

Analisis Faktor-Faktor yang Mempengaruhi Produksi Padi di Provinsi Jawa Timur Menggunakan Regresi Semiparametrik *Spline*

Maulana Ishaq, Agnes Tuti Rumiati, dan Erma Oktania Permatasari

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: agnes_tuti@statistika.its.ac.id, erma@statistika.its.ac.id

Abstrak—Padi merupakan salah satu makanan pokok masyarakat Indonesia. Produksi padi pada tahun 2014 sebesar 70,85 ton Gabah Kering Giling (GKG) mengalami penurunan sebesar 433,24 ribu ton (0,61 persen) dibandingkan tahun 2013. Provinsi Jawa Timur merupakan provinsi penyumbang produksi padi terbesar di Indonesia. Hal ini yang menuntut Provinsi Jawa Timur untuk terus meningkatkan produksi padi demi ketercapaian swasembada pangan khususnya untuk masyarakat Jawa Timur dan umumnya untuk Indonesia. Penelitian ini akan membahas tentang faktor-faktor apa saja yang diduga mempengaruhi produksi padi Jawa Timur tahun 2014 menggunakan regresi semiparametrik *spline*. Hasil dari regresi semiparametrik *spline* menunjukkan factor-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 koefisien determinasi yang didapatkan adalah sebesar 99,17 persen.

Kata Kunci—Indonesia, Jawa Timur, Produksi Padi, Regresi Semiparametrik *Spline*

I. PENDAHULUAN

Oryza sativa L atau yang biasa dikenal dengan tanaman padi merupakan tanaman budidaya yang sangat penting bagi umat manusia. Tanaman padi menjadi sumber bahan pangan utama hampir dari setengah penduduk dunia. Tak terkecuali Indonesia, hampir seluruh penduduk Indonesia memenuhi kebutuhan bahan pangannya dari tanaman padi. Dengan demikian, tanaman padi merupakan tanaman yang mempunyai nilai spiritual, budaya, ekonomi, dan politik yang penting bagi bangsa Indonesia karena memengaruhi hajat hidup orang banyak [1].

Indonesia tercatat sebagai negara dengan konsumsi tanaman padi tertinggi di dunia. Untuk level Asia, Indonesia mengalahkan empat negara yang mengonsumsi tanaman padi tertinggi, seperti Korea, Jepang, Malaysia dan Thailand [2]. Keberadaan komoditi tersebut sebagai makanan pokok bagi hampir seluruh bangsa Indonesia harus tetap terjaga sepanjang tahun. Badan Pusat Statistik (BPS) memprediksikan angka produksi padi pada tahun 2015 sebanyak 74,99 juta ton gabah kering giling (GKG) atau mengalami kenaikan sebanyak 4,15 juta ton (5,85 persen) dibandingkan tahun 2014. Jika prediksi BPS benar, maka produksi padi Tahun 2015 merupakan yang tertinggi dalam 10 tahun terakhir. Menurut publikasi BPS tentang produksi tanaman pangan 2014 didapatkan informasi

bahwa produksi padi Indonesia pada tahun 2014 sebesar 70,85 juta ton gabah kering giling (GKG), mengalami penurunan sebesar 433,24 ribu ton (0,61 persen) dibandingkan tahun 2013. Penurunan produksi padi tersebut disebabkan penurunan produksi di Pulau Jawa sebesar 829,97 ribu ton. Sementara itu produksi padi di luar Pulau Jawa mengalami peningkatan sebesar 396,73 ribu ton. Penurunan produksi terjadi karena adanya penurunan luas panen dan produktivitas masing-masing 37,95 ribu hektar (0,27 persen) sebesar 0,17 kuintal/hektar (0,33 persen). Penurunan produksi padi tahun 2014 yang relatif besar terdapat di Provinsi Jawa Tengah, Jawa Barat, Aceh, Sumatera Utara, dan Nusa Tenggara Barat. Indonesia pernah menjadi salah satu penghasil produksi padi terbesar diantara negara-negara lain, namun pada lima tahun terakhir, seperti yang banyak diberitakan oleh banyak media massa kondisinya menjadi berlawanan. Pada waktu itu pasokan komoditi tersebut menjadi semakin terbatas. Diduga, faktor-faktor yang menjadi penyebab adalah semakin bertambahnya penduduk, adanya sikap berjaga-jaga dikalangan tertentu dan adanya dualisme pendapat ditubuh pemerintah, antara pihak yang menyatakan masih mencukupi dan pihak-pihak yang perlu menjaga stok pangan agar tercukupi pada periode tertentu [3].

Jawa Timur merupakan provinsi penyumbang terbesar produksi padi di Indonesia. Pada Tahun 2014 produksi padi di Jawa Timur mencapai 12.397.049 juta ton atau sebesar 17,5% dari total produksi padi di Indonesia. Produksi padi di Provinsi Jawa Timur pada tahun 2014 naik sebesar 347.707 ribu ton dari tahun 2013. Hal ini yang menuntut Provinsi Jawa Timur untuk terus meningkatkan produksi padi demi ketercapaian swasembada pangan khususnya untuk masyarakat Jawa Timur dan umumnya untuk Indonesia. Untuk mencegah produksi padi menurun ataupun menjaga agar produksi padi Jawa Timur tetap stabil bahkan meningkat adalah dengan cara mengetahui faktor-faktor apa saja yang mempengaruhi produksi padi Jawa Timur, sebagai upaya dalam mengurangi impor dan mencapai hasil akhir yaitu swasembada pangan. Penelitian ini akan membahas tentang faktor-faktor apa saja yang diduga mempengaruhi produksi padi Jawa Timur tahun 2014.

Menurut pakar terkait beberapa faktor yang dapat mempengaruhi naik turunnya rata-rata produksi padi per hektar adalah masalah kesuburan tanah, curah hujan, kelembapan, pemakaian pupuk, pemilihan bibit, cara bercocok tanam, jasad pengganggu dan sebagainya. Penelitian sebelumnya mengenai faktor-faktor yang mempengaruhi produksi padi adalah [4] variabel jumlah pupuk SP-36, variabel jumlah pupuk Urea, jumlah pupuk KCl, jumlah tenaga kerja, jumlah benih secara

individual berpengaruh nyata dan jumlah pestisida berpengaruh tidak nyata terhadap produksi padi. Variabel pestisida dan jam kerja, memberikan pengaruh nyata pada taraf kepercayaan lima persen terhadap produksi padi [5]. Variabel luas lahan garapan, variabel jumlah tenaga kerja efektif, variabel jumlah pupuk, variabel jumlah pestisida, variabel jarak lahan garapan dengan rumah petani, dan variabel sistem irigasi berpengaruh nyata terhadap peningkatan produksi padi. Sedangkan variabel pengalaman petani tidak berpengaruh nyata terhadap peningkatan produksi padi [6].

II. TINJAUAN PUSTAKA

A. Regresi Nonparametrik

Regresi nonparametrik merupakan salah satu model regresi yang sangat tepat digunakan untuk mendeteksi adanya sifat nonlinieritas dan adanya hubungan nonlinier dalam model [7]. Pada pendekatan regresi nonparametrik ini akan menggunakan asumsi yang longgar sehingga peneliti tidak memaksakan memakai bentuk kurva yang kaku namun membiarkan data sendiri yang akan mencari bentuk estimasi dari kurva regresinya. Kurva regresi dalam regresi nonparametrik diasumsikan halus dan termuat dalam suatu ruang fungsi tertentu. Beberapa model regresi nonparametrik yang banyak digunakan diantaranya *spline*, *kernel*, deret *fourier*, dan *Multivariate Analysis Regression Spline* (MARS). Adapun model regresi nonparametrik secara umum disajikan pada persamaan 1.

$$y_i = f(t_i) + \varepsilon_i ; i = 1, 2, \dots, n \quad (1)$$

dimana $f(t_i)$ merupakan fungsi *spline* berorde p dengan titik knot K_1, K_2, \dots, K_r dan bentuk $f(t_i)$ pada persamaan 2 secara lengkap disajikan sebagai berikut.

$$f(t_i) = \sum_{j=1}^p \gamma_j t_i^j + \sum_{k=1}^r \gamma_{p+k} (t_i - K_k)_+^p \quad (2)$$

dengan fungsi *truncated* $(t_i - K_k)_+^p$ dapat dijabarkan dalam persamaan 3.

$$(t_i - K_k)_+^p = \begin{cases} (t_i - K_k)^p, & t_i \geq K_k \\ 0, & t_i < K_k \end{cases} \quad (3)$$

Nilai K_k merupakan titik knot yang memperlihatkan pada perubahan dari fungsi sub interval yang berbeda.

B. Regresi Semiparametrik

Regresi semiparametrik adalah metode analisis statistik yang merupakan gabungan dari regresi parametrik dan regresi nonparametrik [8]. Jika dimisalkan dalam analisis regresi terdapat komponen parametrik dan komponen nonparametrik dengan data berpasangan (x_i, y_i, t_i) dan hubungan antara x_i, y_i, t_i yang telah diasumsikan mengikuti model regresi semiparametrik, maka disajikan dalam persamaan 4 berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_q x_{iq} + f(t_i) + \varepsilon_i \quad (4)$$

Dimana variabel prediktor $x_{ij}, j = 1, 2, \dots, q$ mempunyai hubungan secara parametrik dengan variabel respon y_i . Sedangkan variabel prediktor t_i mempunyai hubungan secara nonparametrik dengan variabel respon y_i .

C. Titik Knot Optimal

Metode yang baik untuk memilih titik knot optimal yaitu dengan menggunakan metode *Generalized Cross Validation* (GCV). Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku pola kurva pada internal yang berlainan [9]. Sehingga titik knot hanya dapat ditemukan pada data yang bentuk kurvanya tidak membentuk pola tertentu. Model spline terbaik dipilih dengan titik knot optimal dengan melihat nilai GCV terkecil. Adapun fungsi GCV diberikan dalam persamaan 5 berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{tr}[I - A(K_1, K_2, \dots, K_r)])^2} \quad (5)$$

Berdasarkan persamaan nilai (K_1, K_2, \dots, K_r) merupakan titik knot sedangkan matriks $A(K_1, K_2, \dots, K_r)$ diperoleh dari persamaan 6.

$$A(K_1, K_2, \dots, K_r) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (6)$$

Nilai dari $MSE(K_1, K_2, \dots, K_r)$ ditunjukkan pada persamaan 7 sebagai berikut.

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

D. Pengujian Parameter Model

Pengujian parameter dilakukan untuk mengetahui apakah suatu variabel memberikan pengaruh secara signifikan terhadap model atau tidak. Pengujian parameter dibagi dua yaitu pengujian secara serentak dan pengujian secara individu.

Pengujian serentak dilakukan untuk mengetahui apakah variabel prediktor secara serentak berpengaruh signifikan terhadap variabel respon atau tidak. Hipotesis yang digunakan pada uji F adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_q = \gamma_1 = \gamma_2 = \dots = \gamma_{p+r} = 0$$

$$H_1 : \text{minimal terdapat satu } \beta_j \neq 0 \text{ atau } \gamma_h \neq 0; j = 1, \dots, q; h = 1, \dots, (p + r)$$

Tabel 1.
Analisis Ragam (ANOVA) Uji Parameter

Sumber Variasi	Derajat Bebas	Sum of Square (SS)	Mean Square (MS)	F_{hitung}
Regresi	$q + p + r$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	
Error	$n - (q + p + r) - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{error}}{df_{error}}$	$\frac{MS_{regresi}}{MS_{error}}$
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	-	

Penarikan kesimpulan:

Jika $|F_{hitung}| \geq F_{tabel}$ atau $P_{value} < \alpha$ maka tolak H_0 yang berarti bahwa minimal terdapat satu variabel prediktor yang berpengaruh terhadap variabel respon, begitu juga sebaliknya. Jika kesimpulan tolak H_0 , maka lanjut pengujian secara individu.

Pengujian secara individu dilakukan untuk melihat pengaruh masing-masing variabel prediktor terhadap variabel respon dengan uji statistik t . Hipotesis yang digunakan dalam pengujian individu untuk komponen parametrik adalah sebagai berikut.

$$H_0 : \beta_j = 0; j = 1, \dots, q$$

$$H_1 : \beta_j \neq 0; j = 1, \dots, q$$

Hipotesis yang digunakan dalam uji individu untuk komponen nonparametrik adalah sebagai berikut.

$$H_0 : \gamma_h = 0; h = 1, \dots, (p + r)$$

$$H_1 : \gamma_h \neq 0; h = 1, \dots, (p + r)$$

Penarikan kesimpulan :

Jika $|t_{hitung}| \geq t_{tabel}$ atau $P_{value} < \alpha$ maka tolak H_0 yang berarti bahwa variabel prediktor berpengaruh signifikan terhadap variabel respon, begitu juga sebaliknya

E. Pengujian Asumsi Residual

Pengujian asumsi residual dilakukan untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yakni identik, independen, dan berdistribusi normal atautkah belum.

a. Asumsi Identik

Asumsi identik (homoskedastisitas) berarti bahwa varians pada residual sama atau disebut juga identic [10]. Secara pengujian, untuk melihat kasus heteroskedastisitas dapat digunakan uji *Glejser* dengan hipotesis sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu pasang } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$$

Nilai statistik uji *Glejser* dapat dirumuskan sebagai berikut.

$$F_{hitung} = \frac{\sum_{i=1}^n (|\hat{\epsilon}_i| - |\hat{\epsilon}_i|^2) / ((q+p+r)-1)}{\sum_{i=1}^n (|\epsilon_i| - |\epsilon_i|^2) / (n-(q+p+r))} \quad (8)$$

$$F_{tabel} = F_{\alpha; (q+p+r, n-(q+p+r)-1)} \quad (9)$$

Penarikan kesimpulan :

Jika $|F_{hitung}| \geq F_{tabel}$ atau $P_{value} < \alpha$ maka tolak H_0 yang berarti bahwa terdapat indikasi adanya kasus heteroskedastisitas, begitu juga sebaliknya.

b. Asumsi Independen

Asumsi independen (otokorelasi) dalam konsep regresi linier berarti komponen residual berkorelasi berdasarkan urutan waktu (pada data berkala) atau urutan ruang (pada data timpang lintang), atau korelasi pada dirinya sendiri [10]. Secara visual, untuk mengetahui apakah tidak terjadi autokorelasi adalah dengan melihat *plot* dari *Autocorrelation Function* (ACF) dari residual.

c. Asumsi Berdistribusi Normal

Secara pengujian, untuk mengetahui apakah residual berdistribusi normal adalah dengan menggunakan uji *Kolmogorov-Smirnov* dengan hipotesis sebagai berikut.

$$H_0 : F_0(x) = F(x)$$

$$H_1 : F_0(x) \neq F(x)$$

Nilai statistik uji *Kolmogorv-Smirnov* dapat dirumuskan sebagai berikut.

$$D = \sup_x |S(x) - F_0(x)| \quad (10)$$

dengan,

$F_0(x)$ = fungsi peluang kumulatif atau fungsi distribusi yang dihipotesiskan

$S(x)$ = fungsi peluang kumulatif yang dihitung dari data sampel atau proporsi nilai-nilai pengamatan dalam sampel yang kurang dari atau sama dengan x .

Penarikan kesimpulan :

Jika $|D| \geq q_{(1-\alpha)}$ atau $P_{value} < \alpha$ maka tolak H_0 yang berarti bahwa residual tidak berdistribusi normal, begitu juga sebaliknya.

F. Faktor-Faktor yang Mempengaruhi Produksi Padi

Menurut Pedoman Pengumpulan Data Tanaman Pangan dan Holtikultura, BPS – Deptan 1993, beberapa faktor yang dapat mempengaruhi naik turunnya rata-rata produksi padi per hektar adalah masalah kesuburan tanah, pemakaian pupuk, bibit, cara bercocok tanam, jasad pengganggu dan sebagainya. Berdasarkan informasi dari pakar terkait aspek anomali seperti kelembapan dan curah hujan berpengaruh terhadap produksi padi. Selain itu penelitian sebelumnya mengenai produksi padi yaitu menunjukkan variabel jumlah pupuk Urea, jumlah pupuk KCl, jumlah tenaga kerja, jumlah benih, jumlah pestisida, jumlah jam kerja, luas lahan garapan, jarak lahan garapan dengan rumah petani, dan sistem irigasi secara individual berpengaruh nyata dan jumlah pestisida berpengaruh tidak nyata terhadap produksi padi.

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder yang didapatkan dari Badan Pusat Statistik Provinsi Jawa Timur dan Dinas Pertanian Provinsi Jawa Timur pada tahun 2014. Kabupaten/Kota di Jawa Timur Tahun 2014. Pada penelitian ini menggunakan unit penelitian sebanyak 38 Kabupaten/Kota di Provinsi Jawa Timur tahun 2014.

B. Variabel Penelitian

Variabel penelitian yang diteliti berdasarkan tinjauan pustaka mengenai faktor-faktor yang diduga mempengaruhinya dari penelitian sebelumnya, beberapa literatur, dan pakar terkait adalah sebagai berikut.

Tabel 2.
Variabel Penelitian

Variabel	Keterangan
y	Produksi Padi (Ton)
x_1	Luas Panen Padi (Ha)
x_2	Realisasi Pupuk Bersubsidi (Ton)
t_1	Luas Puso Padi (Ha)
t_2	Curah Hujan (Mm)
t_3	Ketinggian Rata-Rata dari Permukaan Laut (Liter)

Berikut ini adalah definisi operasional dari masing-masing variabel penelitian.

- Variabel y menyatakan banyaknya produksi padi di suatu wilayah atau daerah.
- Variabel x_1 menyatakan luas tanaman yang dipungut hasilnya paling sedikit 11 persen dari keadaan normal yang bertujuan untuk menghasilkan Gabah Kering Giling (GKG).
- Variabel x_2 menyatakan banyaknya realisasi pupuk urea yang diterima oleh masing-masing Kabupaten/Kota dari alokasi yang disubsidikan Provinsi Jawa Timur.
- Variabel t_1 menyatakan luas tanaman yang mengalami kerusakan yang diakibatkan oleh serangan OPT (Organisme Pengganggu Tumbuhan), DFI (Dampak Fenomena Iklim) dan/atau oleh sebab lainnya (gempa bumi, dll), sedemikian

rupa sehingga hasilnya kurang dari 11 persen dari keadaan normal.

- e. Variabel t_2 menyatakan rata-rata curah hujan selama satu tahun yang diambil dari stasiun pengamatan di masing-masing Kabupaten/Kota.
- f. Variabel t_3 menyatakan tinggi rata-rata Ibukota atau pusat pemerintahan Kabupaten/Kota dari permukaan laut.

Setelah dijelaskan lebih rinci mengenai definisi operasional dari variabel penelitian yang digunakan, selanjutnya dapat disusun struktur data sebagaimana tertera pada Tabel 3.

Tabel 3. Struktur Data Penelitian

Variabel Respon (y)	Variabel Prediktor				
	Komponen Parametrik (x)		Komponen Nonparametrik (t)		
	x_1	x_2	t_1	t_2	t_3
y_1	$x_{1,1}$	$x_{1,2}$	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$
y_2	$x_{2,1}$	$x_{2,2}$	$t_{2,1}$	$t_{2,2}$	$t_{2,3}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
y_{38}	$x_{38,1}$	$x_{38,2}$	$t_{38,1}$	$t_{38,2}$	$t_{38,3}$

C. Langkah Analisis

Langkah-langkah yang akan dilakukan pada penelitian penelitian ini adalah sebagai berikut.

1. Melakukan analisis statistika deskriptif untuk mengetahui gambaran umum dari data meliputi.
 - a. Mendeskripsikan data dengan menggunakan rata-rata, varians, nilai maksimum dan nilai minimum.
 - b. Melihat pola hubungan antara dua variabel dengan menggunakan *scatterplot*.
2. Menentukan komponen parametrik dan komponen nonparametrik.
3. Melakukan pemodelan produksi padi menggunakan regresi semiparametrik *spline* dengan satu, dua, tiga dan kombinasi knot.
4. Memilih model regresi semiparametrik *spline truncated* terbaik berdasarkan titik knot optimal dengan menggunakan karakteristik nilai GCV paling minimum.
5. Menguji signifikansi parameter dari model semiparametrik *spline truncated* terbaik.
6. Menguji asumsi residual dari model semiparametrik *spline truncated* terbaik.
7. Melakukan interpretasi dan membuat kesimpulan.

IV. HASIL DAN PEMBAHASAN

A. Gambaran Umum Produksi Padi dan Faktor-Faktor yang Diduga Mempengaruhi

Langkah pertama yang dilakukan pada penelitian ini adalah melihat gambaran umum data dengan menggunakan nilai rata-rata, nilai varian, nilai minimum, dan nilai maksimum. Berikut ini adalah hasil analisis yang ditunjukkan pada Tabel 4.

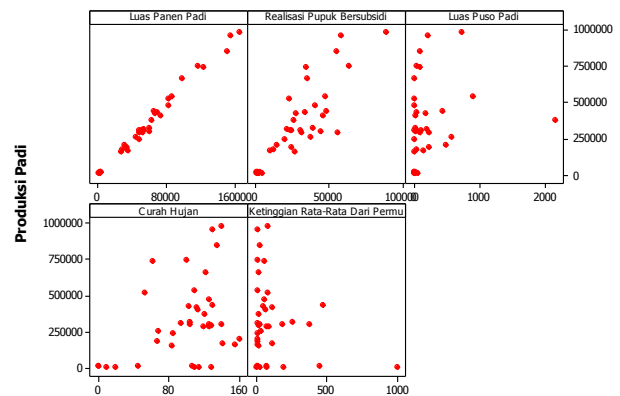
Tabel 4. Gambaran Umum Produksi Padi dan Faktor-Faktor yang Diduga Mempengaruhi

Variabel	Mean	Varians	Min	Maks
y	326238	73901105857	3653	978373
x_1	54543	1977474018	713	164307

x_2	27702	440277982	385	88449
t_1	180,7	152411,2	0	2145
t_2	99,61	1745,94	0	160,25
t_3	105,8	36038,9	2	996

B. Scatterplot Data Produksi Padi di Jawa Timur dengan Faktor-Faktor yang Diduga Mempengaruhinya

Langkah berikutnya adalah melihat *scatterplot* data antara variabel respon dengan masing-masing variabel predictor. Berikut ini merupakan *scatterplot* antara produksi padi dan faktor-faktor yang diduga mempengaruhi.



Gambar 1. Scatterplot Antara Produksi Padi dan Faktor-Faktor yang Diduga Mempengaruhi

Gambar 1 menunjukkan informasi bahwa secara visual luas panen padi (x_1) dan realisasi pupuk bersubsidi (x_2) membentuk pola linier sehingga variabel-variabel tersebut merupakan variabel parametrik (x). Sedangkan luas lahan puso padi (t_1), curah hujan (t_2), dan ketinggian rata-rata dari permukaan laut (t_3) cenderung acak, sehingga variabel-variabel tersebut merupakan variabel nonparametrik (t).

C. Pemilihan Titik Knot Optimum

Setelah mendapatkan nilai GCV dari pemodelan satu knot, dua knot, tiga knot, dan kombinasi knot selanjutnya adalah membandingkan nilai GCV untuk memilih knot mana yang terbaik. Berikut ini adalah nilai GCV terkecil dari masing-masing knot.

Tabel 5. Nilai GCV Masing-Masing Knot

GCV	Jumlah Knot	Jumlah Parameter
1062173893	1	9
64827506842	2	11
1018186986	3	15
1017563053	Kombinasi (2,3,1)	12

Nilai GCV dari masing-masing pemodelan tiap knot pada Tabel 5 menunjukkan bahwa nilai GCV paling minimum terdapat pada model *spline* dengan kombinasi knot (2,3,1). Kombinasi titik knot optimum adalah dua titik knot untuk variabel luas puso padi (t_1), tiga titik knot untuk variabel curah hujan (t_2), dan satu titik knot untuk variabel ketinggian rata-rata dari permukaan laut (t_3). Sehingga, dapat disimpulkan bahwa model regresi semiparametrik *spline* yang paling baik

adalah *spline* dengan kombinasi knot dengan jumlah parameter model sebanyak 12 sudah termasuk β_0 (konstanta).

D. Penaksiran Parameter Model Produksi Padi dengan Titik Knot Optimum

Berdasarkan nilai GCV paling minimum, model *spline* dengan kombinasi knot (2,3,1) merupakan kombinasi titik knot optimum dengan dua titik knot untuk variabel luas puso padi (t_1), tiga titik knot untuk variabel curah hujan (t_2), dan satu titik knot untuk variabel ketinggian rata-rata dari permukaan laut (t_3). Berikut ini merupakan model produksi padi dengan menggunakan titik knot optimal.

$$\hat{y} = 3362,85 + 6,49x_1 - 0,91x_2 + 38,88t_1 - 97,91(t_1 - 1707,24)_+ - 88,12(t_1 - 1751,02)_+ - 107,95t_2 - 13,67(t_2 - 130,61)_+ + 11971,46(t_2 - 133,88)_+ + 9593,14(t_2 - 156,73)_+ + 60,18t_3 - 2867,02(t_3 - 975,71)_+$$

E. Pengujian Parameter Regresi

Setelah mendapatkan model terbaik dengan GCV, selanjutnya adalah melakukan pengujian terhadap parameter model untuk mengetahui variabel prediktor apa saja yang mempengaruhi produksi padi di Provinsi Jawa Timur tahun 2014 secara signifikan.

Tabel 6
Analysis of Variance (ANOVA)

Sumber Variasi	db	Sum Square	Mean Square	F	P-value
Regresi	11	2,679551 x 10 ¹²	243595 538191	281,234	0,00
Error	26	22520277040	8661 64502		
Total	37	2,70207 x 10 ¹²			

Tabel 6 menunjukkan informasi bahwa nilai *p-value* sebesar 0,00 yang kurang dari nilai alpha sebesar 0,05, sehingga keputusannya tolak H_0 yang artinya minimal terdapat satu variabel prediktor yang memberikan pengaruh signifikan terhadap model. Oleh karena itu, pengujian dilanjutkan ke pengujian secara individu yang ditunjukkan pada Tabel.

Tabel 7.
Hasil Pengujian Parameter Secara Individu

Variabel	Parameter	Estimator	T	P-value
Konstan	β_0	3362,85	2,01974	0,05383
x_1	β_1	6,49	27,11508	0,0000
x_2	β_2	-0,91	-1,57069	0,12834
t_1	γ_1	38,88	1,33598	0,19313
	γ_2	-97,91	-1,13345	0,26737
	γ_3	-88,12	-1,13345	0,26737
t_2	γ_4	-107,95	-0,89777	0,37755
	γ_5	-13,67	-214775	0,04123
	γ_6	11971,46	1,99435	0,05669
t_3	γ_7	9593,14	2,06575	0,04895
	γ_8	60,18	1,32511	0,19666
	γ_9	-2867,02	-1,09666	0,28284

Tabel 7 menunjukkan informasi bahwa dari hasil pengujian individu dengan menggunakan nilai *p-value* dari ke-12 parameter tersebut terdapat parameter yang tidak signifikan karena nilai *p-value* > alpha (0,05) yakni $\beta_2, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_6, \gamma_8,$ dan γ_9 . Jika terdapat satu parameter yang signifikan dalam satu variabel prediktor maka variabel tersebut dikatakan signifikan walaupun parameter lain tidak signifikan. Variabel prediktor yang signifikan tersebut adalah variabel x_1 yaitu luas panen padi (Ha) dan Variabel t_2 yaitu curah hujan (Mm). Sedangkan variabel prediktor yang tidak signifikan adalah variabel x_2 yaitu realisasi pupuk bersubsidi (Ton), variabel t_1 yaitu luas puso padi (Ha), dan variabel t_3 yaitu ketinggian rata-rata dari permukaan laut (Meter).

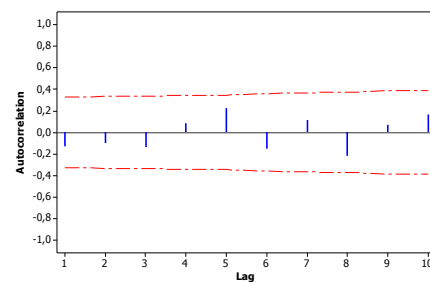
F. Pengujian Asumsi Residual

Asumsi residual yang harus dipenuhi bagi model terbaik dari regresi semiparametrik *spline* adalah asumsi residual identik, independen dan berdistribusi normal. Pengujian asumsi residual identik menggunakan statistik uji Glejser. Pemeriksaan asumsi residual independen menggunakan plot ACF. Pengujian asumsi residual distribusi normal menggunakan statistik uji *Kolmogorov Smirnov*.

Tabel 8.
Hasil Pengujian Glejser

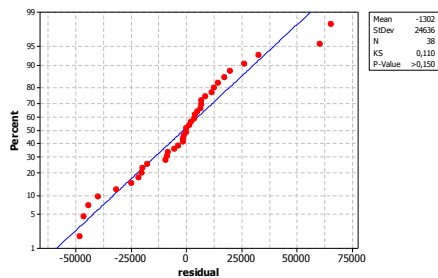
Sumber Variasi	db	Sum Square	Mean Square	F	P-value
Regresi	11	5089159002	462650818	1,718	0,124
Error	26	6999199137	269199967		
Total	37	12088358140			

Nilai *p-value* pengujian glejser untuk residual pada Tabel 8 adalah 0,12465 dimana, nilai tersebut lebih besar dari nilai alpha yaitu 0,05 sehingga keputusan gagal tolak H_0 yang artinya tidak terjadi heteroskedastisitas dalam model sehingga, asumsi residual identik telah dipenuhi.



Gambar 2. Plot ACF Residual

Gambar menunjukkan informasi bahwa residual tidak terdapat nilai autokorelasi (ACF) yang signifikan atau keluar dari batas atas dan batas bawah (*Confidence Interval*). Sehingga, dapat disimpulkan bahwa asumsi residual independen terpenuhi dan tidak terjadi aurokorelasi antar residual.



Gambar 3. Plot Normalitas Residual

Gambar 3 menunjukkan informasi bahwa didapatkan nilai *p-value* hasil pengujian yaitu > 0,150 dimana nilai tersebut lebih besar bila dibandingkan dengan nilai alpha sebesar 0,05. Karena nilai *p-value* > α maka dapat diputuskan pengujian normalitas dengan statistik uji *Kolmogorov-Smirnov* gagal tolak H_0 artinya residual berdistribusi normal. Sehingga, asumsi residual berdistribusi normal telah terpenuhi.

G. Interpretasi Model

Setelah melakukan pengujian signifikansi parameter dan semua pengujian asumsi residual telah terpenuhi maka model regresi terbaik dapat diinterpretasikan. Berdasarkan nilai GCV paling minimum, model *spline* dengan kombinasi knot (2,3,1) Kombinasi titik knot optimum adalah dua titik knot untuk variabel luas puso padi (t_1), tiga titik knot untuk variabel curah hujan (t_2), dan satu titik knot untuk variabel ketinggian rata-rata dari permukaan laut (t_3). Nilai koefisien determinasi yang dihasilkan oleh model semiparametrik *spline truncated* tersebut adalah 99,17 persen. Nilai tersebut menunjukkan bahwa variabel produksi padi di Provinsi Jawa Timur tahun 2014 dapat dijelaskan oleh kelima variabel prediktor sebesar 99,17 persen sedangkan sisanya sebesar 0,83 persen dijelaskan oleh variabel-variabel lain yang tidak terdapat didalam model. Interpretasi model dilakukan terhadap variabel yang berpengaruh signifikan terhadap model yaitu sebagai berikut.

1. Hubungan antara luas panen (x_1) terhadap produksi padi (y) dengan asumsi variabel lain dianggap konstan adalah sebagai berikut.

$$\hat{y} = 3362,85 + 6,49x_1$$

Jika terjadi kenaikan luas panen sebanyak satu Ha, maka produksi padi di Provinsi Jawa Timur akan naik sebanyak 6,49 Ton.

2. Hubungan antara ketinggian rata-rata dari permukaan laut (t_3) terhadap produksi padi (y) dengan asumsi variabel lain dianggap konstan adalah sebagai berikut.

$$\hat{y} = \begin{cases} -107,95t_2 & ; t_2 < 130,61 \\ -121,62t_2 + 1785,439 & ; 130,61 \leq t_2 < 133,88 \\ 11849,84t_2 - 1600953,626 & ; 133,88 \leq t_2 < 156,73 \\ 21442,98t_2 - 3104486,458 & ; t_2 \geq 156,73 \end{cases}$$

Jika curah hujan kurang dari 130,61 dan apabila variabel ini naik satu Mm, maka produksi padi cenderung turun sebesar 107,95 Ton. Daerah yang termasuk dalam interval ini adalah semua Kota Pasuruan, Maidun, Mojokerto, Batu, Probolinggo, Malang, Blitar, Surabaya, Kediri dan Kabupaten Madiun, Ngawi, Sumenep, Situbondo, Pamekasan, Sampang, Bangkalan, Banyuwangi, Jombang,

Bondowoso, Mojokerto, Tuban, Ponorogo, Lumajang, Kediri, Gresik, Pasuruan, Nganjuk, Magetan, Tulungagung, Probolinggo, Lamongan, Malang. Jika curah hujan antara 130,61 sampai 133,88 dan apabila variabel ini naik satu Mm, maka produksi padi cenderung turun sebesar 121,62 Ton. Daerah yang termasuk dalam interval ini tidak ada. Jika curah hujan antara 133,88 sampai 156,73 dan apabila variabel ini naik satu Mm, maka produksi padi cenderung naik sebesar 11849,84 Ton. Daerah yang termasuk dalam interval ini adalah Kabupaten Bojonegoro, Blitar, Jember, Trenggalek, dan Pacitan. Jika curah hujan lebih dari sama dengan 156,73 dan apabila variabel ini naik satu Mm, maka produksi padi cenderung naik sebesar 21442,98 Ton. Daerah yang termasuk dalam interval ini adalah Kabupaten Sidoarjo.

V. PENUTUP

A. Kesimpulan

Setelah analisis dan pembahasan dilakukan, ada beberapa hal yang dapat disimpulkan mengenai produksi padi di Provinsi Jawa Timur tahun 2014 dan faktor-faktor yang diduga mempengaruhinya antara lain sebagai berikut.

1. Gambaran umum produksi padi dan faktor-faktor yang diduga mempengaruhinya adalah luas panen padi dan realisasi pupuk bersubsidi membentuk pola linier sehingga variabel-variabel tersebut merupakan variabel parametrik. Sedangkan luas lahan puso padi, curah hujan, dan ketinggian rata-rata dari permukaan laut cenderung acak, sehingga variabel-variabel tersebut merupakan variabel nonparametrik.
2. Berdasarkan analisis pemodelan dengan metode regresi semiparametrik *spline* terbaik yang dihasilkan adalah *spline* dengan kombinasi knot (2,3,1). Sehingga model yang terbentuk adalah sebagai berikut.

$$\hat{y} = 3362,85 + 6,49x_1 - 0,91x_2 + 38,88t_1 - 97,91(t_1 - 1707,24)_+^1 - 88,12(t_1 - 1751,02)_+^1 - 107,95t_2 - 13,67(t_2 - 130,61)_+^1 + 11971,46(t_2 - 133,88)_+^1 + 9593,14(t_2 - 156,73)_+^1 + 60,18t_3 - 2867,02(t_3 - 975,71)_+^1$$

Dapat diketahui bahwa variabel prediktor yang berpengaruh signifikan adalah variabel luas panen padi (Ha) dan variabel yaitu curah hujan (Mm). Pada pemodelan variabel yang tidak berpengaruh signifikan tetap dimasukkan, karena produksi padi tetap dipengaruhi oleh variabel tersebut walaupun hanya sedikit. Nilai koefisien determinasi yang didapatkan adalah sebesar 99,17 persen yang artinya bahwa variabel produksi padi di Provinsi Jawa Timur tahun 2014 dapat dijelaskan oleh kelima variabel prediktor sebesar 99,17 persen sedangkan sisanya sebesar 0,83 persen dijelaskan oleh variabel-variabel lain yang tidak terdapat didalam model.

B. Saran

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan mengenai produksi padi di Provinsi Jawa Timur tahun 2014 didapatkan beberapa saran sebagai berikut.

1. Penelitian selanjutnya diharapkan dapat mengembangkan model regresi semiparametrik *spline* dengan fungsi yang

lain seperti kuadratik maupun kubik atau estimasi model yang digunakan menggunakan metode fungsi kernel multivariabel dan sebagainya, sehingga bentuk fungsi yang didapatkan lebih sesuai dengan kasus data yang akan digunakan.

2. Model yang terbentuk diharapkan dapat memberikan kontribusi bagi Provinsi Jawa Timur sebagai salah satu pertimbangan dalam penentuan kebijakan mengenai produksi padi.

DAFTAR PUSTAKA

- [1] Utama, M. Z. (2015). *Budidaya Padi pada Lahan Marjinal Kiat Meningkatkan Produksi Padi*. Yogyakarta: CV. Andi Offset.
- [2] Sari, N. I. (2013). KLN Kapanlagi Network. Retrieved from merdeka.com: <http://www.merdeka.com>
- [3] Noorjenah. (2015). *Produksi Tanaman Pangan 2014*. Jakarta: Badan Pusat Statistik.
- [4] Makruf, E., Octavia, Y., & Putra, W. E. (2011). *Faktor-Faktor yang Mempengaruhi Produksi Padi Sawah di Kabupaten Seluma (Bengkulu)*.
- [5] Harianja, S. S. (2011). *Faktor-Faktor yang Mempengaruhi Produksi Padi Sawah di Desa Kebonagung, Kecamatan Inogiri, Kabupaten Bantul*. Yogyakarta: Skripsi Sarjana Universitas Atma Jaya Yogyakarta.
- [6] Mahananto, Sutrisno, S., & Ananda, C. F. (2009). *Faktor-Faktor yang Mempengaruhi Produksi Padi (studi kasus di Kecamatan Nogosari, Boyolali, Jawa Tengah)*. Wacana, 12(1), 179-191.
- [7] Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker, Inc.
- [8] Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric Regression*. New York: Cambridge University Press.
- [9] Budiantara, I. N. (2006). *Regresi Nonparametrik dalam Statistika*. Makalah Pembicara Utama pada Seminar Nasional Matematika. Jurusan Matematika. FMIPA. Makasar: Universitas Negeri Makasar (UNM).
- [10] Gujarati. (2004). *Basic Econometrics Fourth Edition*. The McGraw-Hill Companies.