMS)-*Estimator*, dan *Least Trimmed Squares* (LTS)-*Estimator*.

Penelitian yang dilakukan oleh [3] menyatakan metode estimasi S memiliki *breakdown point* 0,5 yang menjelaskan bahwa banyaknya pencilan hingga separuh data pengamatan tidak akan mempengaruhi model regresi yang dihasilkan. Selain itu penelitian mengenai perbandingan metode estimasi S dan MM dalam [4] menyimpulkan regresi robust estimasi S merupakan metode yang efektif dalam kasus data kemampuan membaca sekelompok anak. Sedangkan penelitian mengenai perbandingan metode estimasi LTS dan MM dilakukan oleh [5] menyatakan metode MM tidak lebih efisien dibanding metode LTS, karena metode LTS memiliki fungsi obyektif yang lebih halus sehingga akan lebih sensitif terhadap efek lokal dan memiliki nilai *breakdown point* yang tinggi. Oleh karena persamaan karakteristik metode estimasi S dan LTS pada efisiensi dan *breakdown point*, peneliti merasa tertarik untuk menggunakan metode tersebut dalam studi kasus produksi padi.

Berdasarkan latar belakang yang telah diuraikan sebelumnya, rumusan masalah yang dapat diambil dari penelitian ini adalah bagaimana model regresi robust dengan estimasi *Least Trimmed Squares* (LTS) dan estimasi S untuk produksi padi di Kabupaten Blitar dan bagaimana perbandingan keefektifan model regresi robust

antara metode *Least Trimmed Squares* (LTS) dan estimasi S untuk produksi padi di Kabupaten Blitar? Sehingga tujuan penelitian ini adalah mengetahui model regresi robust dengan estimasi *Least Trimmed Squares* (LTS) dan estimasi S untuk produksi padi di Kabupaten Blitar serta mengetahui perbandingan keefektifan model regresi robust antara metode *Least Trimmed Squares* (LTS) dan estimasi S untuk produksi padi di Kabupaten Blitar.

2. Kajian Pustaka

2.1. Produksi Padi

Produksi padi secara umum merupakan salah satu hasil yang diperoleh melalui penanaman bibit padi dan perawatannya sehingga menghasilkan padi yang dapat dimanfaatkan. Pada penelitian ini, yang dimaksud produksi padi adalah hasil gabah kering panen (GKP) padi sawah dan ladang. Produksi padi pada tahun 2018 lebih kecil dibandingkan dengan tahun sebelumnya, bila dicermati pada tahun 2018 ada penurunan jumlah produksi padi sebesar 27.764 (GKP) dari tahun sebelumnya yang meliputi padi sawah dan padi ladang.

2.2. Analisis Regresi

Analisis regresi merupakan metode yang digunakan untuk menentukan hubungan antara variabel satu dengan satu atau lebih variabel bebas [2]. Hubungan antara variabel respon dan variabel bebas disimpulkan dalam suatu model matematis guna mengetahui pola perubahan nilai variabel-variabel tersebut pada masa yang akan datang. Analisis regresi sederhana adalah analisis untuk mengetahui hubungan antara variabel respon dan satu variabel bebas. Namun, data suatu pengamatan sering kali tidak hanya diakibatkan oleh satu variabel bebas namun juga dapat diakibatkan oleh dua atau lebih variabel bebas yang sering disebut analisis regresi berganda. Menurut [6] regresi berganda adalah suatu teknik analisis statistik yang mempelajari hubungan antara sebuah **variabel terikat** (*dependent variable*) dengan beberapa **variabel bebas** (*independent variable*) melalui suatu **persamaan statistik** (*statistical equation*), yang sering juga disebut dengan **model statistik** (*statistical model*) yang berdasarkan prinsip hubungan atau fungsi statistik. Bentuk dasar dari model regresi berganda adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_j X_{ij} + \varepsilon_i,$$

dengan:

- Y_i : pengamatan ke i dari variabel terikat Y ,
- $\beta_0, \beta_1, \beta_2$: parameter koefisien regresi,
- X_{ij} : pengamatan ke i dari variabel bebas ke j ,
- β_j : koefisien parameter regresi dari variabel X_j ,
- ε_i : suku random atau suku galat dari pengamatan ke i ,
- i : banyaknya unit pengamatan ($i \in N$),
- j : banyaknya variabel independen ($j \in Z$).

2.3. Uji Asumsi Klasik

Menurut [4], untuk mendapatkan nilai periksa yang efisien dan tidak bias dari persamaan regresi berganda, maka perlu pengujian untuk mengetahui model regresi yang dihasilkan memenuhi persyaratan beberapa asumsi klasik, yaitu:

(1) Uji Normalitas

Uji normalitas bertujuan untuk mengetahui residual pada model regresi mempunyai distribusi normal atau tidak. Cara lain untuk mengetahui asumsi kenormalan adalah rasio *skewness* dan rasio kurtosis. Rasio *skewness* adalah nilai *skewness* dibagi dengan *standart error skewness*, sedangkan rasio kurtosis adalah nilai kurtosis dibagi dengan *standart error kurtosis*. Kriteria pengujian yang digunakan adalah jika rasio *skewness* dan kurtosis berada diantara -2 hingga +2, maka distribusi data adalah normal [7].

(2) Uji Heterokedastisitas

Uji Heterokedastisitas bertujuan untuk mengetahui adanya perbedaan varians variabel pada model regresi. Salah satu cara untuk menguji asumsi heterokedastisitas adalah uji Glejser, dengan statistik uji:

$$F_{hitung} = \frac{KTR}{KTG},$$

dengan

$$\begin{aligned} F_{hitung} &: \text{nilai uji Glejser,} \\ KTR &: \text{Kuadrat tengah regresi,} \\ KTG &: \text{Kuadrat tengah galat.} \end{aligned}$$

Kriteria yang digunakan dalam pengujian adalah $F_{hitung} > F(\alpha, k, n-k-1)$ maka tolak H_0 pada tingkat signifikansi α , artinya bahwa residual tidak identik [4].

(3) Uji Multikolinearitas

Uji multikolinearitas bertujuan untuk mengetahui adanya korelasi yang tinggi atau sempurna antar variabel independen. Salah satu cara untuk menguji asumsi multikolinearitas adalah dengan melihat nilai *variance inflation factor* (VIF), dengan statistik uji [4]:

$$VIF = \frac{1}{(1 - R^2)},$$

dengan

$$\begin{aligned} VIF &: \text{Variance inflation factor,} \\ R &: \text{Nilai korelasi antar variabel dependen} \end{aligned}$$

Kriteria yang digunakan dalam pengujian adalah $VIF > 10$, maka H_1 ditolak, artinya model memiliki masalah multikolinearitas.

(4) Uji Autokorelasi

Uji autokorelasi bertujuan mengetahui adanya korelasi linier antar residual pada

periode t dengan periode $t - 1$. Salah satu cara untuk uji asumsi autokorelasi adalah uji Durbin-Watson, dengan statistik uji:

$$DW = \frac{\sum (e - e_{t-1})^2}{\sum e^2},$$

dengan

DW : Nilai Durbin-Watson,

e : Nilai residual,

e_{t-1} : Nilai residual periode $t - 1$

Kriteria yang digunakan dalam pengujian adalah sebagai berikut [8]: dimana

DW	Kesimpulan
$< dL$	Ada autokorelasi positif
$dL - dU$	Ragu-ragu
$dU - (4 - dU)$	Tidak ada autokorelasi
$(4 - dU) - (4 - dL)$	Ragu-Ragu
$> (4 - dL)$	Ada autokorelasi negatif

Tabel 1. Kriteria Pengujian Autokorelasi dengan Uji Durbin-Watson

dL : Durbin *Lower* atau nilai batas bawah,

dU : Durbin *Upper* atau nilai batas atas.

2.4. Pencilan (*Outlier*)

Menurut Moore dan Mc Cabe dalam [2], *outlier* adalah objek yang terletak jauh atau berbeda jauh dari pola distribusinya. Secara umum data pencilan adalah data yang berbeda dengan data-data lainnya pada suatu hasil pengamatan, data ini nilainya bisa jadi lebih besar atau lebih kecil dari pengamatan pada umumnya. Data yang mengandung pencilan dapat mengganggu proses analisis data, namun pencilan ini kemungkinan mengandung pengaruh besar pada koefisien regresi sehingga keberadaannya tidak boleh dihilangkan begitu saja.

Metode yang dapat digunakan untuk mendeteksi adanya data pencilan yang berpengaruh terhadap parameter regresi, diantaranya: metode grafis, $DfBETA(s)$ dan *Cook's distance*. Namun kelemahan pada metode grafis adalah keputusan memperlihatkan data tersebut merupakan pencilan atau tidak bergantung pada kebijakan peneliti, karena pengamatannya hanya dengan visualisasi gambar. *Cook's Distance* merupakan metode pendeteksi data pencilan dengan menunjukkan besarnya pengaruh data pencilan terhadap semua estimator koefisien regresi. Suatu data yang mempunyai nilai jarak *Cook* lebih besar dari $\frac{4}{n}$ maka didefinisikan sebagai pencilan, dengan n adalah banyaknya observasi [8]. $DfBETA(s)$ digunakan untuk menyatakan pengaruh suatu pengamatan ke- i terhadap koefisien ke- k atau

digunakan untuk mendeteksi pencilan pada variabel bebas. Pengamatan ke- i berpengaruh pada model regresi jika

$$|DfBETA(s)| = \begin{cases} > 1, & n < 30, \\ > \frac{2}{\sqrt{n}}, & n \geq 30, \end{cases} \quad (2.1)$$

dimana n adalah jumlah sampel.

2.5. Regresi Robust

Menurut [10], regresi robust merupakan metode regresi yang digunakan ketika distribusi dari residual tidak normal atau adanya beberapa pencilan yang berpengaruh pada model. Regresi robust memiliki beberapa metode, antara lain: *M-Estimator*, *MM-Estimator*, *S-Estimator*, *Least Median of Squares (LMS)-Estimator*, dan *Least Trimmed Squares (LTS)-Estimator*. Kelima metode regresi robust tersebut masing-masing memiliki kelemahan dan kelebihan. *MM-Estimator* memiliki efisiensi yang tinggi, tetapi nilai *breakdown point* 0. *LMS*, *LTS* dan *S-Estimator* mempunyai *breakdown point* yang tinggi yaitu 0,5, akan tetapi efisiensinya yang rendah.

(a) Least Trimmed Square (LTS)

Least Trimmed Square (LTS) adalah metode estimasi parameter regresi robust yang memiliki nilai *breakdown point* yang tinggi dengan cara meminimalkan jumlah kuadrat h residual. Menurut [11], prosedur estimasi dengan menggunakan *Least Trimmed Square (LTS)* adalah sebagai berikut:

- (i) Mengestimasi koefisien regresi menggunakan MKT.
- (ii) Menentukan n residual $e_i^2 = (Y_i - X_i\beta_0)^2$ yang bersesuaian dengan (β_0) , kemudian menghitung jumlah $h_0 = \lfloor \frac{n}{2} \rfloor + \lfloor \frac{k+1}{2} \rfloor$ pengamatan dengan nilai e_i^2 terkecil.
- (iii) Menghitung $\sum_i^h e_i^2$
- (iv) Mengestimasi parameter β_{baru} dari β_0 observasi
- (v) Ditentukan n kuadrat residual $e_i^2 = (Y_i - X_i\beta_0)^2$ yang bersesuaian dengan β_{baru} kemudian menghitung sejumlah β_{baru} observasi dengan e_i^2 terkecil
- (vi) Menghitung $\sum_i^{h_{baru}} e_i^2$
- (vii) Melakukan iterasi yaitu tahap (iv) sampai (vi) untuk mendapatkan fungsi objektif yang kecil dan konvergen,

dimana:

- n : banyaknya pengamatan,
- e_i^2 : kuadrat residual,
- β_0 : parameter koefisien regresi,
- h_0 : jumlah data yang digunakan pada iterasi selanjutnya,
- Y_i : variabel dependen ke $-i$,

- X_i : variabel independen ke $-i$,
 k : banyaknya parameter,
 β_{baru} : koefisien parameter regresi yang baru,
 i : banyaknya unit pengamatan, $i \in N$

(b) Estimasi S

Estimasi S adalah metode estimasi parameter regresi dengan jumlah kuadrat *error* pada persamaan umum regresi linier. Menurut [12], prosedur estimasi S adalah sebagai berikut:

- (i) Mengestimasi koefisien regresi menggunakan MKT
 (ii) Menghitung nilai residual $e_i = Y_i - \hat{Y}_i$
 (iii) Menghitung nilai estimasi skala robust Standar deviasi sisaan σ_s :

$$\sigma_s = \sqrt{\frac{n \sum_{i=1}^n (e_i^2) - (\sum_{i=1}^n e_i)^2}{n(n-1)}}$$

- (iv) Menghitung nilai $u_i = \frac{e_i}{\sigma_s}$
 (v) Menghitung nilai fungsi pembobot w_i :

$$w_i = \begin{cases} (1 - \frac{u_i^2}{c^2})^2, & |u_i| < c \\ 0, & |u_i| \geq c \end{cases}$$

- (vi) Nilai $c = 1,547$ adalah suatu konstanta yang ditetapkan untuk memberikan efisiensi estimasi sebesar 95% pada fungsi pembobot Tukey Bisquare.
 (vii) Mengestimasi nilai β_s
 (viii) Melakukan langkah (ii) sampai (vi) sehingga diperoleh nilai β_s yang konvergen,

dimana:

- e_i : nilai residual,
 Y_i : variabel dependen ke $-i$,
 \hat{Y}_i : variabel dependen ke $-i$ baru,
 σ_s : standar deviasi sisaan,
 n : banyaknya pengamatan,
 w_i : nilai pembobot data ke $-i$,
 c : nilai fungsi pembobot Tukey Bisquare,
 β_s : koefisien parameter regresi yang baru,
 i : banyaknya unit pengamatan, ($i \in N$)

2.6. Pemilihan Model Terbaik

Pemilihan model terbaik didasarkan pada nilai koefisien determinasi (R^2) dan *Mean Square Error* (MSE). Nilai koefisien determinasi (R^2) memberikan gambaran tentang kesesuaian variabel bebas dalam memprediksi variabel terikat. Kriteria pengujian koefisien determinasi (R^2) adalah jika nilainya semakin besar atau dengan kata lain mendekati 1, maka semakin baik model regresinya [13].

Mean Square Error (MSE) merupakan salah satu ukuran yang sering digunakan untuk menguji kualitas model regresi. MSE adalah jumlah kuadrat kesalahan dan dibagi dengan banyaknya periode waktu n . Menurut [13], semakin kecil nilai MSE maka semakin baik kecocokan suatu model dengan data karena kriteria MSE akan menghasilkan titik optimal yang mendekati target dengan variansi terkecil.

3. Pembahasan

3.1. Data dan Analisis Data

Data yang akan dianalisis dalam penelitian ini adalah data sekunder dari 22 kecamatan di Kabupaten Blitar pada tahun 2018. Data tersebut diantaranya: produksi padi (y), jumlah petani (x_1), pupuk (x_2), curah hujan (x_3), luas panen (x_4), luas tanam (x_5), produktifitas (x_6), dan alat pengolah padi (x_7). Data yang telah terkumpul akan diuji asumsi klasik. Selanjutnya akan diperiksa apakah data tersebut terdapat *outlier*. Jika dideteksi adanya *outlier*, maka digunakan regresi *robust* metode *Least Trimmed Square* (LTS) dan estimasi S. Kedua metode estimasi tersebut akan dibandingkan keefektifannya berdasarkan nilai koefisien determinasi (R^2) dan *Mean Square Error* (MSE).

3.2. Uji Asumsi Klasik

- (a) Uji normalitas Berdasarkan kriteria pengujian, pada tabel di atas da-

Descriptive Statistics					
	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
Standardized Residual	22	.098	.491	-.765	.953
Valid N (listwise)	22				

Tabel 2. Uji Normalitas Menggunakan Metode Skewness dan Kurtosis

pat dihitung rasio $skewness = \frac{0,098}{0,491} = 0,199$, sedangkan rasio $kurtosis = -\frac{0,765}{0,953} = -0,802$. Karena nilai rasio $skewness$ dan $kurtosis$ berada diantara $-1,96$ sampai $+1,96$ maka dapat disimpulkan bahwa data berdistribusi normal.

- (b) Uji Heterokedastisitas

Berdasarkan kriteria pengujian, data tidak terjadi masalah heteroskedastisitas jika nilai signifikansi antara variabel independen dengan absolut residual lebih besar dari 0,05. Pada tabel uji heteroskedastisitas menggunakan uji gejser terlihat nilai signifikansi variabel $x_1 = 0,947, x_2 = 0,556, x_3 = 0,092, x_4 = 0,660, x_5 = 0,717, x_6 = 0,066$ dan $x_7 = 0,499$ yang semuanya diatas 0,05 dapat disimpulkan bahwa seluruh variabel independen memiliki sebaran varian yang sama atau homogen.

(c) Uji multikolinearitas

Coefficients^a

Model		Collinearity	Statistics
		Tolerance	VIF
1	x1	.763	1.311
	x2	.747	1.338
	x3	.812	1.232
	x4	.001	966.455
	x5	.001	979.796
	x6	.609	1.642
	x7	.835	1.197

Tabel 3. Uji Multikolinearitas Menggunakan VIF

^a: Dependent Variable: y

Berdasarkan kriteria pengujian, variabel independen terjadi masalah multikolinearitas jika nilai $VIF > 10$. Dari tabel diatas terlihat semua variabel memiliki nilai $VIF < 10$, kecuali variabel x_4 dan x_5 yang memiliki nilai VIF lebih dari 10, yaitu 966,455 dan 979,796. Dapat disimpulkan bahwa pada model ini terjadi masalah multikolinearitas.

(d) Uji autokorelasi

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.997 ^a	.994	.991	821.739	2.283

Tabel 4. Uji Autokorelasi Menggunakan Durbin Waston

^a : Predictors: (Constant) : $x_7, x_1, x_3, x_4, x_2, x_6, x_5$

^b : Dependent Variable : y.

Nilai dL dan dU ditentukan dengan cara melihat pada tabel menggunakan $\alpha = 0,05$, sampel (n) sebanyak 22, dan variabel independen sebanyak 7 maka didapatkan nilai dL dan dU adalah 0,67719 dan 2,24646. Pada tabel terlihat nilai Durbin-Waston sebesar 2,283. Bandingkan nilai dL dan dU dengan nilai

Durbin-Waston, maka didapat nilai Durbin-Waston memenuhi kriteria $4 - dU < Durbin - Waston < 4 - dL$ yaitu $1,75354 < 2,283 < 3,32281$, maka dapat disimpulkan pengujian tidak meyakinkan sehingga tidak dapat disimpulkan ada atau tidaknya autokorelasi antara residual.

3.3. Pendeteksian Pencilan (Outlier)

Pendeteksian pencilan dilakukan menggunakan metode *cook's distance* dan *DfBETA(s)*. Pada metode *cook's distance* data diduga sebagai pencilan apabila nilai *cook's distance* $> \frac{4}{n}$, dengan $n = 22$ yang merupakan banyaknya data, maka diperoleh nilai $\frac{4}{n} = 0,18181$. Nilai *cook's distance* diperoleh dengan bantuan SPSS 16.0. terlihat data ke-5, ke-6, ke-15 dan ke-17 memiliki nilai *cook's distance* $> 0,18181$ yaitu data ke-5 = 0,51576, ke-6 = 0,18225, ke-14 = 0,20018, ke-15 = 0,45847 dan ke-17 = 1,73035. Dapat disimpulkan bahwa data ke-5, ke-6, ke-14, ke-15 dan ke-17 merupakan pencilan (*outlier*).

Pada metode *DfBETA(s)* digunakan kriteria pengujian pada persamaan 2.1, pada penelitian ini jumlah sampel (n) = 22, jadi suatu data diduga sebagai pencilan jika nilai *DfBETA(s)* > 1 . Nilai *DfBETA(s)* diperoleh dengan bantuan SPSS 16.0. Perhatikan bahwa beberapa data (seperti: data ke-5, ke-9, ke-13, ke-14, ke-15, ke-16, dan ke-20) berpengaruh untuk lebih dari satu estimasi. Dapat disimpulkan bahwa data ke-5, ke-9, ke-13, ke-14, ke-15, ke-16, dan ke-20 merupakan data pencilan (*outlier*).

3.4. Regresi Robust Metode Least Trimmed Square (LTS)

Estimasi parameter analisis regresi robust metode *Least Trimmed Square* (LTS) dilakukan dengan menggunakan bantuan *software* SPSS 16.0 dan *Microsoft Excel* 2016 untuk proses iterasinya. Langkah iterasi dengan metode *Least Trimmed Square* (LTS) dapat dilihat pada Tabel 5 berikut.

Iterasi	h_i	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	$\sum_i^h e_i^2$
MKT	22	-19419,711	0,071	0,338	-9,241	12,076	-5,517	284,723	-1,153	9.454.375,94
1	15	-13104,586	-0,0007	0,063	-13,098	3,074	3,004	212,268	-0,742	505.132,01
2	12	-13489,145	-0,006	0,095	-17,421	3,716	2,476	213,897	-0,523	115.316,89
3	10	-11874,006	-0,008	0,047	-18,239	2,682	3,413	195,118	-0,401	17.102,60
4	9	-11262,756	-0,01	0,031	-14,304	2,292	3,741	188,274	-0,419	1.084,66

Tabel 5. Iterasi Metode LTS

Iterasi 4 merupakan iterasi terakhir, hal ini dikarenakan nilai $h_0 = 9$ sama dengan iterasi selanjutnya. Dari Tabel 5 di atas diperoleh model regresi robust metode *Least Trimmed Square* (LTS) untuk produksi padi yaitu:

$$Y = -11262,756 - 0,01x_1 + 0,031x_2 - 14,304x_3 + 2,292x_4 + 3,741x_5 + 188,274x_6 - 0,419x_7$$

3.5. Regresi Robust Metode Estimasi S

Sama halnya dengan metode *Least Trimmed Square* (LTS), metode estimasi S ini juga memerlukan beberapa iterasi dalam pengerjaannya. Langkah iterasi dengan metode estimasi S dapat dilihat pada Tabel 6 berikut.

Iterasi	σ	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
MKT	670,95	-19419,711	0,071	0,338	-9,241	12,076	-5,517	284,723	-1,153
1	815,64	-14614,424	-0,00027	0,149	-13,103	4,710	1,480	227,809	-0,734
2	1.030,59	-10644,333	-0,012	-0,014	-17,789	0,971	4,950	184,698	-0,531
3	1.666,72	-9698,949	-0,14	-0,49	-19,531	0,133	5,714	175,018	-0,507

Tabel 6. Iterasi Metode Estimasi S

Iterasi ke 3 merupakan iterasi terakhir, hal ini dikarenakan nilai $w_{i4} = 0$ dan mengakibatkan analisis parameter tidak bisa dijalankan lagi. Dari Tabel 6 di atas diperoleh model regresi robust metode estimasi S untuk produksi padi yaitu:

$$Y = -9698,949 - 0,14x_1 - 0,49x_2 - 19,531x_3 + 0,133x_4 + 5,714x_5 + 175,018x_6 - 0,507x_7$$

3.6. Pemilihan Metode Terbaik

Tahapan pemilihan metode terbaik digunakan untuk mengetahui metode estimasi yang memberikan hasil yang lebih baik, maka kriteria yang digunakan adalah dengan membandingkan nilai koefisien determinasi (R^2) dan *Mean Square Error* (MSE).

No	Metode Regresi Robust	Koefisien Determinasi (R^2)	Mean Square Error (MSE)
1	<i>Least Trimmed Square</i> (LTS)	0,99999	0,62105
2	Estimasi S	0,99882	9,04800

Tabel 7. Perbandingan Nilai Koefisien R^2 dan MSE

Dari Tabel 7 di atas dapat dilihat bahwa nilai koefisien determinasi (R^2) dari metode *Least Trimmed Square* (LTS) lebih besar daripada metode Estimasi S dan nilai *Mean Square Error* (MSE) dari metode *Least Trimmed Square* (LTS) lebih kecil dibandingkan metode Estimasi S. Dengan melihat perbandingan nilai koefisien determinasi (R^2) dan *Mean Square Error* (MSE) pada kedua metode estimasi tersebut dapat disimpulkan bahwa dalam mengestimasi parameter data pada produksi padi di Kabupaten Blitar metode *Least Trimmed Square* (LTS) memiliki hasil yang lebih baik dibandingkan dengan metode Estimasi S.

4. Kesimpulan

Berdasarkan tinjauan pustaka, hasil penelitian dan pembahasan dapat disimpulkan sebagai berikut.

- (1) Estimasi regresi robust menggunakan metode *Least Trimmed Square* (LTS) pada produksi padi di Kabupaten Blitar diperoleh model:

$$Y = -11262,756 - 0,01x_1 + 0,031x_2 - 14,304x_3 + 2,292x_4 + 3,741x_5 + 188,274x_6 - 0,419x_7$$

Dari estimasi regresi robust menggunakan metode Estimasi S pada produksi padi di Kabupaten Blitar diperoleh model:

$$Y = -9698,949 - 0,14x_1 - 0,49x_2 - 19,531x_3 + 0,133x_4 + 5,714x_5 + 175,018x_6 - 0,507x_7$$

- (2) Tingkat keefektifan kedua metode dilihat dari nilai koefisien determinasi (R^2) dan *Mean Square Error* (MSE). Berdasarkan analisis dan pembahasan maka dapat disimpulkan regresi robust metode *Least Trimmed Square* (LTS) merupakan metode yang menghasilkan model terbaik, karena metode *Least Trimmed Square* (LTS) memiliki nilai koefisien determinasi (R^2) sebesar 0,99999 yang lebih besar dibandingkan nilai koefisien determinasi (R^2) metode Estimasi S sebesar 0,99882, dan metode *Least Trimmed Square* (LTS) memiliki nilai *Mean Square Error* (MSE) sebesar 0,62105 yang lebih kecil dibandingkan nilai *Mean Square Error* (MSE) metode Estimasi S sebesar 9,04800.

Daftar Pustaka

- [1] Badan Pusat Statistika Kabupaten Blitar, 2019, Blitar Dalam Angka, Blitar: BPS Kabupaten Blitar
- [2] Z. Aflakah, Jajang dan A. T. B. Sb, 2019, *Least Square* dan Robust Estimasi M pada Model Regresi Linier Sederhana yang Memuat Outlier, *Jurnal Ilmiah Matematika dan Pendidikan Matematika* (JMP) Vol. 11 No. 1, pp. 21 – 32
- [3] Y. Susanti, N. Qona'ah, K. Ferawati dan C. Qumillaila, 2020, Prediction Modelling Of Annual Parasite Incidence (API) of Malaria in Indonesia Using Robust Regression of M-Estimation and S-Estimation, dalam *AIP Conferences Proceeding 2296*, Amerika
- [4] A. Semar, F. Virgantari dan H. Wijayanti, 2020, Perbandingan Estimasi S (Scale) dan Estimasi MM (Method Of Moment) pada Model Regresi Robust dengan Data Pencilan, *Jurnal Statistika dan Matematika* (STATMAT), pp. 21 – 33
- [5] A. Shodiqin, A. N. Aini dan M. R. Rubowo, 2018, Perbandingan Dua Metode Regresi Robust Yakni Metode Least Trimmed Square (LTS) dengan Metode Estimator MM (Estimasi MM) (Studi Kasus Data Ujian Tulis Masuk Terhadap IPK Mahasiswa UPGRIS), *Jurnal Ilmiah Teknosains*, pp. 35 – 42
- [6] A. Asra, M. A. N. H. Pusponegoro dan A. P. Utomo, 2017, *Analisis Multivariabel : Suatu Pengantar*, In Media
- [7] A. Setyadharma, 2010, *Uji Asumsi Klasik dengan SPSS 16.0*, Universitas Negeri Semarang

- [8] L. S. Febrianto, N. K. Dwijayati dan P. Hendikawati, 2018, Perbandingan Metode Robust Least Median Of Square (LMS) dan Penduga S untuk Menangani Outlier pada Regresi Linier Berganda, *UNNES Journal Mathematic* (UJM)
- [9] E. Widodo dan A. A. Dewayanti, 2016, Perbandingan Metode Estimasi LTS, Estimasi M, dan Estimasi MM pada Regresi Robust, Yogyakarta: Universitas Islam Indonesia
- [10] F. P. Hidayatulloh, D. Yuniarti dan S. Wahyuningsih, 2015, Regresi Robust dengan Metode Estimasi S, *Jurnal EKSPONENSIAL*, pp. 163 – 170
- [11] M. B. Rahman dan E. Widodo, 2018, Perbandingan Metode Regresi Robust Estimasi Least Trimmed Square, Estimasi S dan Estimasi Method Of Moment, dalam *Prosiding Seminar Nasional Matematika (PRISMA)*, Semarang
- [12] P. WD dan L. E, 2016, Perbandingan Metode Estimasi-M, Estimasi-S, dan Estimasi-MM pada Model Regresi Robust untuk Memprediksi Produksi Kedelai di Indonesia, *Jurnal Pendidikan Matematika dan Sains*, Vol. 5 No. 5
- [13] A. Rohmawati, N. K. Dwijayanti dan Sugiman, 2018, Perbandingan Metode Least Trimmed Square (LTS) dan Scale (S) pada Response Surface Methodology, dalam *Prosiding Seminar Nasional Matematika (PRISMA)*, Semarang