

Pemodelan Faktor-Faktor yang Mempengaruhi Produksi Padi di Jawa Timur Tahun 2012 dengan Kasus Pencilan dan Autokorelasi Error

Ria Kumala Dewi dan Wiwiek Setya Winahju
 Statistika, FMIPA, Institut Teknologi Sepuluh Nopember (ITS)
 Jl. Arief Rahman Hakim, Surabaya 60111
 E-mail: wiiwiek.statistika@gmail.com

Abstrak—Jawa Timur merupakan daerah penghasil padi terbe-sar kedua di Pulau Jawa setelah Jawa Barat. Pada tahun 2011, produksi padinya mengalami penurunan sebesar 1,07 ton diban-dingkan dengan tahun 2010. Penurunan produksi padi ini dise-babkan oleh penurunan luas areal panen, produktivitas padi, dan perubahan fungsi lahan pertanian untuk pengembangan in-dustri, dan sebagainya akibat dari bertambahnya penduduk. Se-lama ini, pengukuran produktivitas dilakukan dengan membagi hasil produksi dengan luas areal panen. Pengukuran ini belum cukup menggambarkan performansi tanaman padi yang sebe-narnya, kecuali dengan memperhitungkan pula faktor lainnya. Agar diketahui faktor-faktor potensial untuk peningkatan pro-duksi padi, diperlukan penelitian yang lebih seksama dalam pe-ngukuran produktivitas tanaman padi di Jawa Timur. Dalam penelitian ini, digunakan regresi OLS (*Ordinary Least Square*) untuk mendapatkan model produksi padi di Jawa Timur. Dari hasil pemeriksaan asumsi residual, terdapat data pencilan (*out-lier*) dan autokorelasi antar residual. Sehingga, digunakan regre-si robust dengan estimasi M untuk mengatasi data pencilan dan GLS (*Generalized Least Square*) untuk mengatasi penyimpangan asumsi autokorelasi antar residual. Hasil model terbaik menun-jukkan bahwa variabel yang berpengaruh signifikan terhadap produksi padi di Jawa Timur adalah luas panen (X_1), luas puso (X_2), dan penggunaan pupuk (X_3) dengan nilai koefisien determi-nasi (R^2) sebesar 99,3%.

Kata kunci—Produksi padi di Jawa Timur, Regresi OLS, Regresi Robust dengan Estimasi M, GLS

I. PENDAHULUAN

Beralihnya lahan pertanian di Jawa Timur menjadi lahan non pertanian menyebabkan penurunan hasil produksi padi. Pada tahun 2011, produksi padinya sebesar 10,58 juta ton mengalami penurunan dibandingkan dengan tahun 2010 [1]. Jawa Timur sebagai provinsi yang memiliki sumba-ngan cukup besar terhadap produksi padi nasional perlu melaku-kan intensifikasi pertanian pada lahan yang tersedia dengan meningkatkan produktivitas tanaman padi. Selama ini, peng-uukuran produktivitas dilakukan dengan membagi hasil produksi dengan luas areal panen. Pengukuran ini belum cukup meng-gambarkan performansi tanaman padi yang sebenarnya, ke-cuali dengan memperhitungkan pula faktor lainnya. Sehingga, diperlukan suatu penelitian yang lebih seksama dalam peng-uukuran produktivitas tanaman padi agar diketahui faktor-faktor yang potensial berpengaruh.

[2] telah melakukan penelitian mengenai faktor-faktor yang mempengaruhi produksi padi di Jawa Timur. Hasil penelitiannya menunjukkan bahwa variabel yang mempenga-ruhi produktivitas tanaman padi selama periode tahun 1995-2000 adalah luas areal tanam, tenaga kerja, pupuk, dan benih. Penelitian serupa juga pernah dilakukan oleh [3], hasil peneli-tiannya menyimpulkan terdapat tiga pengelompokan wilayah berdasarkan luas lahan pertanian dan produksi padi, serta vari-abel yang mempengaruhi produksi padi sawah Jawa Timur ta-hun 2002 adalah jumlah penggunaan pupuk dan harga gabah.

Pada penelitian ini, digunakan metode OLS (*Ordinary Least Square*) untuk mengetahui faktor-faktor yang mempe-ngaruhi produksi padi di Jawa Timur tahun 2012. Pada metode OLS, terdapat asumsi klasik yang harus dipenuhi. Jika terda-pat pelanggaran terhadap asumsi tersebut, maka metode OLS tidak dapat digunakan. Oleh karena itu, diperlukan metode lain untuk mengatasi pelanggaran terhadap asumsi tersebut. Penelitian ini akan membahas mengenai metode regresi robust yang digunakan untuk mengatasi adanya data pencilan dan metode GLS (*Generalized Least Square*) untuk mengatasi adanya autokorelasi antar residual pada faktor-faktor yang mempengaruhi produksi padi di Jawa Timur tahun 2012.

II. TINJAUAN PUSTAKA

A. Regresi Linear Berganda

Regresi linear berganda adalah teknik statistika yang digu-nakan untuk menentukan model hubungan satu variabel respon (Y) dengan melibatkan lebih dari satu variabel prediktor hingga p variabel prediktor dimana banyaknya p kurang dari jumlah observasi (n). Adapun model regresi berganda sebagai berikut [4].

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \quad (1)$$

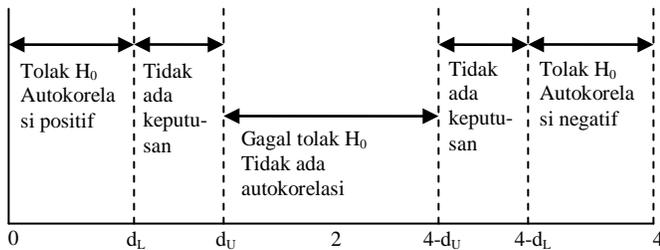
dimana Y_i merupakan nilai variabel dependen dalam penga-matan ke- i , $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ adalah parameter yang tidak dike-tahui nilainya, $X_{i1}, X_{i2}, \dots, X_{ip}$ adalah nilai dari variabel inde-penden dari pengamatan ke- i , dan ε_i adalah error random dan berdistribusi normal dengan rata-rata nol dan varians σ^2 .

Persamaan (1) dapat ditransformasikan ke dalam bentuk matriks sebagai berikut.

$$Y = X\beta + \varepsilon \quad (2)$$

Metode estimasi parameter yang digunakan adalah metode OLS (*Ordinary Least Square*), yaitu menduga koefisien regre-si (β) dengan meminimumkan kesalahan (error) [4]. Adapun penaksir parameternya adalah sebagai berikut.

(3)



Gambar 1 Kriteria Uji Durbin-Watson

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

dimana $\hat{\beta}$ adalah vektor dari parameter yang ditaksir $(p+1) \times 1$, X adalah matriks variabel bebas ukuran $n \times (p+1)$, dan Y adalah vektor observasi dari variabel respon berukuran $n \times 1$. Dan berikut merupakan bentuk penaksir variannya.

$$var(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \quad (4)$$

Uji signifikansi parameter dilakukan secara serentak dan parsial. Berikut merupakan hipotesis untuk uji signifikansi parameter secara serentak.

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_p = 0$$

$$H_1: \text{paling tidak terdapat satu } \beta_j \neq 0, j=1,2,3,\dots,p$$

Statistik uji yang digunakan adalah $F = \frac{MSR}{MSE}$. Tolak H_0 jika nilai $F > F_{\alpha; v_1, v_2}$ dimana $v_1 = p$ dan $v_2 = n - p - 1$.

Hipotesis untuk uji signifikansi parameter secara parsial adalah sebagai berikut.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, j=1,2,3,\dots,p$$

Statistik uji yang digunakan adalah $t = \frac{\hat{\beta}_j}{stdev(\hat{\beta}_j)}$. Tolak H_0 jika

$$|t_{hitung}| > t_{\alpha/2, (n-p-1)}$$

Dalam analisis regresi, terdapat asumsi residual yang harus dipenuhi antara lain antar residual tidak terjadi autokorelasi (independen), varian error pada setiap nilai variabel bebas harus sama (identik), dan residual berdistribusi normal. Sehingga, perlu dilakukan uji asumsi residual. Uji Durbin-Watson dapat digunakan untuk mendeteksi adanya autokorelasi. Berikut merupakan hipotesis uji asumsi residual independen [5].

$$H_0: \text{residual independen}$$

$$H_1: \text{residual tidak independen}$$

Statistik uji yang digunakan adalah $d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n (e_t)^2}$. Tolak H_0 apabila nilai d berada dalam interval d_U dan $4 - d_U$. Kriteria uji Durbin-Watson secara lengkap dapat dilihat pada Gambar 1.

Uji glejser dapat digunakan untuk mendeteksi adanya varian error yang heterogen. Berikut ini merupakan hipotesis uji asumsi residual identik [5].

$$H_0: \text{residual identik}$$

$$H_1: \text{residual tidak identik}$$

Statistik uji yang digunakan adalah $|U_t| = \beta X_t + vt$ dimana $|U_t|$ merupakan nilai absolut residual (ϵ) dan X_t adalah variabel bebas. Tolak H_0 apabila nilai $p\text{-value} < \alpha$ yang artinya varian error tidak homogen.

Pemeriksaan asumsi residual berdistribusi normal dapat menggunakan uji Kolmogorov-Smirnov dan berikut merupakan hipotesisnya [6].

$$H_0: \text{residual berdistribusi normal}$$

$$H_1: \text{residual tidak berdistribusi normal}$$

$$D = \sup_x |S(x) - F_0(x)|$$

Statistik uji yang digunakan adalah $S(x)$ merupakan sebaran kumulatif sampel dan $F_0(x)$ merupakan sebaran kumulatif normal. Tolak H_0 apabila $D > D_{(\alpha, n)}$ atau nilai $p\text{-value} < \alpha$.

Pada analisis regresi tidak boleh terjadi multikolinieritas. Multikolinieritas adalah terjadinya hubungan linier antar variabel bebas dalam suatu regresi linier berganda [5]. VIF (Variance Inflation Factor) dapat digunakan untuk mendeteksi adanya multikolinieritas dengan rumus sebagai berikut [7].

$$VIF(i) = \frac{1}{1 - R^2(i)} \quad (5)$$

dengan $R^2(i)$ merupakan nilai koefisien determinasi hasil regresi antara variabel X_i dengan variabel X lainnya. Jika nilai $VIF(i) > 10$, maka terdapat indikasi multikolinieritas.

B. Regresi Robust

Regresi robust merupakan metode regresi yang digunakan apabila distribusi dari residual tidak normal dan atau mengandung pencilan yang berpengaruh pada model [7]. Tujuan dari regresi robust adalah untuk memperoleh model terbaik yang robust atau resisten terhadap pencilan. Metode estimasi parameter yang digunakan adalah estimasi M, yaitu dengan menggunakan fungsi Huber untuk meminimalkan fungsi galatnya. Adapun fungsi Huber adalah sebagai berikut.

$$\psi(e_i^*) = \begin{cases} e_i^* & |e_i^*| \leq r \\ = r & e_i^* > r \\ = -r & e_i^* < -r \end{cases} \quad (6)$$

dimana r adalah permisalan suatu nilai tertentu (biasanya 1, 1,5 atau 2).

Bentuk penaksir parameter regresi robust adalah sebagai berikut.

$$b = (X^T W X)^{-1} X^T W Y \quad (7)$$

dengan X adalah matrik variabel bebas berukuran $n \times (p+1)$, W adalah matrik diagonal berukuran $n \times n$ dengan w_i sebagai elemen diagonalnya, dan Y adalah vektor observasi dari variabel respon berukuran $n \times 1$. Pembobot dalam estimasi M bergantung pada residual dan koefisien. Prosedur iterasi yang digunakan adalah IRLS (Iteratively Reweighted Least Squares) atau melakukan weighted least squares secara iterasi sampai residual konvergen [8].

C. Generalized Least Square (GLS)

Metode GLS adalah metode yang digunakan apabila terdapat autokorelasi antar residual. GLS merupakan regresi OLS dengan transformasi variabel yang dapat memenuhi asumsi standar least square [5]. Misalkan model regresi untuk satu prediktor sebagai berikut.

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t \quad (8)$$

dan diasumsikan residual mengikuti AR(1),

$$\epsilon_t = \rho \epsilon_{t-1} + u_t, \quad -1 < \rho < 1 \quad (9)$$

sehingga, model GLS untuk satu prediktor dapat dinotasikan sebagai berikut.

$$Y_t^* = \beta_0^* + \beta_1^* X_t^* + u_t \quad (10)$$

dengan $Y_t^* = Y_t - \rho Y_{t-1}$, $\beta_0^* = \beta_0(1 - \rho)$, $\beta_1^* = \beta_1$, $X_t^* = X_t - \rho X_{t-1}$, dan $u_t = \epsilon_t - \rho \epsilon_{t-1}$. Setelah variabel persamaan (8) ditransformasi ke dalam persamaan (10), dilakukan regresi dengan metode OLS. Dengan kata lain, GLS merupakan regresi OLS dengan transformasi variabel respon dan prediktor.

Jika nilai ρ tidak diketahui, maka perlu melakukan estimasi ρ . Salah satu metode yang digunakan untuk mencari nilai

estimasi ρ yaitu *Cochrane-Orcutt Iterative Procedure*. Prinsip dari metode tersebut adalah meregresikan residual ϵ_t dengan ϵ_{t-1} hingga diperoleh nilai ρ yang konvergen. Apabila jumlah observasi yang digunakan sedikit, maka perlu dilakukan transformasi Prais-Winsten pada pengamatan pertama [5].

D. Tanaman Padi

Tanaman padi dapat hidup dengan baik di daerah yang berhawa panas dan banyak mengandung uap air. Padi dapat ditanam di dataran rendah sampai ketinggian 1300 m di atas permukaan laut. Curah hujan yang dikehendaki sekitar 1500-2000 mm per tahun. Padi menghendaki tempat yang memiliki intensitas sinar matahari cukup atau di lingkungan yang terbuka. Suhu yang baik untuk pertumbuhannya adalah 23°C. Tanah yang baik untuk pertumbuhan padi adalah tanah sawah yang kandungan fraksi pasir, debu dan lempung dalam perbandingan tertentu serta memerlukan air dalam jumlah yang cukup. Padi dapat tumbuh dengan baik pada tanah yang ketebalan lapisan atasnya antara 18-22 cm dengan pH antara 4-7 [9].

III. METODOLOGI PENELITIAN

A. Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari Dinas Pertanian, yaitu data produksi sektor petanian khususnya tanaman padi tahun 2012.

B. Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini adalah produksi padi sebagai variabel respon dan variabel prediktornya meliputi luas panen, luas puso, jumlah penggunaan pupuk, jumlah curah hujan, dan jumlah hari hujan dengan unit pengamatan 29 kabupaten dan 7 kota di Jawa Timur.

C. Metode Analisis Data

Langkah-langkah dalam analisis data adalah sebagai berikut.

1. Mendeskripsikan karakteristik variabel penelitian
2. Membuat model regresi kuadrat terkecil (*Ordinary Least Square*) dengan langkah-langkah sebagai berikut.
 - a. Identifikasi pola hubungan variabel jumlah produksi padi dengan faktor-faktor yang mempengaruhinya.
 - b. Menyusun model regresi antara jumlah produksi padi dengan faktor-faktor yang mempengaruhinya.
 - c. Melakukan pemeriksaan asumsi residual $\epsilon_i \sim \text{IIDN}(0, \sigma^2)$.
 - d. Mendeteksi adanya kasus multikolinieritas pada variabel bebas.
 - e. Mendeteksi adanya data pencilan.
3. Membuat model regresi robust untuk mengatasi adanya data pencilan (*outlier*) dengan langkah-langkah sebagai berikut.
 - a. Menaksir koefisien parameter (**b**) menggunakan OLS, sehingga didapatkan $\hat{y}_{i,0}$ dan $\epsilon_{i,0} = y_i - \hat{y}_{i,0}$ $i=1, 2, \dots, n$.
 - b. Menghitung $\hat{\sigma}_0$ dari nilai residual dan pembobot awal $w_{i,0} = \frac{\psi(\epsilon_{i,0}^*)}{(\epsilon_{i,0}^*)}$. Nilai $\psi(\epsilon_{i,0}^*)$ dihitung sesuai dengan fungsi Huber, dan $\epsilon_{i,0}^* = \epsilon_{i,0} / \hat{\sigma}_0$.
 - c. Menyusun matrik pembobot berupa matrik diagonal yang berisi elemen $w_{1,0}, w_{2,0}, \dots, w_{n,0}$ yang dinamakan W_0 .

Tabel 1.

Statistik Deskriptif Variabel Penelitian

Variabel	Mean	Std. Deviasi	Minimum	Maksimum
Produksi padi	338117	255572	4878	968505
Luas panen	54780	41304	925	158568
Luas puso	1564	2439	0	9557
Penggunaan pupuk	69463	49662	1123	166584
Curah hujan	1375,6	391,2	266	2226
Hari hujan	83,28	21,64	16	124

- d. Menghitung penaksir koefisien regresi dengan persamaan $b_{\text{Robust ke-1}} = (X^T W_0 X)^{-1} X^T W_0 Y$ dan $\sum_{i=1}^n |y_i - \hat{y}_{i,1}|$ atau $\sum_{i=1}^n |\epsilon_{i,1}|$.
- e. Mengulangi langkah b) sampai dengan d) hingga didapatkan $\sum_{i=1}^n |\epsilon_{i,m}|$ yang konvergen.
- f. Melakukan pemeriksaan asumsi residual $\epsilon_i \sim \text{IIDN}(0, \sigma^2)$.
4. Membuat model regresi *Generalized Least Square* untuk mengatasi penyimpangan terhadap asumsi residual independen. Karena nilai koefisien autokorelasi (ρ) tidak diketahui, maka dilakukan pendugaan atau estimasi nilai ρ secara iterasi dengan langkah-langkah sebagai berikut.
 - a. Meregresikan **Y** terhadap **X** dengan metode OLS menggunakan nilai awal dari estimasi parameter regresi robust, sehingga didapatkan nilai residual yang pertama yang disimbolkan dengan $e_i(1)$.
 - b. Meregresikan residual $e_i(1)$ dengan $e_{t-1}(1)$, sehingga diperoleh nilai koefisien parameter yang selanjutnya dijadikan sebagai nilai estimasi ρ .
 - c. Mensubstitusikan nilai estimasi ρ pada variabel **X*** dan **Y*** sesuai persamaan (9).
 - d. Meregresikan **Y*** dengan **X*** menggunakan regresi OLS yang kemudian diperoleh nilai residual $e_i(2)$.
 - e. Mengulangi langkah b) sampai dengan d) hingga diperoleh nilai estimasi ρ yang konvergen.
 - f. Melakukan pemeriksaan asumsi residual $\epsilon_i \sim \text{IIDN}(0, \sigma^2)$.
 - g. Melakukan tranformasi kembali ke dalam bentuk model OLS.

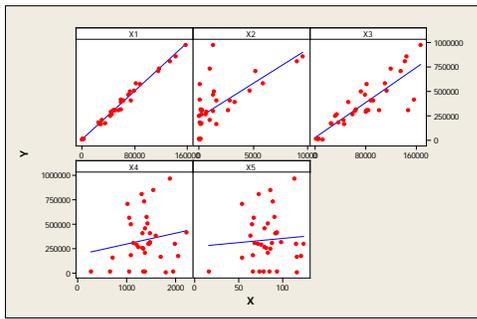
IV. ANALISIS PEMBAHASAN

A. Deskripsi Karakteristik Variabel Penelitian

Rata-rata produksi padi (Y) tiap kabupaten/kota di Jawa Timur adalah 338.117 ton dengan standard deviasi 255.572 ton. Kabupaten Jember merupakan kabupaten yang menghasilkan produksi padi terbesar di Jawa Timur, yaitu sebesar 968.505 ton dengan luas panen 158.568 ha. Rata-rata luas puso (X_2) yang terjadi di Jawa Timur adalah 1.564 ha dengan standard deviasi 2.439 ha. Kabupaten Lamongan merupakan kabupaten dengan luas puso terbesar, yaitu 9.557 ha. Hal ini terjadi karena pengaruh kondisi iklim yang tidak kondusif dan meningkatnya Organisme Pengganggu Tumbuhan (OPT). Statistik deskriptif variabel penelitian lebih jelas dapat dilihat pada Tabel 1.

B. Identifikasi Pola Hubungan antara Variabel Respon dengan Variabel Prediktor

Gambar 2 menunjukkan semua variabel prediktor memiliki pola hubungan yang positif terhadap variabel respon. Hasil pengujian korelasi dengan menggunakan $\alpha=5\%$, diketahui bahwa variabel luas panen, luas puso, dan penggunaan pupuk memiliki hubungan nyata terhadap produksi padi. Sedangkan,

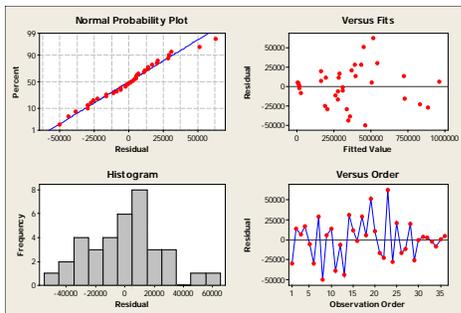


Gambar 2 Diagram Pencar Antara Variabel Respon Dengan Variabel Prediktor

Tabel 2.

Pengujian Parameter Regresi OLS secara Parsial

Variabel	t_{hitung}	p -value	VIF	Keterangan
β_1	23,28	0,000	5,223	Signifikan
β_2	0,59	0,562	1,834	Tidak signifikan
β_3	1,43	0,163	4,469	Tidak signifikan
β_4	0,03	0,973	2,742	Tidak signifikan
β_5	-0,61	0,548	2,545	Tidak signifikan



Gambar 3 Plot Residual

variabel curah hujan dan hari hujan tidak memiliki hubungan nyata terhadap produksi padi.

C. Pemodelan dengan Metode OLS

Berikut ini merupakan hasil model regresi OLS pada produksi padi di Jawa Timur tahun 2012.

$$Y = 13881 + 5,82X_1 + 1,47X_2 + 0,275X_3 + 0,64X_4 - 202X_5$$

Model tersebut memiliki nilai koefisien determinasi (R^2) sebesar 98,9% yang artinya model regresi dapat menjelaskan keragaman variabel respon sebesar 98,9%, sedangkan 1,1% dijelaskan oleh variabel atau faktor lain yang tidak ada dalam model.

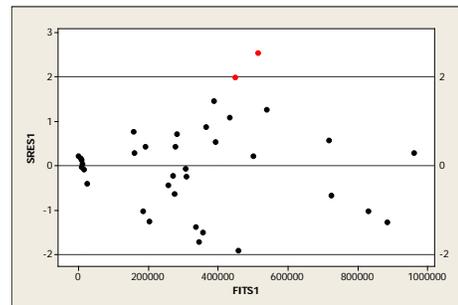
Pengujian parameter regresi secara serentak menggunakan ANOVA diperoleh nilai $F_{hitung} = 633,83$ dan p -value = 0,000. Dengan menggunakan $\alpha = 5\%$, dapat disimpulkan bahwa minimal terdapat satu parameter yang berpengaruh signifikan terhadap variabel produksi. Hasil pengujian parameter secara parsial menggunakan $\alpha = 10\%$ dapat dilihat pada Tabel 2. Nilai VIF masing-masing prediktor kurang dari 10, hal ini menunjukkan antar variabel prediktor tidak terjadi kasus multikolinieritas.

Langkah selanjutnya adalah pemeriksaan asumsi residual. Gambar 3 merupakan pemeriksaan asumsi residual secara visual.

Tabel 3 merupakan hasil uji hipotesis masing-masing asumsi residual. Pada Tabel 3, diketahui terdapat satu asumsi residual yang tidak terpenuhi yaitu terdapat autokorelasi antar residual. Sehingga, pemodelan produksi padi dengan metode

Tabel 3. Uji Hipotesis Asumsi Residual

Asumsi Residual	Statistik Uji	Keputusan
Independen	$d_{hitung} = 2,62096$ lebih besar dari $4 - d_U = 2,2013$	Tidak independen
Identik	Tidak ada parameter yang signifikan (p -value > 0,05)	Identik
Berdistribusi normal	KS (D) = 0,074 kurang dari $D_{(\alpha,n)} = 0,221$	Berdistribusi normal



Gambar 4 Deteksi Pencilan

OLS ini tidak dapat digunakan karena terdapat pelanggaran asumsi. Selain itu, terdapat titik yang jauh dari garis regresi yang diduga merupakan pencilan yang ditunjukkan oleh Gambar 3 bagian *Normal Probability Plot*. Sehingga, dilakukan identifikasi pencilan dengan menggunakan Gambar 4.

Gambar 4 merupakan plot *standardized residual versus fits*. Gambar 4 memperlihatkan terdapat dua titik yang melebihi batas nilai 2. Titik tersebut merupakan pengamatan ke-19 dan 23. Sehingga, kedua pengamatan tersebut diduga merupakan pencilan. Pada *output software Minitab*, juga menunjukkan terdapat residual yang mengandung huruf R yang berarti data memiliki nilai sisa atau residual yang besar (*R denotes an observation with large standardized residual*). Oleh karena itu, akan digunakan metode regresi robust untuk mengatasi adanya pencilan dan GLS untuk mengatasi adanya autokorelasi antar residual.

D. Pemodelan dengan Regresi Robust

Berikut ini merupakan model akhir regresi robust dengan estimasi M pada data produksi padi di Jawa Timur. Model ini menggunakan nilai awal (*initial value*) dari estimasi parameter metode OLS.

$$Y = 11659,5 + 5,836X_1 + 0,251X_2 + 0,28X_3 + 1,806X_4 - 186,047X_5$$

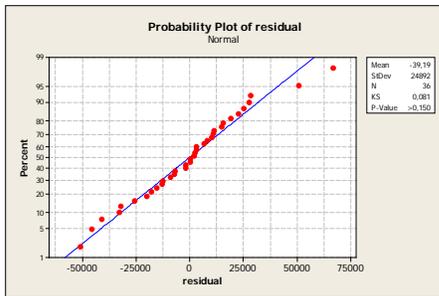
Nilai koefisien determinasi dari model tersebut adalah 98,89%, artinya model regresi dapat menjelaskan keragaman variabel respon sebesar 98,89%, sedangkan 1,11% dijelaskan oleh variabel atau faktor lain yang tidak ada dalam model.

Pengujian parameter regresi robust juga dilakukan secara serentak dan parsial. Pengujian parameter regresi robust secara serentak menggunakan F_{hitung} . Nilai F_{hitung} yang diperoleh adalah 621,679 lebih besar dari $F_{tabel} = 2,534$ pada taraf signifikansi 5%, sehingga dapat disimpulkan bahwa paling tidak terdapat satu parameter yang berpengaruh signifikan terhadap variabel produksi. Hasil pengujian parameter model regresi robust secara parsial menggunakan $t_{tabel} = 1,3014$ dapat dilihat pada Tabel 4.

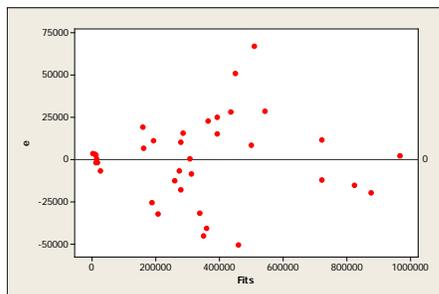
Selanjutnya, dilakukan pemeriksaan asumsi residual. Hasil pemeriksaan asumsi residual berdistribusi normal secara visual pada Gambar 5 menyatakan bahwa residual telah berdistribusi

Tabel 4.
Penguujian Parameter Regresi Robust

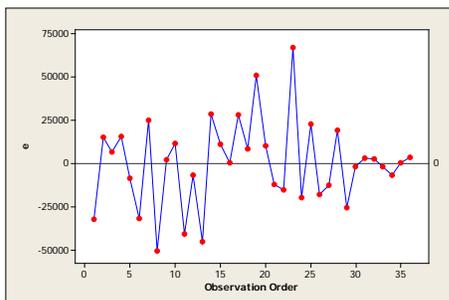
Variabel	t _{hitung}	t _(0,10;30)	Keterangan
β ₁	26,4888	1,3014	Signifikan
β ₂	0,115531	1,3014	Tidak signifikan
β ₃	1,60795	1,3014	Signifikan
β ₄	0,111640	1,3014	Tidak signifikan
β ₅	-0,664225	1,3014	Tidak signifikan



Gambar 5 Normal Probability Plot Regresi Robust



Gambar 6 Residual Versus Fits Plot Regresi Robust



Gambar 7 Residual Versus Order Plot Regresi Robust

busi normal karena titik-titik pengamatannya mengikuti pola garis linear. Pada pengujian dengan Kolmogorov-Smirnov diperoleh nilai $D=0,081$ dan p -value lebih besar dari 0,150. Jika nilai D dibandingkan dengan nilai $D_{(\alpha,n)}=0,221$ dengan menggunakan $\alpha=5\%$, maka diputuskan gagal tolak H_0 yang artinya residual telah berdistribusi normal.

Hasil pemeriksaan asumsi residual identik secara visual pada Gambar 6 menyatakan bahwa asumsi residual identik terpenuhi karena titik-titik pengamatannya menyebar secara random/acak dan tidak membentuk suatu pola. Pada uji glejser, nilai p -value yang dihasilkan lebih dari nilai $\alpha=5\%$ yang artinya tidak ada parameter yang signifikan berpengaruh terhadap model. Sehingga, dapat diputuskan bahwa asumsi residual identik terpenuhi.

Hasil pemeriksaan asumsi residual independen secara visual pada Gambar 7 menyatakan bahwa asumsi residual independen terpenuhi karena titik-titik pengamatannya menyebar secara random/acak. Namun, hasil uji Durbin-Watson diperoleh nilai $DW=2,466$. Nilai tersebut lebih besar dari batas

Tabel 5.
Penguujian Parameter Regresi GLS secara Parsial

Variabel	t _{hitung}	t _(0,10;30)	Keterangan
β ₁	29,075	1,3014	Signifikan
β ₂	1,853	1,3014	Signifikan
β ₃	1,861	1,3014	Signifikan
β ₄	-0,492	1,3014	Tidak signifikan
β ₅	-0,057	1,3014	Tidak signifikan

nilai $4-d_U=2,2013$. Sehingga, diputuskan bahwa asumsi residual independen tidak terpenuhi. Selanjutnya, dilakukan pemodelan dengan GLS untuk mengatasi pelanggaran asumsi residual independen.

E. Pemodelan dengan metode GLS

Berikut ini merupakan model akhir GLS dengan estimasi *Cochrane-Orcutt Iterative Procedure* pada produksi padi di Jawa Timur. Model ini menggunakan nilai awal dari estimasi parameter regresi robust.

$Y=14224,89+5,762X_1+3,479X_2+0,286X_3-8,687X_4-15,988X_5$
 Nilai koefisien determinasi model tersebut adalah 99,4%, artinya model regresi dapat menjelaskan keragaman variabel respon sebesar 99,4%, sedangkan 0,6% dijelaskan oleh variabel atau faktor lain yang tidak ada dalam model.

Hasil pengujian parameter regresi secara serentak diperoleh nilai $F_{hitung}=1202,193$ yang lebih besar dari $F_{tabel}=2,534$ pada $\alpha=5\%$, sehingga dapat disimpulkan paling tidak terdapat satu parameter yang berpengaruh signifikan terhadap variabel produksi. Untuk hasil pengujian parameter regresi secara parsial menggunakan $\alpha=10\%$ dapat dilihat pada Tabel 5.

Setelah itu, dilakukan pemeriksaan asumsi residual kembali. Hasil pemeriksaan asumsi residual berdistribusi normal secara visual menyatakan bahwa residual telah berdistribusi normal karena titik-titik pengamatannya mengikuti pola garis linear. Pada pengujian dengan Kolmogorov-Smirnov diperoleh nilai $D=0,071$ dan p -value lebih besar dari 0,150. Jika nilai D dibandingkan dengan nilai $D_{(\alpha,n)}=0,221$ dengan menggunakan $\alpha=5\%$, maka diputuskan gagal tolak H_0 yang artinya residual telah berdistribusi normal.

Hasil pemeriksaan asumsi residual identik secara visual menyatakan bahwa asumsi residual identik terpenuhi karena titik-titik pengamatannya menyebar secara random/acak dan tidak membentuk suatu pola. Pada uji glejser, nilai p -value yang dihasilkan lebih dari nilai $\alpha=5\%$ yang artinya tidak ada parameter yang signifikan berpengaruh terhadap model. Sehingga, dapat diputuskan bahwa asumsi residual identik terpenuhi.

Hasil pemeriksaan asumsi residual independen secara visual menyatakan bahwa asumsi residual independen terpenuhi karena titik-titik pengamatannya menyebar secara random atau acak. Pada uji Durbin-Watson diperoleh nilai $DW = 2,1333$. Nilai tersebut berada pada batas nilai $d_U < DW < 4-d_U$. Sehingga, diputuskan bahwa asumsi residual independen telah terpenuhi.

Setelah didapatkan model yang telah memenuhi semua asumsi residual, langkah selanjutnya melakukan transformasi kembali ke dalam bentuk OLS. Berikut ini merupakan model GLS yang telah dikembalikan ke dalam bentuk model OLS.

$$Y=13385+5,78X_1+3,08X_2+0,281X_3-6,7X_4-49X_5$$

Model tersebut menjelaskan bahwa apabila luas panen (X_1) mengalami kenaikan sebesar satu hektar, maka produksi padi

akan meningkat sebesar 5,78 ton dengan syarat variabel yang lainnya konstan. Nilai produksi padi juga akan menurun 49 ton apabila jumlah hari hujan (X_5) meningkat satu hari dengan syarat variabel yang lainnya konstan. Interpretasi serupa juga berlaku untuk variabel prediktor lain yang terdapat dalam model regresi. Nilai koefisien determinasi yang didapat adalah 99,3% yang berarti model regresi dapat menjelaskan keragaman variabel respon sebesar 99,3%, sedangkan 0,7% dijelaskan oleh variabel atau faktor lain yang tidak ada dalam model.

V. KESIMPULAN DAN SARAN

Pemodelan data produksi padi di Jawa Timur Tahun 2012 tidak dapat dilakukan dengan regresi OLS saja karena terdapat pelanggaran asumsi residual independen (autokorelasi) dan data pencilan (*outlier*), sehingga dilakukan regresi robust untuk mengatasi data pencilan dan GLS untuk mengatasi penyimpangan asumsi independen. Model akhir yang didapat menghasilkan variabel yang berpengaruh signifikan terhadap produksi padi di Jawa Timur Tahun 2012 adalah luas panen (X_1), luas puso (X_2) dan penggunaan pupuk (X_3) dengan nilai koefisien determinasi (R^2) sebesar 99,3%.

Dalam penelitian ini, variabel curah hujan dan banyak hari hujan tidak terlihat berpengaruh signifikan terhadap model karena diduga terdapat kasus multikolinieritas. Sehingga, disarankan untuk penelitian selanjutnya menggunakan metode analisis yang dapat mengatasi kasus multikolinieritas dan melakukan penambahan variabel prediktor lain yang diduga juga mempengaruhi produksi padi, seperti suhu, ketinggian tempat, kelembaban udara, jenis tanah, dll.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih pada jurusan Statistika ITS dan teman-teman yang telah membimbing penulis dalam pengerjaan Tugas Akhir sebagai syarat kelulusan. Penulis juga mengucapkan terima kasih kepada PT Angkasa Pura II yang telah memberikan dukungan finansial melalui Beasiswa Program BUMN Peduli Pendidikan tahun 2010-2013.

DAFTAR PUSTAKA

- [1] BPS. 2011. *Tanaman Pangan*. Dipetik Februrari 3, 2013 dari Badan Pusat Statistik: http://bps.go.id/tmn_pgn.php?kat=3
- [2] Lukmandono. 2002. *Menentukan Variabel-variabel yang Potensial Untuk Meningkatkan Produktivitas Sektor Pertanian Tanaman Padi Provinsi Jawa Timur*. Tesis Tidak Dipublikasikan, Institut Teknologi Sepuluh Nopember, Surabaya.
- [3] Norman, I. 2004. *Analisis Statistik Terhadap Faktor-faktor yang Mempengaruhi Produksi Padi di Wilayah Jawa Timur*. Tugas Akhir Tidak Dipublikasikan, Institut Teknologi Sepuluh Nopember, Surabaya.
- [4] Draper, N.R., & Smith, H. 1992. *Analisis Regresi Terapan* (Second ed.). Jakarta: PT Gramedia Pustaka Utama
- [5] Gujarati, N. D. 2004. *Basic Econometrics* (Fourth ed.). New York: The McGraw-Hill Companies.
- [6] Daniel, W. W. 1989. *Statistika Nonparametrik Terapan* (Second ed.). Jakarta: PT Gramedia.
- [7] Ryan, T. P. 1997. *Modern Regression Methods*. New York: John Wiley & Sons, Inc.
- [8] Fox, J. 2002. Robust Regression. *Appendix to An R and S-PLUS Companion to Applied Regression*.
- [9] Sastrapradja, S. D., & Widjaja, E. A. 2010. *Keanekaragaman Hayati Pertanian Menjamin Kedaulatan Pangan*. Jakarta: LIPI Press.