



## Pemodelan Produksi Padi di Provinsi Jawa Timur dengan Regresi Nonparametrik B-Spline

Sri Sulistijowati Handajani\* , Hasih Pratiwi , Yuliana Susanti, Respatiwan , Muhammad Bayu Nirwana , Arik Mahmudah

Program Studi Statistika, Universitas Sebelas Maret, Surakarta, Indonesia

\* Corresponding Author. E-mail: [rr\\_ssh@staff.uns.ac.id](mailto:rr_ssh@staff.uns.ac.id)

### ARTICLE INFO

#### Article History:

Received: 14-Nov. 2023

Revised: 13-Jan. 2024

Accepted: 21-Mar. 2024

#### Keywords:

B-spline, GCV, produksi padi, regresi, titik knot

### ABSTRACT

Kebutuhan pangan merupakan kebutuhan primer masyarakat yang harus terpenuhi. Makanan pokok yang banyak dikonsumsi masyarakat Indonesia salah satunya beras. Beras yang berasal dari padi selalu diusahakan memenuhi untuk kebutuhan konsumsi masyarakat terutama di sekitarnya. Jawa Timur adalah salah satu provinsi penyumbang terbesar produksi padi di Indonesia. Oleh sebab itu perlunya melihat pengaruh faktor-faktor iklim di beberapa wilayah produksi padi terbesar di provinsi Jawa Timur yaitu kabupaten Tuban, Nganjuk dan Gresik terhadap besarnya produksi padi di wilayah tersebut. Tujuan penelitian ini adalah menganalisis faktor-faktor meliputi suhu, kelembaban, curah hujan dan luas panen padi terhadap jumlah produksi padi. Data diambil dari *website* BMKG dan BPS tahun 2020-2022 di Kabupaten Tuban, Nganjuk dan Gresik. Metode analisis yang digunakan dengan memodelkan regresi nonparametrik B-spline dengan beberapa kombinasi titik knot dari beberapa variabel prediktor yang menghasilkan GCV terkecil dari kemungkinan banyaknya titik knot yang dicobakan. Hasil pemodelan mendapatkan knot optimum pada variabel  $X_1$  (suhu) berorde 2 dengan tiga titik knot bernilai 23,45584; 24,32467; 26,93116. Knot optimum pada variabel  $X_2$  (kelembaban) berorde 2 dengan satu titik knot bernilai 83,3828. Knot optimum pada variabel  $X_3$  (curah hujan) berorde 2 dengan dua titik knot bernilai 5,177247 dan 15,51238. Knot optimum pada variabel  $X_4$  (luas panen padi) berorde 2 dengan satu titik knot bernilai 16939,25. Nilai GCV minimum yang diperoleh adalah 18462458. Hasil analisis menunjukkan semua variabel berpengaruh signifikan walaupun untuk variabel iklim terdapat beberapa segmen yang kurang signifikan, dengan nilai  $R_{adj}^2$  sebesar 0,987.

*The need for food is a primary requirement of society that must be fulfilled. One of the staple foods widely consumed by the Indonesian society is rice. Rice, which comes from paddy fields, is always cultivated to fulfill the consumption needs of the community, especially in the surrounding areas. East Java is one of the largest contributors to rice production in Indonesia. Therefore, it is necessary to examine the influence of climate factors in several rice-producing regions in East Java, namely Tuban, Nganjuk, and Gresik regencies, on the level of rice production in those areas. This research aims to analyze factors including temperature, humidity, rainfall, and rice cultivation area on rice production quantity. The data was collected from the BMKG (Meteorology, Climatology, and Geophysics Agency) and BPS (Central Statistics Agency) websites for the years 2020-2022 in Tuban, Nganjuk, and Gresik regencies. The analysis method used involves modeling non-parametric B-splines with various combinations of knot points from multiple predictor variables, resulting in the smallest Generalized Cross-Validation (GCV) among the possible knot points tested. The modeling results obtained the optimal knots for variable  $X_1$  (temperature) of order 2 with three knot points at values 23.45584, 24.32467, and 26.93116. The optimal knot for variable  $X_2$  (humidity) of order 2 was at one knot point with a value of 83.3828. The optimal knots for variable  $X_3$  (rainfall) of order 2 were two knot points with values of 5.177247 and 15.51238. The optimal knot for variable  $X_4$  (rice cultivation area) of order 2 was at one knot point with a value of 16,939.25. The minimum GCV value obtained was 18,462,458. The analysis results indicate that all variables have a significant influence, although for climate variables, there were some segments that were less significant, with an  $R_{adj}^2$  value of 0.987.*




This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license



**How to Cite:**

Handajani, S. S., Pratiwi, H., Susanti, Y., Respatiwan, R., Nirwana, M. B., Mahmudah, A. (2023). Pemodelan produksi padi di Provinsi Jawa Timur dengan regresi nonparametrik B-spline. *Pythagoras: Jurnal Matematika dan Pendidikan Matematika*, 18(2), 159-175. <https://doi.org/10.21831/pythagoras.v18i2.67475>

 <https://doi.org/10.21831/pythagoras.v18i2.67475>

## PENDAHULUAN

Produksi padi di Indonesia memiliki peran penting dalam sektor pertanian dan perekonomian negara. Padi merupakan salah satu tanaman pangan utama di Indonesia dan merupakan sumber karbohidrat yang penting bagi penduduk Indonesia. Indonesia memiliki kondisi geografis yang mendukung pertanian padi. Iklim tropis, curah hujan yang cukup, serta lahan pertanian yang luas menjadikan Indonesia sebagai salah satu produsen padi terbesar di dunia. Indonesia telah berhasil mencapai swasembada beras, yaitu produksi beras yang cukup untuk memenuhi kebutuhan dalam negeri. Namun, masih terdapat tantangan dalam meningkatkan produktivitas dan efisiensi produksi padi, seperti pengelolaan air yang lebih baik, pengendalian hama dan penyakit, serta peningkatan akses petani terhadap teknologi pertanian modern.

Pada tahun 2021, Jawa Timur merupakan salah satu produsen padi terbesar di Indonesia. Sementara pada tahun 2020, produksi padi di Jawa Timur mencapai sekitar 14,9 juta ton padi gabah kering, khususnya di wilayah kabupaten Tuban, Nganjuk dan Gresik yang merupakan penghasil padi tertinggi di Jawa Timur. Jumlah ini menyumbang sekitar 19% dari total produksi padi di seluruh Indonesia (BPS, 2020). Provinsi Jawa Timur memiliki luas panen pertanian yang signifikan, serta kondisi iklim dan curah hujan yang mendukung pertumbuhan tanaman padi. Produksi padi di Jawa Timur dipengaruhi oleh beberapa faktor, seperti teknologi pertanian yang digunakan, varietas padi yang ditanam, metode irigasi, dan kebijakan pemerintah dalam sektor pertanian. Pemerintah Jawa Timur juga telah berusaha meningkatkan produktivitas dan efisiensi produksi padi melalui program-program yang mendukung petani, seperti penyediaan benih unggul, pelatihan pertanian, dan pengembangan infrastruktur pertanian. Akan tetapi kondisi saat ini seperti El Nino yaitu fenomena kurang hujan telah terjadi di sebagian pulau Jawa yang selama ini menjadi penghasil utama beras nasional, sehingga faktor curah hujan perlu dipertimbangkan dalam melihat pengaruhnya pada besarnya produksi padi. Pemodelan produksi padi menjadi tujuan penelitian ini karena perlu dilakukan dengan melihat kondisi iklim seperti curah hujan, kelembaban dan temperatur serta luas panen untuk melihat seberapa besar produksi padi dapat dihasilkan.

Pemodelan dengan menggunakan regresi dapat dilakukan untuk memodelkan pola hubungan fungsional antara satu atau lebih variabel. Hal ini dapat dimulai dengan melihat pola data antara variabel produksi padi dengan masing-masing variabel yang diperkirakan memengaruhi untuk menentukan pemodelan yang lebih tepat. Pemodelan secara parametrik, semiparametrik dan nonparametrik dapat dilakukan berdasarkan pola kecenderungan data tersebut melalui *scatter plot*. Di samping pola kecenderungan data dalam bentuk kurva nampak terlihat, dalam regresi parametrik diharapkan juga mempunyai informasi pola data sebelumnya sehingga dapat diperoleh pemodelan yang baik (Antoniadis et al., 2001; Eubank, 1999; Kayri & Zirrhoglu, 2009; Wahba, 1990; Wu & Zhang, 2006). Pemodelan tentang faktor-faktor yang memengaruhi produksi beras di Kabupaten Lombok Timur juga telah dilakukan oleh Bidayani et al. (2019) dengan regresi semiparametrik *spline* dan menghasilkan koefisien determinasi yang cukup tinggi yang menunjukkan model cukup baik.

Pada perkembangan akhir-akhir ini terdapat gejala-gejala alam yang menunjukkan perubahan perilaku alam sehingga mengarah pada pola yang tidak seperti biasanya (seolah-olah abnormal). Pada beberapa tahun yang lalu para petani masih bisa memperkirakan kapan mulai dan berakhirnya musim kemarau dan penghujan sehingga dapat mempersiapkan diri kapan untuk memulai menanam padi dan kapan memanennya, tetapi sekarang seolah-olah sulit untuk diprediksi. Oleh karena itu, diperlukan suatu pendekatan metode yang dapat digunakan dan memberikan pemodelan dan prediksi yang lebih baik. Pendekatan regresi non parametrik dan semi parametrik dapat digunakan untuk mendapatkan model yang baik dengan *error* yang kecil. Beberapa estimator yang digunakan untuk mengestimasi fungsi regresi telah banyak digunakan diantaranya histogram, kernel (Kayri & Zirrhoglu, 2009), *spline* (Becher et al., 2009; Chamidah et al., 2019; Cox & O'Sullivan, 1996; Fatmawati et al., 2019; Huang & Liu, 2006; Khan & Shahna, 2019; Koenker, R. Ng. & Portnoy, 1994; Lestari et al., 2010; Oehlert, 1992; Shahna & Khan, 2019; Wahba, 1990), dan *penalized weighted least square* (PWLS) (Lestari et al., 2020; Lestari et al.,

2022), *local linear* (Chamidah et al., 2019; Chamidah & Lestari, 2019). Selanjutnya estimator kernel dan *spline* dalam model nonparametrik dan model multirespon telah didiskusikan oleh Osmani et al. (2019), Lestari et al. (2018; 2019). Estimator di atas kecuali *spline* membutuhkan penentuan *bandwidth*.

Pemodelan B-*spline* merupakan model yang cukup baik dalam menginterpolasi dan memberikan estimasi dalam interval tertentu. Hal ini terlihat dari hasil penelitian Handajani et al. (2023) yang memodelkan data *temperature* pada beberapa stasiun yaitu Semarang, Juanda, Serang, Sleman, Bandung dan Kemayoran dengan model B-*spline* yang menghasilkan nilai *mean square error* (MSE) cukup kecil dan nilai koefisien determinasi cukup besar dari beberapa model tersebut. Fungsi dengan basis B-*spline* berdasarkan titik-titik knot pada interval data yang dapat digunakan untuk memprediksi nilai variabel responnya pada titik yang belum diketahui. Kemampuan serta fleksibilitasnya dalam memodelkan hubungan antara variabel respon dan prediktornya menyebabkan model B-*spline* menghasilkan hasil yang lebih baik, termasuk data yang mengandung pencilan. Dengan kelebihan model ini dimana plot data yang dipunyai tidak beraturan sehingga pemodelan produksi padi di beberapa wilayah di provinsi Jawa Timur dilakukan dengan pendekatan model regresi B-*spline*.

## METODE

### Jenis dan sumber data penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Jawa Timur (BPS Jawa Timur, 2023) dan *website* Badan Meteorologi Klimatologi dan Geofisika (BMKG, 2023). Variabel yang digunakan adalah jumlah produksi padi (ton) terbesar di Provinsi Jawa Timur tahun 2020-2022 yaitu di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel respon (Y), suhu rata-rata ( $^{\circ}\text{C}$ ) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_1$ ), kelembaban rata-rata (persen) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_2$ ), curah hujan (mm) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_3$ ), dan luas panen padi (hektar) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_4$ ). Unit pengamatan berupa data per bulan dari variabel-variabel tersebut sebanyak 108 data.

### Langkah-langkah penelitian

Langkah-langkah yang digunakan dalam penelitian ini yaitu:

1. Mengumpulkan data variabel respons dan variabel prediktor.
2. Membuat *scatterplot* antara variabel respons dengan masing-masing variabel prediktor.
3. Mengkombinasikan orde dan titik knot.
4. Menghitung nilai *generalized cross validation* (GCV) untuk setiap kombinasi orde dan titik knot.
5. Menentukan titik knot optimal yang dilihat dari nilai GCV minimum pada masing-masing kombinasi orde dan titik knot.
6. Menentukan model B-*spline* berdasarkan titik knot dan orde yang telah ditentukan sebelumnya.
7. Melakukan pengujian serentak untuk mengetahui ada tidaknya pengaruh variabel prediktor terhadap variabel respons.
8. Menentukan signifikansi kebaikan model dengan nilai koefisien determinasi.
9. Membuat interpretasi, menarik kesimpulan dan saran untuk model regresi nonparametrik B-*spline* yang diperoleh.

### Regresi B-Spline

Lyche dan Morken (2008) mengatakan bahwa salah satu model regresi nonparametrik yang memiliki interpretasi statistik dan visual khusus yang baik adalah B-*spline*. B-*spline* merupakan model polinomial tersegmen pada suatu titik fokus yang disebut knot sehingga memberikan sifat fleksibilitas yang baik. Sifat tersegmen ini memberikan fleksibilitas yang lebih dari polinomial biasa sehingga memungkinkan untuk menyesuaikan diri secara lebih efektif.

Pendekatan *spline* mempunyai suatu basis fungsi yang biasa dipakai antara lain *truncated power basis* dan basis B-*spline*. *Spline* dengan *truncated power basis* mempunyai kelemahan saat orde tinggi, knot yang banyak dan knot yang terlalu dekat akan membentuk matriks dalam persamaan normal untuk menghitung parameter  $\beta$  yang hampir singular sehingga persamaan normal sulit diselesaikan (Budiantara, 2006). Untuk mengatasi permasalahan

pada matriks persamaan normal yang singular digunakan pendekatan spline menggunakan basis fungsi yaitu basis B-spline (Eubank, 1999). Bentuk umum regresi nonparametrik B-spline berorde  $m$  dengan  $k$  titik knot dapat dinyatakan pada persamaan 1.

$$y_i = \sum_{h=1}^p \left( \sum_{j=1}^{m+k} \beta_j B_{(j-m),m}(x_{hi}) \right) + \varepsilon_i \quad (1)$$

$$= \sum_{h=1}^p \sum_{j=1}^{m+k} \beta_j B_{(j-m),m}(x_{hi}) + \varepsilon_i$$

dengan

- $\beta_j$  : parameter model untuk  $j = 1, 2, \dots, (m + k)$
- $B_{(j-m),m}$  : basis B-spline ke- $(j - m)$  berorde  $m$
- $x_{hi}$  : variabel prediktor ke- $h$  pada unit pengamatan ke- $i$
- $\varepsilon_i$  : error ke- $i$ .

Untuk membangun fungsi B-spline yang berorde  $m$  dengan  $k$  titik knot  $\xi_1, \dots, \xi_k$  dimana  $a_0 < \xi_1 < \dots < \xi_k < a_1$ , terlebih dahulu didefinisikan knot tambahan sebanyak  $2m$ , yaitu  $\xi_{-(m-1)}, \dots, \xi_{-1}, \xi_0, \xi_{k+1}, \dots, \xi_{k+m}$ , dengan  $\xi_{-(m-1)} = \dots = \xi_0 = a_0$  dan  $\xi_{k+1} = \dots = \xi_{k+m} = a_1$ . Biasanya  $a_0$  diambil dari nilai minimum  $x$  dan  $a_1$  diambil dari nilai maksimum  $x$  (Budiantara, 2006).

Menurut Eberly (1999), basis fungsi B-spline pada orde  $m$  dengan titik-titik knot di  $\xi_l$  dimana  $l = -(m - 1), \dots, k$  didefinisikan secara rekursif (pengulangan) sebagai berikut:

$$B_{l,m}(x_{hi}) = \frac{x_{hi} - \xi_l}{\xi_{l+m-1} - \xi_l} B_{l,m-1}(x_{hi}) + \frac{\xi_{l+m} - x_{hi}}{\xi_{l+m} - \xi_{l+1}} B_{l+1,m-1}(x_{hi})$$

dengan

- $B_{l,m}(x_{hi})$  : basis B-spline ke- $l$ , berorde  $m$
- $\xi_l$  : titik knot ke- $l$

untuk  $l = -(m - 1), \dots, k$  dan  $B_{l,m}(x_{hi}) = \begin{cases} 1, & x_{hi} \in [\xi_l, \xi_{l+1}] \\ 0, & \text{untuk } x_{hi} \text{ yang lainnya.} \end{cases}$

Fungsi B-spline berorde  $m$  dan  $k$  knot  $\xi_1, \dots, \xi_k$  pada persamaan (1) dapat ditulis sebagai:

$$y = \beta_1 B_{(1-m),m}(x_{1i}) + \beta_2 B_{(2-m),m}(x_{1i}) + \dots + \beta_{(m+k)} B_{k,m}(x_{1i}) + \dots + \beta_1 B_{(1-m),m}(x_{pi}) + \beta_2 B_{(2-m),m}(x_{pi}) + \dots + \beta_{(m+k)} B_{k,m}(x_{pi}) + \varepsilon_i \quad (2)$$

Persamaan (2) jika dinyatakan dalam bentuk matriks dapat ditulis menjadi:

$$\mathbf{y} = \mathbf{B}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3)$$

dengan matriks  $\mathbf{y}$  vektor variabel respon berordo  $(n \times 1)$ ,  $\mathbf{B}$  matriks basis B-spline berordo  $n \times (m + k)p$ ,  $\boldsymbol{\beta}$  parameter model berordo  $(m + k)p \times 1$ , dan  $\boldsymbol{\varepsilon}$  vektor error berordo  $(n \times 1)$ , dengan:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

dan matriks basis B-spline

$$\mathbf{B} = \begin{bmatrix} B_{(1-m),m}(x_{11}) & \dots & B_{k,m}(x_{11}) & \dots & B_{(1-m),m}(x_{p1}) & \dots & B_{k,m}(x_{p1}) \\ B_{(1-m),m}(x_{12}) & \dots & B_{k,m}(x_{12}) & \dots & B_{(1-m),m}(x_{p2}) & \dots & B_{k,m}(x_{p2}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_{(1-m),m}(x_{1n}) & \dots & B_{k,m}(x_{1n}) & \dots & B_{(1-m),m}(x_{pn}) & \dots & B_{k,m}(x_{pn}) \end{bmatrix}$$

parameter model

$$\beta = \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{(m+k)1} \\ \vdots \\ \beta_{1p} \\ \beta_{2p} \\ \vdots \\ \beta_{(m+k)p} \end{bmatrix}$$

dengan error

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

### Basis fungsi B-Spline

Orde model dapat ditentukan berdasarkan pola yang terjadi pada data, sedangkan banyaknya knot dan letak knot ditentukan berdasarkan perubahan pola di daerah tertentu pada kurva. Salah satu jenis basis fungsi B-spline yang dikategorikan berdasarkan orde  $m$  dan digunakan dalam analisis ini adalah orde  $m = 2$  memberikan basis fungsi B-spline linear, dengan fungsi sebagai berikut:

$$B_{l,2}(x_{hi}) = \frac{x_{hi}-\xi_l}{\xi_{l+1}-\xi_l} B_{l,1}(x_{hi}) + \frac{\xi_{l+2}-x_{hi}}{\xi_{l+2}-\xi_{l+1}} B_{l+1,1}(x_{hi})$$

dengan  $l = -1, 2, \dots, k; i = 1, 2, \dots, n; h = 1, 2, \dots, p$ .

### Estimasi parameter dalam model B-Spline

Kurva regresi pada persamaan (1) jika didekati dengan fungsi B-spline berorde  $m$  dengan  $k$  titik knot maka persamaannya menjadi

$$y = \beta_1 B_{(1-m),m}(x_{1i}) + \beta_2 B_{(2-m),m}(x_{1i}) + \dots + \beta_{(m+k)} B_{k,m}(x_{1i}) + \dots + \beta_1 B_{(1-m),m}(x_{pi}) + \beta_2 B_{(2-m),m}(x_{pi}) + \dots + \beta_{(m+k)} B_{k,m}(x_{pi}) + \varepsilon_i$$

Model regresi B-spline tersebut apabila dibuat dalam bentuk matriks maka menjadi:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} B_{(1-m),m}(x_{11}) & \dots & B_{k,m}(x_{11}) & \dots & B_{(1-m),m}(x_{p1}) & \dots & B_{k,m}(x_{p1}) \\ B_{(1-m),m}(x_{12}) & \dots & B_{k,m}(x_{12}) & \dots & B_{(1-m),m}(x_{p2}) & \dots & B_{k,m}(x_{p2}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ B_{(1-m),m}(x_{1n}) & \dots & B_{k,m}(x_{1n}) & \dots & B_{(1-m),m}(x_{pn}) & \dots & B_{k,m}(x_{pn}) \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{(m+k)1} \\ \vdots \\ \beta_{1p} \\ \beta_{2p} \\ \vdots \\ \beta_{(m+k)p} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

dapat ditulis menjadi:

$$y = B\beta + \varepsilon$$

Dalam regresi nonparametrik B-spline untuk mencari nilai estimasi  $\beta$  dapat menggunakan metode kuadrat terkecil (Valentina, 2022).

$$\hat{\beta} = (B^T B)^{-1} B^T y$$

Estimasi model untuk regresi nonparametrik B-spline adalah:

$$\hat{y} = B\hat{\beta}$$

$$\hat{y} = B(B^T B)^{-1} B^T y.$$

### Pemilihan model B-Spline terbaik

Untuk memperoleh model B-spline terbaik diperlukan titik knot yang optimal. Metode yang digunakan untuk memilih titik knot yang optimal yaitu *Generalized Cross Validation* (GCV), metode GCV dapat ditulis dengan (Wulandari & Budiantara, 2014).

$$GCV = \frac{MSE(k)}{\left(\frac{1}{n} \text{trace}[I - S]\right)^2}$$

$$MSE(k) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}(x_i))^2$$

$$S = B(B^T B)^{-1} B^T$$

dengan

- $n$  : banyaknya pengamatan
- $I$  : matriks identitas  $n \times n$
- $MSE$  : *Mean Square Error* (Rata-rata Kesalahan Kuadrat)
- $x_i$  : variabel prediktor
- $y_i$  : variabel respon
- $n$  : banyaknya pengamatan.

Pemilihan model terbaik dilakukan dengan cara membandingkan nilai GCV dari masing-masing orde dan titik knot yang memiliki nilai GCV paling kecil.

### Koefisien determinasi *adjusted* ( $R_{adj}^2$ )

Koefisien determinasi (salah satu ukuran *goodness of fit*) merupakan suatu ukuran yang penting dalam analisis regresi. Koefisien ini dapat menunjukkan baik atau tidaknya model regresi yang terestimasi. Nilai dari  $R^2$  sendiri berkisar antara  $0 \leq R^2 \leq 1$ . Jika  $R^2$  bernilai 1 berarti model estimasi yang diperoleh menunjukkan ketepatan yang sempurna, sedangkan jika bernilai 0 menunjukkan bahwa tidak ada hubungan antara variabel respon dengan variabel prediktor. Koefisien determinasi yang telah disesuaikan ( $R_{adj}^2$ ) merupakan koefisien yang digunakan telah dikoreksi dengan memasukkan jumlah variabel dan ukuran sampel yang digunakan. Model regresi dapat disimpulkan lebih baik jika memiliki nilai koefisien determinasi yang telah disesuaikan ( $R_{adj}^2$ ) lebih tinggi. Perhitungan  $R_{adj}^2$  dengan menggunakan berikut ini.

$$R_{adj}^2 = 1 - \frac{MSE}{JKT/(p-1)}$$

dengan

- $R_{adj}^2$  : koefisien determinasi yang disesuaikan
- $MSE$  : *mean squared error* (rata-rata kuadrat error)
- $JKT$  : jumlah kuadrat total
- $p$  : jumlah variabel.

### Uji signifikansi parameter model

Uji serentak (Uji F) merupakan pengujian untuk mengukur secara bersama-sama/serentak pengaruh variabel prediktor (X) terhadap variabel respon (Y). Uji hipotesisnya yaitu:

Hipotesis :

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  (semua variabel prediktor (X) tidak berpengaruh secara signifikan terhadap variabel respon (Y)).

$H_1$ : paling tidak terdapat satu  $\beta_i \neq 0; i = 1, 2, 3, \dots, (m + k)$

(paling tidak terdapat satu variabel prediktor (X) berpengaruh secara signifikan terhadap variabel respon (Y)).



Tingkat signifikansi  $\alpha$ : nilai  $\alpha$  yang biasa digunakan yaitu 0,05 atau 0,1.

Daerah kritis :  $H_0$  ditolak jika nilai  $p - value < \alpha$  atau nilai  $F_{hitung} > F_{n-(m+k)-1}$ . Dengan nilai  $(m+k)$  merupakan banyak parameter dalam regresi nonparametrik B-spline kecuali  $\beta_0$ .

Statistik uji :

$$F_{hitung} = \frac{\sum_{i=1}^n (\hat{y}_i - \hat{y})^2}{m+k} \div \frac{\sum_{i=1}^n (y_i - \hat{y}(x_i))^2}{n - (m+k) - 1}$$

Kesimpulan: tolak  $H_0$  jika  $F_{hitung} \in$  daerah kritis.

Uji parsial (uji  $t$ ) merupakan pengujian untuk mengukur secara individu pengaruh variabel prediktor (X) terhadap variabel respon (Y). Uji hipotesisnya yaitu:

Hipotesis :

$$H_0: \beta_i = 0; i = 1, 2, 3, \dots, (m+k)$$

(parameter  $\beta_i$  tidak berpengaruh secara signifikan terhadap model regresi).

$$H_1: \beta_i \neq 0; i = 1, 2, 3, \dots, (m+k)$$

(paling tidak terdapat satu parameter  $\beta_i$  (variabel  $x_i$ ) yang berpengaruh secara signifikan terhadap model regresi).

Tingkat signifikansi  $\alpha$ : nilai  $\alpha$  yang biasa digunakan yaitu 0,05 atau 0,1.

Daerah kritis:  $H_0$  ditolak jika nilai  $p - value < \alpha$  atau nilai  $t_{hitung} > t_{(\frac{\alpha}{2}, n-h)}$ .

Statistik uji:

$$t_{hitung} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$$

dengan  $\hat{\beta}_i$  adalah hasil estimasi parameter ke- $i$  dan  $SE(\hat{\beta}_i)$  adalah simpangan baku dari  $\hat{\beta}_i$ .

Kesimpulan : tolak  $H_0$  jika  $t_{hitung} \in$  daerah kritis.

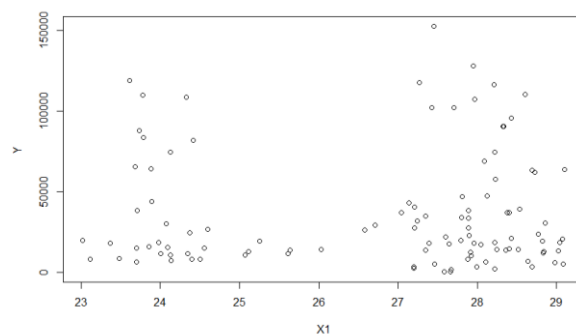
Uji normalitas dilakukan untuk validitas dari uji signifikansi parameter yang dilakukan dengan menguji apakah residu berdistribusi normal atau tidak. Pada penelitian ini, uji normalitas menggunakan uji Kolmogorov Smirnov. Pengujian parameter model regresi nonparametrik yang dihasilkan baik apabila residu berdistribusi normal. Data dikatakan berdistribusi normal jika  $p - value$  lebih besar dari taraf signifikansi ( $p - value > 0,05$ ). Begitu sebaliknya, jika  $p - value$  lebih kecil dari taraf signifikansi ( $p - value < 0,05$ ), maka data dikatakan tidak berdistribusi normal.

Uji autokorelasi merupakan pengujian untuk mengetahui apakah model regresi memiliki korelasi antar residu yang satu dengan residu yang lain. Untuk model regresi nonparametrik B-spline ini, lebih baik jika tidak terdapat autokorelasi. Pada penelitian ini, uji autokorelasi dilakukan dengan uji Durbin Watson. Data dikatakan tidak terdapat autokorelasi jika nilai  $p - value > \alpha$ .

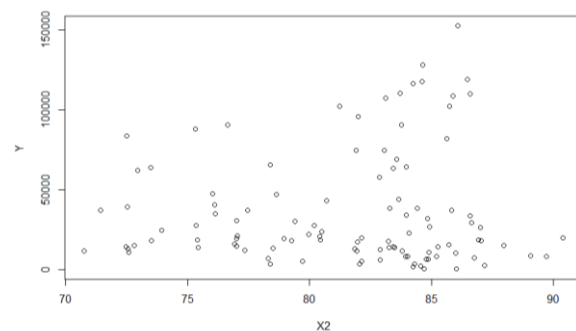
Uji heteroskedastisitas merupakan pengujian yang dilakukan untuk menguji apakah dalam model regresi yang diperoleh terdapat ketidaksamaan variansi pada residu dalam satu pengamatan dengan pengamatan lainnya. Pada penelitian ini, uji heteroskedastisitas dilakukan dengan uji Glejser. Residu dikatakan tidak terjadi heteroskedastisitas jika nilai  $p - value > \alpha$ .

## HASIL PENELITIAN

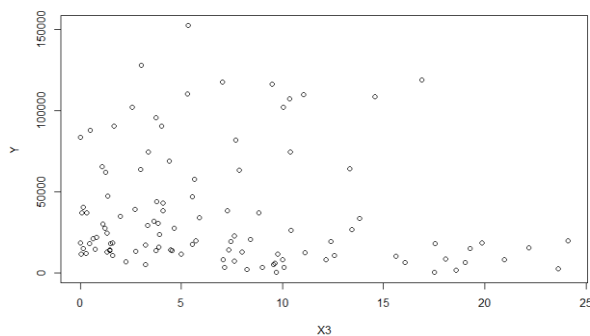
Bentuk pola hubungan antara variabel prediktor dan variabel respon dapat dilihat dari *scatterplot* yang terbentuk. *Scatterplot* antara variabel Y dengan masing-masing variabel  $X_1, X_2, X_3$  dan  $X_4$  dapat dilihat pada Gambar 1. Gambar 1 menunjukkan *scatterplot* antara suhu rata-rata ( $^{\circ}C$ ) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_1$ ), kelembaban rata-rata (persen) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_2$ ), curah hujan ( $mm$ ) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_3$ ), dan luas panen padi (hektar) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel prediktor ( $X_4$ ) terhadap jumlah produksi padi (ton) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik sebagai variabel respon (Y).



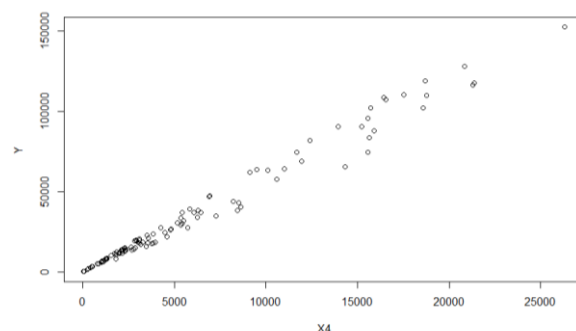
(a)



(b)



(c)



(d)

**Gambar 1.** Scatterplot antar variabel (a) Y dan  $X_1$ , (b) Y dan  $X_2$ , (c) Y dan  $X_3$ , (d) Y dan  $X_4$

Berdasarkan Gambar 1 dapat diketahui bahwa scatterplot antar variabel Y dengan masing-masing variabel  $X_1, X_2, X_3$  tidak mengikuti pola tertentu dan variabel Y dengan  $X_4$  menunjukkan hubungan yang mendekati linier. Karena ketiga variabel tersebut menunjukkan pola tidak teratur sehingga penelitian ini menggunakan pendekatan model



regresi nonparametrik *B-spline*. Pemilihan titik knot optimal dilakukan untuk mendapatkan estimasi model *B-spline* terbaik. Model *B-spline* yang optimal didapatkan dengan menentukan letak dan banyaknya knot dalam beberapa orde. Titik knot optimal yang nantinya akan dipilih berdasarkan nilai GCV terkecil. Dalam penelitian ini, orde yang digunakan yaitu orde 2, sedangkan jumlah titik knot yang digunakan dalam setiap kombinasi orde pada setiap variabel prediktor yaitu 1, 2, 3, dan 4 titik knot. Percobaan dalam menentukan nilai GCV menggunakan R diperoleh hasil pada Tabel 1, Tabel 2, Tabel 3, Tabel 4, dan Tabel 5.

Tabel 1. Satu titik knot optimum untuk masing-masing variabel

Indikator	Nilai
GCV minimum	21122365
Knot optimum variabel ke-1	26,93116
Knot optimum variabel ke-2	83,3828
Knot optimum variabel ke-3	15,51238
Knot optimum variabel ke-4	16939,25

Tabel 2. Dua titik knot optimum untuk masing-masing variabel

Indikator	Nilai
GCV minimum	19333631
Knot optimum variabel ke-1	24,32467; 26,93116
Knot optimum variabel ke-2	74,97074; 83,3828
Knot optimum variabel ke-3	5,177247; 15,51238
Knot optimum variabel ke-4	5677,751; 16939,25

Tabel 3. Tiga titik knot optimum untuk masing-masing variabel

Indikator	Nilai
GCV minimum	19727121
Knot optimum variabel ke-1	23,45584; 24,32467; 26,93116
Knot optimum variabel ke-2	72,16672; 74,97074; 83,3828
Knot optimum variabel ke-3	1,7322; 5,177247; 15,51238
Knot optimum variabel ke-4	1923,917; 5677,751; 16939,25

Tabel 4. Empat titik knot optimum untuk masing-masing variabel

Indikator	Nilai
GCV minimum	19370710
Knot optimum variabel ke-1	26,06233; 26,49674; 26,93116; 27,79998
Knot optimum variabel ke-2	80,57878; 81,98079; 83,3828; 86,18682
Knot optimum variabel ke-3	12,06734; 13,78986; 15,51238; 18,95743
Knot optimum variabel ke-4	13185,42; 15062,34; 16939,25; 20693,09

Tabel 5. Kombinasi knot optimum dan GCV minimum

Indikator	Nilai
GCV minimum	18462458
Knot optimum variabel ke-1	23,45584; 24,32467; 26,93116
Knot optimum variabel ke-2	83,3828
Knot optimum variabel ke-3	5,177247; 15,51238
Knot optimum variabel ke-4	16939,25

Berdasarkan penentuan titik knot dari beberapa kemungkinan yang telah dilakukan, Tabel 5 menunjukkan kombinasi knot optimum dan nilai GCV minimum yang diperoleh. Knot optimum pada variabel  $X_1$  berorde 2 dengan tiga titik knot bernilai 23,45584; 24,32467; 26,93116. Knot optimum pada variabel  $X_2$  berorde 2 dengan satu titik knot bernilai 83,3828. Knot optimum pada variabel  $X_3$  berorde 2 dengan dua titik knot bernilai 5,177247 dan 15,51238. Knot optimum pada variabel  $x_4$  berorde 2 dengan satu titik knot bernilai 16939,25. Nilai GCV minimum yang diperoleh adalah 18462458.

Setelah diperoleh nilai GCV terkecil dari model regresi nonparametrik B-spline untuk orde linier, selanjutnya dipilih model regresi nonparametrik B-spline berdasarkan titik knot optimal tersebut. Dari perhitungan diperoleh nilai estimasi parameter  $\beta$  untuk model B-spline terbaik dapat dilihat pada Tabel 6.

Tabel 6. Estimasi parameter model regresi

Variabel	Parameter	Estimasi Parameter
$X_1$	$\beta_{11}$	12955,815
	$\beta_{12}$	3223,183
	$\beta_{13}$	19541,814
	$\beta_{14}$	17404,213
	$\beta_{15}$	18415,406
$X_2$	$\beta_{21}$	23719,413
	$\beta_{22}$	27420,278
	$\beta_{23}$	17325,108
$X_3$	$\beta_{31}$	19301,392
	$\beta_{32}$	31213,067
	$\beta_{33}$	13906,287
	$\beta_{34}$	18362,826
$X_4$	$\beta_{41}$	21032,777
	$\beta_{42}$	25910,454
	$\beta_{43}$	16047,501

Berdasarkan Tabel 6 diperoleh persamaan estimasi model B-spline terbaik yang ditulis pada persamaan berikut:

$$\hat{y}_i = 12955,815B_{-1,2}(X_{1i}) + 3223,183B_{0,2}(X_{1i}) + 19541,814B_{1,2}(X_{1i}) + 17404,213B_{2,2}(X_{1i}) + 18415,406B_{3,2}(X_{1i}) + 23719,413B_{-1,2}(X_{2i}) + 27420,278B_{0,2}(X_{2i}) + 17325,108B_{1,2}(X_{2i}) + 19301,392B_{-1,2}(X_{3i}) + 31213,067B_{0,2}(X_{3i}) + 13906,287B_{1,2}(X_{3i}) + 18362,826B_{2,2}(X_{3i}) + 2103,777B_{-1,2}(X_{4i}) + 25910,454B_{0,2}(X_{4i}) + 16047,501B_{1,2}(X_{4i})$$

dengan,

$$B_{-1,2}(X_{1i}) = \begin{cases} \frac{23,46 - X_1}{0,44}, & 23,02 \leq X_1 < 23,46 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases}$$

$$B_{0,2}(X_{1i}) = \begin{cases} \frac{X_1 - 23,02}{0,44}, & 23,02 \leq X_1 < 23,46 \\ \frac{24,33 - X_1}{0,87}, & 23,46 \leq X_1 < 24,33 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases}$$

$$B_{1,2}(X_{1i}) = \begin{cases} \frac{X_1 - 23,46}{0,87}, & 23,46 \leq X_1 < 24,33 \\ \frac{26,93 - X_1}{2,6}, & 24,33 \leq X_1 < 26,93 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases}$$

$$\begin{aligned}
 B_{2,2}(X_{1i}) &= \begin{cases} \frac{X_1 - 24,33}{2,6}, & 24,33 \leq X_1 < 26,93 \\ \frac{29,10 - X_1}{2,17}, & 26,93 \leq X_1 < 29,10 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases} \\
 B_{3,2}(X_{1i}) &= \begin{cases} \frac{X_1 - 26,93}{2,17}, & 26,93 \leq X_1 < 29,10 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases} \\
 B_{-1,2}(X_{2i}) &= \begin{cases} \frac{83,38 - X_2}{12,61}, & 70,77 \leq X_2 < 83,38 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases} \\
 B_{0,2}(X_{2i}) &= \begin{cases} \frac{X_2 - 70,77}{12,61}, & 70,77 \leq X_2 < 83,38 \\ \frac{90,39 - X_2}{7,01}, & 83,38 \leq X_2 < 90,39 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases} \\
 B_{1,2}(X_{2i}) &= \begin{cases} \frac{X_2 - 83,38}{7,01}, & 83,38 \leq X_2 < 90,39 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases} \\
 B_{-1,2}(X_{3i}) &= \begin{cases} \frac{5,18 - X_3}{5,17}, & 0,0097 \leq X_3 < 5,18 \\ 0, & \text{untuk } X_3 \text{ lainnya} \end{cases} \\
 B_{0,2}(X_{3i}) &= \begin{cases} \frac{X_3 - 0,0097}{5,17}, & 0,0097 \leq X_3 < 5,18 \\ \frac{15,51 - X_3}{10,33}, & 5,18 \leq X_3 < 15,51 \\ 0, & \text{untuk } X_3 \text{ lainnya} \end{cases} \\
 B_{1,2}(X_{3i}) &= \begin{cases} \frac{X_3 - 5,18}{10,33}, & 5,18 \leq X_3 < 15,51 \\ \frac{15,51 - X_3}{8,62}, & 15,51 \leq X_3 < 24,13 \\ 0, & \text{untuk } X_3 \text{ lainnya} \end{cases} \\
 B_{2,2}(X_{3i}) &= \begin{cases} \frac{X_3 - 15,51}{8,62}, & 15,51 \leq X_3 < 24,13 \\ 0, & \text{untuk } X_3 \text{ lainnya.} \end{cases} \\
 B_{-1,2}(X_{4i}) &= \begin{cases} \frac{16939,25 - X_4}{16892,25}, & 47 \leq X_4 < 16939,25 \\ 0, & \text{untuk } X_4 \text{ lainnya} \end{cases} \\
 B_{0,2}(X_{4i}) &= \begin{cases} \frac{X_2 - 47}{16892,25}, & 47 \leq X_4 < 16939,25 \\ \frac{26323,8 - X_2}{9384,55}, & 16939,25 \leq X_2 < 26323,8 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases} \\
 B_{1,2}(X_{4i}) &= \begin{cases} \frac{X_2 - 16839,84}{9384,55}, & 16939,25 \leq X_2 < 26323,8 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases}
 \end{aligned}$$

Selanjutnya menguji secara serentak/bersama-sama pengaruh variabel prediktor terhadap variabel respons. Berikut langkah-langkah uji hipotesis pada uji serempak menggunakan uji F:

Hipotesis

$H_0: \beta_1 = \beta_2 = \dots = \beta_9 = 0$  (Semua parameter (variabel) model tidak berpengaruh signifikan).

$H_1$ : paling tidak terdapat satu  $\beta_i \neq 0$ ;  $i = 1, 2, \dots, 9$  (paling tidak terdapat satu parameter (variabel) yang berpengaruh signifikan pada model).

Tingkat signifikansi :  $\alpha = 0,05$

Daerah kritis :  $H_0$  ditolak jika nilai  $p - value < \alpha$

Statistik uji diberikan pada Tabel 7.

Tabel 7. Hasil uji serentak

SV	DB	JK	RK	F	p-value
Regresi	14	131524608867	9394614919	502,772	0,046
Residu	93	1737765402	18685649		
Total	107	133262374269			

Tabel 7 menunjukkan bahwa nilai  $F$  sebesar 502,772 dan  $p - value = 0,046 \leq \alpha = 0,05$ . Berdasarkan hasil tersebut, dapat diambil keputusan bahwa  $H_0$  ditolak yang artinya paling tidak terdapat satu variabel prediktor yang berpengaruh secara signifikan terhadap variabel respons.

Hasil pengujian signifikansi secara parsial dapat dilihat pada Tabel 8.

Tabel 8. Hasil Uji parsial

Variabel	Parameter	t	p-value	Keterangan
$X_1$	$\beta_{11}$	2,599	0,011	Signifikan
	$\beta_{12}$	0,713	0,478	Tidak Signifikan
	$\beta_{13}$	0,013	0,989	Tidak Signifikan
	$\beta_{14}$	0,337	0,737	Tidak Signifikan
	$\beta_{15}$	2,517	0,014	Signifikan
$X_2$	$\beta_{21}$	3,046	0,003	Signifikan
	$\beta_{22}$	2,300	0,024	Signifikan
	$\beta_{23}$	0,148	0,883	Tidak Signifikan
$X_3$	$\beta_{31}$	1,658	0,101	Tidak Signifikan
	$\beta_{32}$	2,495	0,014	Signifikan
	$\beta_{33}$	3,135	0,002	Signifikan
	$\beta_{34}$	0,266	0,791	Tidak Signifikan
$X_4$	$\beta_{41}$	17,997	4,708e-32	Signifikan
	$\beta_{42}$	60,182	2,842e-76	Signifikan
	$\beta_{43}$	32,990	4,007e-53	Signifikan

Apabila nilai  $p - value < \alpha = 0,05$  maka  $H_0$  ditolak yang berarti bahwa variabel prediktor berpengaruh secara signifikan terhadap variabel respon. Tabel 9 menunjukkan bahwa suhu rata-rata ( $^{\circ}\text{C}$ ) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik ( $X_1$ ), kelembaban rata-rata (persen) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik ( $X_2$ ), curah hujan ( $\text{mm}$ ) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik ( $X_3$ ), dan luas panen padi (hektar) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik ( $X_4$ ) berpengaruh secara signifikan terhadap jumlah produksi padi (ton) tahun 2020-2022 di Kabupaten Tuban, Nganjuk, dan Gresik ( $Y$ ).

### Uji Asumsi Model

Uji normalitas ini dilakukan untuk mengetahui apakah residu berdistribusi normal atau tidak. Pengujian asumsi distribusi normal menggunakan uji Kolmogorov Smirnov. Berdasarkan hasil pengujian yang telah dilakukan dengan  $\alpha = 0,05$  diperoleh nilai  $p - value = 0,837 > \alpha$  sehingga  $H_0$  gagal ditolak. Hal ini berarti bahwa residu berdistribusi normal. Uji autokorelasi dilakukan dengan menggunakan uji Durbin Watson untuk mengetahui apakah ada atau tidak autokorelasi pada residu. Berdasarkan hasil pengujian yang telah dilakukan dengan  $\alpha =$

0,05 diperoleh nilai  $p - value = 0,2 > \alpha$  sehingga  $H_0$  gagal ditolak. Hal ini berarti bahwa tidak terdapat autokorelasi pada residu. Uji heteroskedastisitas dilakukan dengan menggunakan uji *Gleyser* untuk mengetahui apakah model regresi terjadi ketidaksamaan varian error pada satu pengamatan dengan pengamatan yang lain. Berdasarkan hasil pengujian yang telah dilakukan dengan  $\alpha = 0,05$  diperoleh nilai  $p - value = 1 > \alpha$  sehingga  $H_0$  gagal ditolak. Hal ini berarti bahwa tidak terjadi heteroskedastisitas pada model regresi.

Untuk mengukur kebaikan model dan ketepatan dalam prediksi dilakukan validasi model. Nilai  $R_{adj}^2$  yang diperoleh pada penelitian ini adalah sebesar 0,987. Nilai  $R_{adj}^2$  tersebut menunjukkan bahwa pengaruh variabel suhu rata-rata, kelembaban rata-rata, curah hujan, dan luas panen padi berpengaruh secara signifikan terhadap jumlah produksi padi sebesar 98,7%, sedangkan 1,3% jumlah produksi padi di Provinsi Jawa Timur dipengaruhi oleh faktor lain yang tidak diteliti pada penelitian ini.

### PEMBAHASAN

Hasil analisis menghasilkan model regresi nonparametrik *B-spline* terbaik dari kombinasi titik-titik knot optimal dan diuraikan dengan mempertimbangkan nilai dari basis *B-spline* yang diperoleh. Interpretasi model dapat diuraikan sebagai berikut:

Apabila variabel  $X_2, X_3, X_4$  diasumsikan tetap, maka suhu rata-rata ( $X_1$ ) terhadap jumlah produksi padi di Jawa Timur ( $y_i$ ) dapat dimodelkan menjadi

$$\hat{y}_i = 12955,815B_{-1,2}(X_{1i}) \tag{4}$$

dengan

$$B_{-1,2}(X_{1i}) = \begin{cases} \frac{23,46 - X_1}{0,44}, & 23,02 \leq X_1 < 23,46 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases}$$

$$\hat{y}_i = 18415,406B_{3,2}(X_{1i}) \tag{5}$$

dengan

$$B_{3,2}(X_{1i}) = \begin{cases} \frac{X_1 - 26,93}{2,17}, & 26,93 \leq X_1 < 29,10 \\ 0, & \text{untuk } X_1 \text{ lainnya} \end{cases}$$

Ketika suhu rata-rata bernilai minimum yaitu 23,02°C, maka nilai minimum tersebut disubstitusikan ke persamaan (4) sehingga diperoleh jumlah produksi padi tersebut meningkat sebanyak 12956 ton. Selain itu, ketika suhu rata-rata bernilai maksimum yaitu 29,10°C, maka nilai maksimum tersebut disubstitusikan ke persamaan (5) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 18415 ton. Dengan demikian, ketika suhu rata-rata rendah maka jumlah produksi padi di wilayah Kabupaten Tuban, Nganjuk dan Gresik akan bertambah dan sebaliknya jika suhu rata-rata tinggi maka jumlah produksi padi di Jawa Timur juga akan lebih meningkat dibanding saat suhu 23,02°C.

Apabila variabel  $X_1, X_3, X_4$  diasumsikan tetap, maka kelembaban rata-rata ( $X_2$ ) terhadap jumlah produksi padi di Jawa Timur ( $Y_i$ ) dapat dimodelkan menjadi

$$\hat{y}_i = 23719,413B_{-1,2}(X_{2i}) \tag{6}$$

dengan

$$B_{-1,2}(X_{2i}) = \begin{cases} \frac{83,38 - X_2}{12,61}, & 70,77 \leq X_2 < 83,38 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases}$$

$$\hat{y}_i = 17325,108B_{1,2}(X_{2i}) \tag{7}$$

dengan

$$B_{1,2}(X_{2i}) = \begin{cases} \frac{X_2 - 83,38}{7,01}, & 83,38 \leq X_2 < 90,39 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases}$$

Ketika kelembaban rata-rata bernilai minimum yaitu 70,77, maka nilai minimum tersebut disubstitusikan ke persamaan (6) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 23719 ton. Selain itu, ketika kelembaban rata-rata bernilai maksimum yaitu 90,39, maka nilai maksimum tersebut disubstitusikan ke persamaan (7) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 17325 ton. Dengan demikian, ketika kelembaban rata-rata rendah maka jumlah produksi padi di Jawa Timur juga akan meningkat dan sebaliknya jika kelembaban rata-rata tinggi maka jumlah produksi padi di Jawa Timur juga akan meningkat tetapi tidak setinggi di saat rata-rata kelembaban lebih rendah.

Apabila variabel  $X_1, X_2, X_4$  diasumsikan tetap, maka curah hujan ( $X_3$ ) terhadap jumlah produksi padi di Jawa Timur ( $y_i$ ) dapat dimodelkan menjadi

$$\hat{y}_i = 19301,392B_{-1,2}(X_{3i}) \quad (8)$$

dengan

$$B_{-1,2}(X_{3i}) = \begin{cases} \frac{5,18-X_3}{5,17}, & 0,0097 \leq X_3 < 5,18 \\ 0, & \text{untuk } X_3 \text{ lainnya} \end{cases}$$

$$\hat{y}_i = 18362,826B_{2,2}(X_{3i}) \quad (9)$$

dengan

$$B_{2,2}(X_{3i}) = \begin{cases} \frac{X_3-15,51}{8,62}, & 15,51 \leq X_3 < 24,13 \\ 0, & \text{untuk } X_3 \text{ lainnya.} \end{cases}$$

Ketika curah hujan bernilai minimum yaitu 0,0097, maka nilai minimum tersebut disubstitusikan ke persamaan (8) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 19301 ton. Selain itu, ketika curah hujan bernilai maksimum yaitu 24,13, maka nilai maksimum tersebut disubstitusikan ke persamaan (9) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 18363 ton. Dengan demikian, ketika curah hujan rendah atau tinggi yang berada dalam rentang nilai minimum dan nilai maksimum maka jumlah produksi padi di Jawa Timur diperoleh meningkat.

Apabila variabel  $X_1, X_2, X_3$  diasumsikan tetap, maka luas panen padi ( $X_4$ ) terhadap jumlah produksi padi di Jawa Timur ( $y_i$ ) dapat dimodelkan menjadi

$$\hat{y}_i = 2103,777B_{-1,2}(X_{4i}) \quad (10)$$

dengan

$$B_{-1,2}(X_{4i}) = \begin{cases} \frac{16939,25-X_4}{16892,25}, & 47 \leq X_4 < 16939,25 \\ 0, & \text{untuk } X_4 \text{ lainnya} \end{cases}$$

$$\hat{y}_i = 16047,501B_{1,2}(X_{4i}) \quad (11)$$

dengan

$$B_{1,2}(X_{4i}) = \begin{cases} \frac{X_2-16839,84}{9384,55}, & 16939,25 \leq X_2 < 26323,8 \\ 0, & \text{untuk } X_2 \text{ lainnya} \end{cases}$$

Ketika luas panen padi bernilai minimum yaitu 47, maka nilai minimum tersebut disubstitusikan ke persamaan (10) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 2103 ton. Selain itu, ketika luas panen padi bernilai maksimum yaitu 26323,8, maka nilai maksimum tersebut disubstitusikan ke persamaan (11) sehingga diperoleh jumlah produksi padi tersebut bertambah sebanyak 16048 ton. Dengan demikian, ketika luas panen padi rendah maka jumlah produksi padi di Jawa Timur juga akan rendah dan sebaliknya jika luas panen padi tinggi maka jumlah produksi padi di Jawa Timur juga akan tinggi.

### SIMPULAN

Berdasarkan pembahasan di atas dapat disimpulkan bahwa model regresi B-spline untuk memodelkan produksi padi di Kabupaten Tuban, Nganjuk dan Gresik berdasarkan variabel suhu, curah hujan, kelembaban dan luas panen menghasilkan kontribusi  $R^2$  adjusted sebesar 0,987. Model terbaik diperoleh dari knot optimum pada variabel  $x_1$  berorde 2 dengan tiga titik knot bernilai 23,45584; 24,32467; 26,93116, knot optimum pada variabel  $x_2$  berorde 2 dengan satu titik knot bernilai 83,3828, knot optimum pada variabel  $x_3$  berorde 2 dengan dua titik



knot bernilai 5,177247 dan 15,51238, serta knot optimum pada variabel  $x_4$  berorde 2 dengan satu titik knot bernilai 16939,25. Nilai GCV minimum sebesar 18462458. Produksi padi dipengaruhi secara signifikan oleh suhu, curah hujan, kelembaban dan luas panen. Ketika suhu rata-rata bernilai minimum yaitu 23,02°C, maka jumlah produksi padi tersebut bertambah sebanyak 12956 ton. Sebaliknya, ketika suhu rata-rata bernilai maksimum yaitu 29,10°C, maka jumlah produksi padi akan lebih meningkat yaitu bertambah sebanyak 18415 ton. Ketika kelembaban rata-rata bernilai minimum yaitu 70,77, maka jumlah produksi padi tersebut bertambah sebanyak 23719 ton. Sebaliknya, ketika kelembaban rata-rata bernilai maksimum yaitu 90,39, maka jumlah produksi padi tersebut bertambah sebanyak 17325 ton. Ketika curah hujan diantara 0,0097-24,13, maka jumlah produksi padi tersebut bertambah sebanyak 18363-19301 ton. Faktor terakhir ketika luas panen padi rendah maka jumlah produksi padi di wilayah tersebut juga akan rendah dan sebaliknya jika luas panen padi tinggi maka jumlah produksi padi di Jawa Timur juga akan tinggi.

### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih atas dukungan finansial dari Hibah Penelitian Fundamental di Universitas Sebelas Maret Surakarta. Penulis juga mengucapkan terima kasih kepada pemberi referensi yang telah membantu dalam menyempurnakan makalah ini.

### DAFTAR PUSTAKA

- Antoniadis, A., Bigot, J., & Spatinas, T. (2001). Wavelet estimators in nonparametric regression: a comparative simulation study. *Journal of Statistical Software*, 6, 1–83. <https://doi.org/10.18637/jss.v006.i06>
- Badan Pusat Statistik (BPS). (2020). Badan Pusat Statistik Indonesia. <https://www.bps.go.id/id>.
- Badan Pusat Statistik Provinsi Jawa Timur (BPS Jawa Timur). (2023). Badan Pusat Statistik Provinsi Jawa Timur. [jatim.bps.go.id](http://jatim.bps.go.id).
- Data Online Pusat Database BMKG. (2023). Data Online Database BMKG Provinsi Jawa Timur. <https://dataonline.bmkg.go.id/home>
- Becher, H., Kauermann, G., Khomski, P., & Kouyate, B. (2009). Using penalized splines to model age and season of birth dependent effects of childhood mortality risk factors in rural burkina faso. *Biometrical Journal*, 51, 110–122. <https://doi.org/10.1002/bimj.200810496>
- Bidayani, B., Hadijati, M., & Fitriyani, N. (2019). Model regresi semiparametrik spline hasil produksi padi di Kabupaten Lombok Timur. *EIGEN MATHEMATICS JOURNAL*, 2(1), 11–17. <https://doi.org/10.29303/emj.v1i1.31>
- Chamidah, N., Gusti, K. H., Tjahjono, E., & Lestari, B. (2019). Improving of classification accuracy of cyst and tumor using local polynomial estimator. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 17(3), 1492–1500. <https://doi.org/10.12928/telkomnika.v17i3.12240>
- Chamidah, N., & Lestari, B. (2019). Estimation of covariance matrix using multi-response local polynomial estimator for designing children growth charts: A theoretically discussion. *Journal of Physics: Conference Series*, 1397(1). <https://doi.org/10.1088/1742-6596/1397/1/012072>
- Chamidah, N., Lestari, B., & Saifudin, T. (2019). Modeling of blood pressures based on stress score using least square spline estimator in bi-response nonparametric regression. *International Journal of Innovation, Creativity and Change*, 5(3). [https://www.ijcc.net/images/Vol\\_5\\_Iss\\_3/Part\\_2\\_2020/5321\\_Chamidah\\_2019\\_E\\_R.pdf](https://www.ijcc.net/images/Vol_5_Iss_3/Part_2_2020/5321_Chamidah_2019_E_R.pdf)
- Cox, D. D., & O'Sullivan, F. (1996). Penalized type estimator for generalized nonparametric regression. *Journal of Multivariate Analysis*, 56, 185–206. <https://doi.org/10.1006/jmva.1996.0010>

- Eberly, D. (1999). *B-Splines interpolation on Lattices*.  
<https://www.geometrictools.com/Documentation/BSplineInterpolation.pdf>
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. Marcel Dekker Inc.
- Fatmawati, F., Budiantara, I. N., & Lestari, B. (2019). Comparison of smoothing and truncated spline estimators in estimating blood pressures models. *International Journal of Innovation, Creativity and Change*, 5(3), 1177–1199. [https://www.ijicc.net/images/Vol\\_5\\_Iss\\_3/Part\\_2\\_2020/5320\\_Fatmawati\\_2019\\_E\\_R.pdf](https://www.ijicc.net/images/Vol_5_Iss_3/Part_2_2020/5320_Fatmawati_2019_E_R.pdf)
- Huang, J. Z., & Liu, L. (2006). Polynomial spline estimation and inference of proportional hazards regression models with flexible relative risk form. *Biometrics*, 62, 793–802. <https://doi.org/10.1111/j.1541-0420.2005.00519.x>
- Handajani, S. S., Pratiwi, H., Respatiwan, Qona'ah, N., Ramadhania, M., Evitasari, N., Apsari, N. E. (2023). Comparison of B-spline and truncated spline regression models for temperature forecast. *Journal of Mathematics and Its Applications*, 17(4), 1969-1984. <https://doi.org/10.30598/barekengvol17iss4pp1969-1984>
- Kayri, M., & Zirhoglu, G. (2009). Kernel smoothing function and choosing bandwidth for nonparametric regression methods. *Ozean Journal of Applied Sciences*, 2, 49–60. <https://www.semanticscholar.org/paper/KERNEL-SMOOTHING-FUNCTION-AND-CHOOSING-BANDWIDTH-Kayri-Z%20C4%B1rh%C4%B1o/b9aa75efe9bdaaa4131f2c209a865ad14021d326>
- Khan, A., & Shahna, S. (2019). Non-polynomial quadratic spline method for solving fourth order singularly perturbed boundary value problems. *Journal of King Saud University - Science*, 31(4), 479–484. <https://doi.org/10.1016/j.jksus.2017.08.006>
- Koenker, R. Ng., P., & Portnoy, S. (1994). Quantile smoothing splines. *Biometrika*, 81(4). <https://doi.org/10.2307/2337070>
- Lestari, B., Budiantara, I. N., Sunaryo, S., & Madhuri, M. (2010). Spline estimator in-multi response nonparametric regression model. *Journal of Mathematics and Statistics*, 6, 327–332. <https://jurnal.unej.ac.id/index.php/JID/article/view/102/74>
- Lestari, B., Chamidah, N., Aydin, D., & Yilmaz, E. (2022). Reproducing kernel hilbert space approach to multiresponse smoothing spline regression function. *Symmetry*, 14(11), 2227. <https://doi.org/10.3390/sym14112227>
- Lestari, B., Fatmawati, F., & Budiantara, I. N. (2020). Spline estimator and its asymptotic properties in multiresponse nonparametric regression model. *Songklanakarin Journal Of Science and Technology*, 42(3), 533–548. <https://doi.org/10.14456/sjst-psu.2020.68>
- Lestari, B., Fatmawati, F., Budiantara, I. N., & Chamidah, N. (2018). Estimation of regression function in multi-response nonparametric regression model using smoothing spline and kernel estimators. *Journal of Physics: Conference Series*, 1097, 012091. <https://doi.org/10.1088/1742-6596/1097/1/012091>
- Lestari, B., Fatmawati, F., Budiantara, I. N., & Chamidah, N. (2019). Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches. *Journal of Physics: Conference Series*, 1397(1), 012064. <https://doi.org/10.1088/1742-6596/1397/1/012064>

- Lyche, T., & Morken, K. (2008). *Spline Methods*. Retrieved September 6, 2022, from <https://www.uio.no/studier/emner/matnat/ifi/nedlagte-emner/INF-MAT5340/v10/undervisningsmateriale/book.pdf>
- Oehlert, G. W. (1992). Relaxed boundary smoothing spline. *The Annals of Statistics*, 20(1146–1160). <http://www.jstor.org/stable/2242154>
- Osmani, F., Hajizadeh, E., & Mansouri, P. (2019). Kernel and regression spline smoothing techniques to estimate coefficient in rates model and its application in psoriasis. *Medic. J. Islamic Rep. Iran*, 33(90), 1–5. <https://doi.org/10.34171/mjiri.33.90>
- Shahna, S., & Khan, A. (2019). Approximations for higher order boundary value problems using non-polynomial quadratic spline based on off-step points. *Journal of King Saud University - Science*, 31(4), 737–745. <https://doi.org/10.1016/j.jksus.2018.06.004>
- Valentina, M. F. (2022). Pemodelan kemiskinan di jawa timur menggunakan regresi nonparametrik B-spline. *Surabaya: Universitas Islam Negeri Sunan Ampel*. <http://digilib.uinsa.ac.id/id/eprint/52923>
- Wahba, G. (1990). Spline models for observation data. In *SIAM Pennsylvania*. <https://doi.org/10.1137/1.9781611970128>
- Wu, H., & Zhang, J. T. (2006). *Nonparametric Regression Method for Longitudinal Data Analysis: Mixed Effects Modeling Approaches*. John Wiley and Sons.
- Wulandari, I. D., & Budiantara, I. N. (2014). Analisis faktor-faktor yang mempengaruhi persentase penduduk miskin dan pengeluaran perkapita makanan di jawa timur menggunakan regresi nonparametrik biresponden spline. *Jurnal Sains Dan Seni POMITS*, 3(1). <https://doi.org/10.12962/j23373520.v3i1.6110>