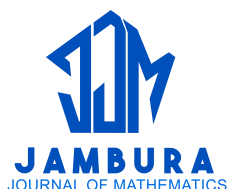


Estimasi Produksi Beras dengan Estimator Campuran Spline Truncated – Kernel di Jawa Timur

Andrea Tri Rian Dani, Fachrian Bimantoro Putra, I Nyoman Budiantara, dan Vita Ratnasari



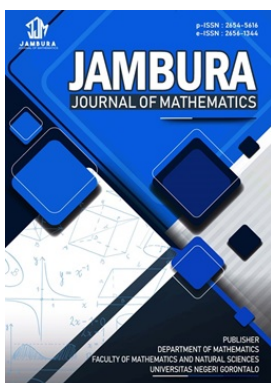
Volume 7, Issue 2, Pages 166–176, August 2025

Diterima 11 Juli 2025, Direvisi 13 Agustus 2025, Disetujui 17 Agustus 2025, Diterbitkan 19 Agustus 2025

To Cite this Article : A. T. R. Dani, F. B. Putra, I. N. Budiantara, dan V. Ratnasari ,“Estimasi Produksi Beras dengan Estimator Campuran Spline Truncated – Kernel di Jawa Timur”, *Jambura J. Math*, vol. 7, no. 2, pp. 166–176, 2025, <https://doi.org/10.37905/jjom.v7i2.33379>

© 2025 by author(s)

JOURNAL INFO • JAMBURA JOURNAL OF MATHEMATICS

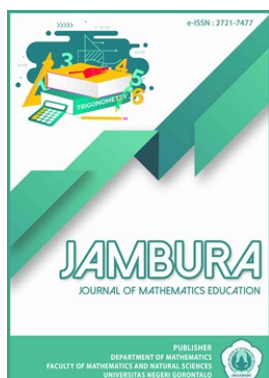


	Homepage	:	http://ejurnal.ung.ac.id/index.php/jjom/index
	Journal Abbreviation	:	Jambura J. Math.
	Frequency	:	Biannual (February and August)
	Publication Language	:	English (preferable), Indonesia
	DOI	:	https://doi.org/10.37905/jjom
	Online ISSN	:	2656-1344
	Editor-in-Chief	:	Hasan S. Panigoro
	Publisher	:	Department of Mathematics, Universitas Negeri Gorontalo
	Country	:	Indonesia
	OAI Address	:	http://ejurnal.ung.ac.id/index.php/jjom/oai
	Google Scholar ID	:	iWLjgaUAAAAJ
	Email	:	info.jjom@ung.ac.id

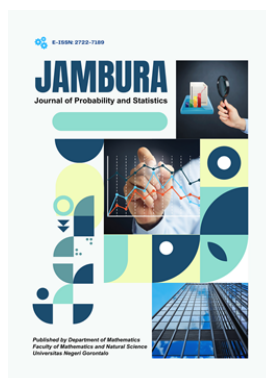
JAMBURA JOURNAL • FIND OUR OTHER JOURNALS



Jambura Journal of Biomathematics



Jambura Journal of Mathematics Education



Jambura Journal of Probability and Statistics



EULER : Jurnal Ilmiah Matematika, Sains, dan Teknologi

Estimasi Produksi Beras dengan Estimator Campuran Spline Truncated – Kernel di Jawa Timur

Andrea Tri Rian Dani^{1,2,*}, Fachrian Bimantoro Putra³, I Nyoman Budiantara⁴, Vita Ratnasari⁴

¹Program Studi Statistika, Universitas Mulawarman, Samarinda 75119, Indonesia

²Doktor Program MIPA, Universitas Airlangga, Surabaya 60286, Indonesia

³BNI Staff Banking, Future Relationship Manager, Bank Negara Indonesia, Indonesia

⁴Departemen Statistika, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

ARTICLE HISTORY

Diterima 11 Juli 2025

Direvisi 13 Agustus 2025

Disetujui 17 Agustus 2025

Diterbitkan 19 Agustus 2025

KATA KUNCI

Estimator Campuran
Spline Truncated-Kernel
GCV
Produksi Beras
Regresi Nonparametrik

KEYWORDS

Mixed Estimators
Truncated Spline-Kernel
GCV
Rice Production
Nonparametric Regression

ABSTRAK. Penelitian ini bertujuan untuk mengaplikasikan model regresi nonparametrik dengan menggunakan estimator campuran Spline Truncated dan Kernel untuk mengestimasi Produksi Beras di Provinsi Jawa Timur. Model ini mengkombinasikan beberapa variabel prediktor, yaitu Luas Panen Tanaman Padi, Produktivitas Tanaman Padi, Jumlah Penduduk, dan Indeks Pembangunan Manusia. Pemilihan kombinasi variabel terbaik dilakukan berdasarkan nilai Generalized Cross-Validation (GCV) terendah guna memperoleh model yang stabil dan akurat. Hasil penelitian menunjukkan bahwa model dengan kombinasi variabel Luas Panen Tanaman Padi dan Produktivitas Tanaman Padi yang ditetapkan sebagai komponen Spline Truncated dengan tiga titik knot, serta variabel Jumlah Penduduk dan Indeks Pembangunan Manusia sebagai komponen Kernel menghasilkan nilai GCV minimum sebesar 85.504.949, RMSE sebesar 242.723,6, dan R^2 sebesar 91,24%. Model ini berhasil menangkap pola hubungan non-linier dan memberikan estimasi yang lebih stabil. Implikasi dari temuan ini adalah model yang dihasilkan dapat digunakan untuk merancang kebijakan pertanian yang lebih efisien, dengan memperhatikan faktor-faktor yang berinteraksi secara dinamis dalam produksi beras.

ABSTRACT. This study aims to apply a nonparametric regression model using a mixed estimator of Truncated Spline and Kernel to estimate Rice Production in East Java Province. This model combines several predictor variables, namely Harvested Area of Rice Plants, Rice Productivity, Population, and Human Development Index. The selection of the best combination of variables is based on the lowest Generalized Cross-Validation (GCV) value to obtain a stable and accurate model. The results show that the model with a combination of variables Harvested Area of Rice Plants and Rice Productivity set as Truncated Spline components with three knot points, and Population and Human Development Index as Kernel components produces a minimum GCV value of 85,504,949, RMSE of 242,723.6, and R^2 of 91.24%. This model successfully captures non-linear relationship patterns and provides more stable estimates. The implication of this finding is that the resulting model can be used to design more efficient agricultural policies, by considering the factors that interact dynamically in rice production.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License. *Editorial of JJoM:* Department of Mathematics, Universitas Negeri Gorontalo, Jln. Prof. Dr. Ing. B. J. Habibie, Bone Bolango 96554, Indonesia.

1. Pendahuluan

Dalam banyak masalah dunia nyata, kita sering menghadapi situasi di mana lebih dari satu variabel saling berhubungan, namun hubungan di antara variabel-variabel tersebut sulit diprediksi atau dipisahkan dengan jelas menggunakan model yang sederhana [1, 2]. Fenomena ini sering kali membuat prediksi dan estimasi menjadi sangat menantang [3]. Salah satu metode yang banyak digunakan untuk memahami hubungan antara variabel-variabel ini adalah analisis regresi [4]. Tujuan dari analisis regresi adalah untuk mengidentifikasi pola hubungan antara satu atau lebih variabel yang disebut variabel respon dan variabel prediktor, sehingga kita bisa memprediksi atau mengestimasi nilai tertentu [5, 6]. Namun, seiring berjalannya waktu, kita semakin menyadari

bahwa masalah hubungan antar variabel semakin kompleks dan membutuhkan pendekatan yang lebih fleksibel dan tepat untuk menangani kerumitan data yang ada [7–9].

Meski banyak pendekatan dalam analisis regresi telah berkembang, baik dalam bentuk regresi parametrik, nonparametrik, maupun semiparametrik, tantangan terbesar tetap ada pada pemilihan model yang paling sesuai dengan pola hubungan antar variabel dalam data yang sangat bervariasi [10–12]. Regresi parametrik, contohnya, mengandalkan asumsi bahwa hubungan antar variabel sudah diketahui, namun dalam banyak kasus data dunia nyata tidak selalu mengikuti pola yang dapat diprediksi [13]. Di sisi lain, regresi nonparametrik yang tidak mengasumsikan bentuk hubungan tertentu, lebih fleksibel dalam menangani pola data yang kompleks, tetapi sering kali memiliki keterbatasan dalam memodelkan data yang memiliki karakteristik prediktor yang sa-

*Penulis Korespondensi.

ngat berbeda-beda [14, 15]. Seringkali, meskipun regresi nonparametrik seperti spline truncated atau kernel dapat efektif untuk menangani pola yang bervariasi, pemaksaan penggunaan satu jenis estimator untuk semua prediktor dapat menghasilkan model yang tidak akurat [16].

Penggunaan satu jenis estimator untuk memodelkan semua variabel prediktor, memungkinkan akan diperoleh model yang kurang tepat, artinya pola hubungan tidak sesuai dengan estimator yang digunakan. Untuk mengatasi masalah ini, pendekatan estimator campuran antara spline truncated dan kernel dikembangkan oleh Budiantara dkk. [14], meskipun penggunaannya sebelumnya terbatas pada satu prediktor per komponen spline truncated dan kernel. Penelitian lebih lanjut oleh Ratnasari dkk. [16] mengembangkan pendekatan ini dengan memungkinkan lebih dari satu prediktor digunakan dalam setiap komponen spline truncated dan kernel, memberikan fleksibilitas yang lebih besar dalam memodelkan hubungan yang lebih kompleks antar variabel.

Estimator campuran adalah pendekatan yang menggabungkan dua atau lebih metode estimasi untuk memperoleh hasil yang lebih baik dalam memodelkan data yang kompleks [5, 17–20]. Pendekatan ini bertujuan untuk mengatasi kelemahan yang mungkin ada pada metode tunggal dalam menangani pola atau karakteristik data yang bervariasi. Dalam konteks regresi nonparametrik, estimator campuran menggabungkan dua jenis estimator yang berbeda, misalnya spline truncated dan kernel, untuk mendapatkan model yang lebih fleksibel dan lebih akurat. Dengan menggabungkan kedua metode ini, estimator campuran dapat menghasilkan model yang lebih baik dibandingkan jika hanya menggunakan salah satu metode saja [21].

Kebaruan utama dari estimator campuran spline truncated dan kernel dalam penelitian ini adalah pendekatan estimasi yang mencoba berbagai kombinasi variabel prediktor untuk dimodelkan dengan kedua estimator tersebut. Dengan pendekatan ini, diharapkan dapat diperoleh model yang paling sesuai dengan data, yakni model yang paling mencerminkan hubungan antar variabel prediktor dan respon. Pemilihan model terbaik dilakukan dengan menggunakan Generalized Cross-Validation (GCV) yang minimal, yang memastikan bahwa model yang dihasilkan memiliki akurasi tinggi dan mampu menangani variabilitas dalam data [21]. Pendekatan ini sangat penting, karena sering kali pola hubungan antar variabel prediktor sangat bervariasi. Dengan mencoba berbagai kombinasi variabel dan memilih model dengan GCV terkecil, model yang diperoleh akan lebih sesuai dengan data yang ada, memberikan estimasi yang lebih akurat dan fleksibel.

Dalam praktik dunia nyata, seperti dalam pemodelan produksi beras di Jawa Timur di mana berbagai faktor seperti jumlah penduduk, luas panen padi, produktivitas tanaman padi, dan indeks pembangunan manusia saling mempengaruhi, pendekatan ini memungkinkan peneliti untuk memodelkan hubungan yang kompleks dan dinamis antara variabel-variabel tersebut dengan lebih efektif. Urgensi dari penelitian ini adalah kemampuannya untuk menghasilkan model regresi yang lebih tepat dan akurat. Estimator campuran ini diharapkan dapat memberikan kontribusi besar dalam meningkatkan kualitas estimasi produksi beras, yang sangat bergantung pada interaksi antara berbagai variabel. Dengan model yang lebih akurat, kebijakan pertanian yang lebih efisien bisa dirancang, dan langkah-langkah yang lebih efektif bi-

sa diambil untuk meningkatkan ketahanan pangan serta keberlanjutan pertanian di Jawa Timur.

Melalui pendekatan inovatif ini, diharapkan penelitian ini dapat memberikan kontribusi signifikan dalam pengembangan metode regresi nonparametrik dan aplikasinya di berbagai bidang lainnya, termasuk pertanian, ekonomi, dan sektor-sektor lain yang menghadapi tantangan hubungan antar variabel yang kompleks. Pendekatan ini juga memberikan gambaran bagaimana kita bisa mengatasi tantangan data yang sangat dinamis, dan membuka jalan bagi pemodelan yang lebih canggih dan akurat di masa depan.

2. Metode

2.1. Regresi Nonparametrik Spline Truncated

Fungsi spline adalah model polinomial tersegmen yang dirancang untuk mengatasi kelemahan pada penggunaan fungsi polinomial global. Untuk meningkatkan fleksibilitasnya, fungsi spline yang telah dimodifikasi, yaitu fungsi spline truncated [22, 23]. Fungsi spline truncated ini mempertahankan sifat dasar polinomial namun dengan penambahan titik-titik knot yang memungkinkan perubahan kurva pada interval yang berbeda-beda [24, 25]. Fungsi spline truncated dengan derajat m dan knot K_1, K_2, \dots, K_r dinyatakan pada pers. (1):

$$f(x_i) = \sum_{j=0}^m \varphi_j x_i^j + \sum_{k=1}^r \varphi_{m+k} (x_i - K_k)_+^m. \quad (1)$$

Model regresi nonparametrik spline truncated dituliskan pada pers. (2):

$$y_i = \sum_{j=0}^m \varphi_j x_i^j + \sum_{k=1}^r \varphi_{m+k} (x_i - K_k)_+^m + \varepsilon_i, \quad (2)$$

dengan fungsi truncated diberikan oleh:

$$(x_i - K_k)_+^m = \begin{cases} (x_i - K_k)^m & , \quad x \geq K_k \\ 0 & , \quad x < K_k \end{cases}.$$

Model regresi pada pers. (2), dapat ditulis dalam bentuk matriks pada pers. (3):

$$\mathbf{y} = \mathbf{X}(\mathbf{K}) \boldsymbol{\varphi} + \boldsymbol{\varepsilon}, \quad (3)$$

di mana:

- \mathbf{y} : vektor variabel respon berukuran $(n \times 1)$,
- $\mathbf{X}(\mathbf{K})$: matriks dari fungsi basis spline berukuran $(n \times (m + r + 1))$,
- $\boldsymbol{\varphi}$: vektor parameter koefisien regresi yang akan diestimasi berukuran $((m + r + 1) \times 1)$,
- $\boldsymbol{\varepsilon}$: vektor error acak berukuran $(n \times 1)$.

Estimasi parameter $\boldsymbol{\varphi}$ dapat dilakukan dengan menggunakan metode *Least Squares* (LS) yang meminimumkan jumlah kuadrat error dengan hasil estimasi parameter dituliskan pada pers. (4):

$$\hat{\boldsymbol{\varphi}} = \left(\mathbf{X}(\mathbf{K})^T \mathbf{X}(\mathbf{K}) \right)^{-1} \mathbf{X}(\mathbf{K})^T \mathbf{y}. \quad (4)$$

2.2. Regresi Nonparametrik Kernel

Misalkan diberikan data berpasangan q_i dan y_i yang mengikuti model regresi nonparametrik kernel. Model umum regresi nonparametrik kernel dituliskan pada pers. (5):

$$y_i = h(q_i) + \varepsilon_i, \quad (5)$$

di mana fungsi regresi $h(q_i)$ merupakan kurva yang tidak diketahui pola hubungannya. Estimator kernel adalah estimator yang digunakan untuk memperkirakan fungsi regresi $h(q_i)$ dengan menggunakan suatu kernel yang memberikan bobot lebih besar pada titik data yang lebih dekat dengan titik estimasi [26, 27]. Estimasi fungsi regresi kernel diberikan pada pers. (6):

$$h(q_i) = n^{-1} \sum_{i=1}^n \left(\frac{K_\alpha(q - q_i)}{n^{-1} \sum_{i=1}^n K_\alpha(q - q_i)} \right) y_i$$

$$h(q_i) = n^{-1} \left(\sum_{i=1}^n W_{\alpha i}(q) \right) y_i, \tag{6}$$

di mana:

$$W_{\alpha i}(q) = \frac{K_\alpha(q - q_i)}{n^{-1} \sum_{i=1}^n K_\alpha(q - q_i)},$$

$$K_\alpha(q - q_i) = \frac{1}{\alpha} K \left(\frac{q - q_i}{\alpha} \right).$$

K merupakan fungsi kernel. Pada penelitian ini fungsi kernel yang digunakan adalah Gaussian, yang dinyatakan sebagai berikut:

$$K(z) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} z^2 \right); \quad I_{[-\infty, \infty]}(z), \tag{7}$$

dengan $z = \frac{q - q_i}{\alpha}, i = 1, 2, \dots, n, \alpha$ adalah bandwidth, q adalah suatu nilai yang ditentukan dari variabel prediktor dan q_i adalah nilai ke- i dari variabel prediktor. Estimator kernel sangat bergantung pada fungsi kernel K dan bandwidth. Berdasarkan pers. (7) yang berlaku untuk setiap $q = q_1; q = q_2; \dots; q = q_3$, maka diperoleh bentuk matriks sebagai berikut:

$$\begin{bmatrix} \hat{h}(q_1) \\ \hat{h}(q_2) \\ \vdots \\ \hat{h}(q_n) \end{bmatrix} = \begin{bmatrix} n^{-1} \sum_{i=1}^n W_{\alpha i}(q_1) y_i \\ n^{-1} \sum_{i=1}^n W_{\alpha i}(q_2) y_i \\ \vdots \\ n^{-1} \sum_{i=1}^n W_{\alpha i}(q_n) y_i \end{bmatrix}$$

$$= \begin{bmatrix} n^{-1} W_{\alpha 1}(q_1) & n^{-1} W_{\alpha 2}(q_1) & \dots & n^{-1} W_{\alpha n}(q_1) \\ n^{-1} W_{\alpha 1}(q_2) & n^{-1} W_{\alpha 2}(q_2) & \dots & n^{-1} W_{\alpha n}(q_2) \\ \vdots & \vdots & \ddots & \vdots \\ n^{-1} W_{\alpha 1}(q_n) & n^{-1} W_{\alpha 2}(q_n) & \dots & n^{-1} W_{\alpha n}(q_n) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\hat{h}_\alpha(q) = D(\alpha) y. \tag{8}$$

Berdasarkan pers. (8), maka diperoleh estimator kernel yaitu $\hat{h}_\alpha(q) = D(\alpha) y$. Vektor $\hat{h}_\alpha(q)$ berukuran $(n \times 1)$, vektor y berukuran $(n \times 1)$ dan matriks $D(\alpha)$ berukuran $(n \times n)$ [21].

2.3. Estimator Campuran Spline Truncated – Kernel

Misalkan diberikan data berpasangan $(x_1, x_2, \dots, x_p, q_1, q_2, \dots, q_q)$ dan variabel respon y_i yang

mengikuti model regresi nonparametrik estimator campuran:

$$y_i = f(x_1, x_2, \dots, x_p) + h(q_1, q_2, \dots, q_q) + \varepsilon_i,$$

$$y_i = \sum_{p=1}^P f_p(x_{pi}) + \sum_{q=1}^Q h_q(q_{qi}) + \varepsilon_i \rightarrow y_i = \mu(x_i, q_i) + \varepsilon_i. \tag{9}$$

Estimator campuran pada pers. (9) dapat disajikan dalam bentuk matriks sebagai berikut:

$$y = X(K) \varphi + D(\alpha) y. \tag{10}$$

Estimasi parameter φ diperoleh dengan metode LS dan dinyatakan pada pers. (11):

$$\hat{\varphi} = \left(X(K)^T X(K) \right)^{-1} X(K)^T (I - D(\alpha)) y,$$

$$\hat{\varphi} = C(K, \alpha) y, \tag{11}$$

dengan $C(K, \alpha) = \left(X(K)^T X(K) \right)^{-1} X(K)^T (I - D(\alpha))$.

Estimator kurva regresi nonparametrik spline truncated dapat dituliskan pada pers. (12):

$$\sum_{p=1}^P f_p(x_{pi}) = X(K) \hat{\varphi} \rightarrow$$

$$X(K) \left(\left(X(K)^T X(K) \right)^{-1} X(K)^T (I - D(\alpha)) y \right), \tag{12}$$

di mana

$$A(K, \alpha) = X(K) \left(\left(X(K)^T X(K) \right)^{-1} X(K)^T (I - D(\alpha)) \right),$$

dan

$$\sum_{q=1}^Q h_q(q_{qi}) = D(\alpha) y, \tag{13}$$

sehingga estimator campuran regresi nonparametrik spline truncated dan kernel dituliskan pada pers. (14):

$$\mu(x_i, q_i) = A(K, \alpha) + D(\alpha) y. \tag{14}$$

Matriks $A(K, \alpha)$ dan $D(\alpha)$ sangat dependen dengan titik knot dan bandwidth optimal. Untuk memperoleh estimator campuran yang optimal, perlu dilakukan pemilihan titik knot dan bandwidth yang optimal menggunakan Generalized Cross-Validation (GCV). Modifikasi GCV untuk estimator campuran dituliskan pada pers. (15):

$$GCV(K_{opt}, \alpha_{opt}) = \left[\frac{MSE(K, \alpha)}{n^{-1} tr(I - A(K, \alpha) + D(\alpha))^2} \right], \tag{15}$$

di mana:

$$MSE(K, \alpha) = n^{-1} \sum_{i=1}^n y_i - \hat{y}_i,$$

dan \hat{y}_i adalah taksiran variabel respon berukuran $(n \times 1)$ dari model estimator campuran.

Tabel 1. Rincian variabel penelitian

Variabel	Notasi	Definisi Operasional	Satuan
Produksi Beras	Y	Jumlah total beras yang diproduksi di setiap kabupaten/kota di Provinsi Jawa Timur selama periode tertentu.	Ton
Jumlah Penduduk	X_1	Total jumlah penduduk yang tinggal di setiap kabupaten/kota di Provinsi Jawa Timur pada tahun yang relevan.	Orang
Produktivitas Tanaman Padi	X_2	Rata-rata jumlah gabah (padi) yang dihasilkan per satuan luas lahan yang digunakan untuk menanam padi di setiap kabupaten/kota.	Kuintal / Hektar (Ku/ha)
Luas Panen Tanaman Padi	X_3	Luas total lahan yang dipanen untuk tanaman padi di setiap kabupaten/kota di Provinsi Jawa Timur pada periode yang relevan.	Hektar (Ha)
Indeks Pembangunan Manusia (IPM)	X_4	Ukuran yang mengindikasikan tingkat kesejahteraan dan kualitas hidup penduduk di suatu kabupaten/kota berdasarkan dimensi kesehatan, pendidikan, dan standar hidup.	Indeks

Sumber: Badan Pusat Statistik Provinsi Jawa Timur (2024)

2.4. Kriteria Kebaikan Model

Kriteria kebaikan yang digunakan pada penelitian ini selain GCV, adalah Root Mean Squared Error (RMSE) dan Koefisien Determinasi (R^2). RMSE adalah akar kuadrat dari MSE dan ini mengubah kesalahan kuadrat menjadi satuan yang sama dengan data aslinya. Formula RMSE dituliskan pada pers. (16):

$$RMSE = \sqrt{MSE} \rightarrow \sqrt{n^{-1} \sum_{i=1}^n y_i - \hat{y}_i}. \quad (16)$$

Semakin kecil nilai RMSE, semakin baik model dan bisa diinterpretasikan model memiliki prediksi yang lebih akurat. Selain itu digunakan Koefisien Determinasi untuk mengukur proporsi variansi dalam variabel respon yang dapat dijelaskan oleh model regresi. Formula R^2 dituliskan pada pers. (17):

$$R^2 = 1 - \frac{SSE}{SST}, \quad (17)$$

di mana SSE adalah jumlah kuadrat residual dan SST adalah jumlah kuadrat total.

2.5. Data, Sumber Data, dan Variabel Penelitian

Pada penelitian ini, fokus utama adalah produksi beras dengan prediktor meliputi jumlah penduduk, produktivitas tanaman padi, luas panen tanaman padi, dan indeks pembangunan manusia (IPM). Semua data yang digunakan dalam penelitian direkapitulasi dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur Tahun 2024, dengan unit observasi berupa Kabupaten/Kota di Provinsi Jawa Timur sebanyak 38. Definisi operasional dari variabel penelitian dituliskan pada Tabel 1.

2.6. Tahapan Penelitian

Tahapan penelitian untuk estimasi produksi beras dengan estimator campuran spline truncated – kernel di Jawa Timur dirincikan sebagai berikut:

1. Melakukan eksplorasi pada data dengan membuat Box Plot dan Bar Plot;
2. Melakukan identifikasi pola hubungan menggunakan diagram pencar;
3. Titik knot yang dicobakan untuk estimator spline truncated 1 hingga 3 titik knot;

4. Titik knot pada estimator spline truncated ditentukan dengan membagi rentang setiap variabel spline menjadi 30 titik merata, lalu menghilangkan titik awal dan akhir sehingga tersedia 28 kandidat; untuk setiap variabel spline dicobakan 1 sampai 3 titik knot menggunakan kombinasi tanpa pengulangan sesuai kebutuhan model. Grid bandwidth untuk komponen kernel dibuat dengan membagi rentang setiap variabel kernel menjadi 30 titik, menghilangkan titik ekstrem sehingga terdapat 28 kandidat bandwidth per variabel. Seluruh kombinasi parameter dievaluasi secara eksensif melalui grid search;
5. Membuat kombinasi setiap variabel prediktor yang di setting untuk dimodelkan dengan estimator spline truncated dan Kernel pada model estimator campuran spline truncated – Kernel;
6. Melakukan estimasi parameter model estimator campuran spline truncated – kernel dengan pers. (11) hingga pers. (14);
7. Menghitung nilai GCV, RMSE, dan R^2 untuk setiap kombinasi variabel prediktor pada model estimator campuran spline truncated – kernel dengan pers. (15) hingga pers. (17);
8. Memilih model terbaik estimator campuran spline truncated – kernel berdasarkan nilai GCV yang minimum.

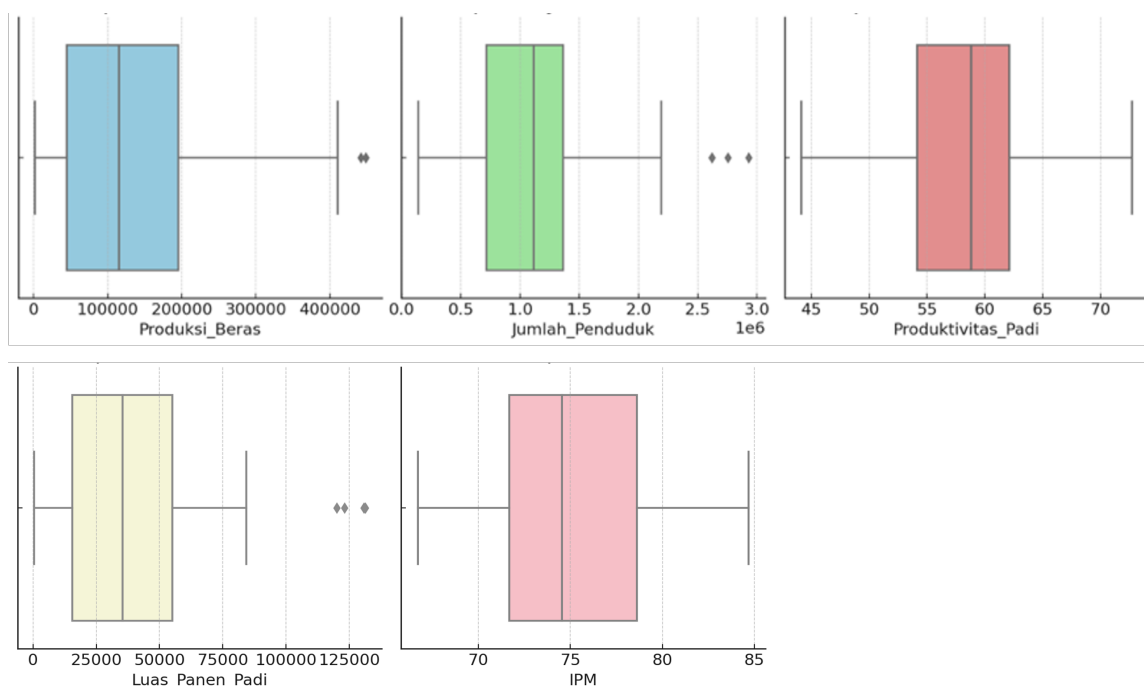
3. Hasil dan Pembahasan

3.1. Eksplorasi Data

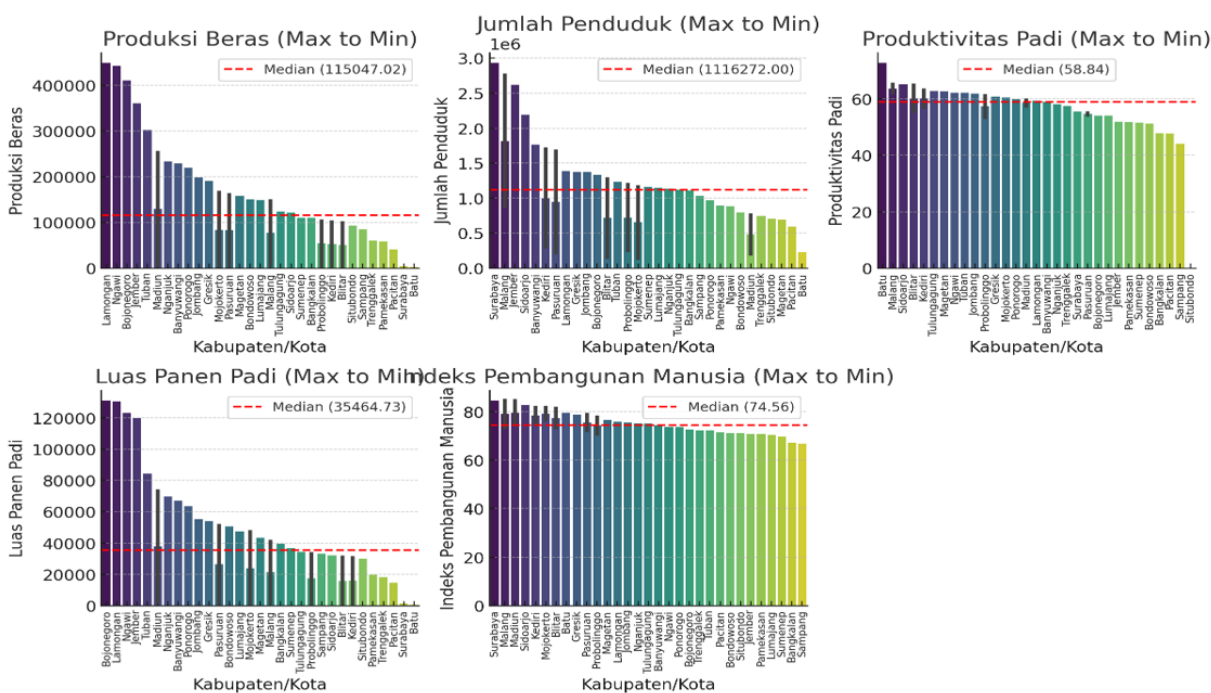
Eksplorasi data untuk variabel respon dan prediktor pada penelitian ini menjadi langkah awal dalam analisis untuk mengetahui informasi awal dan gambaran umum. Eksplorasi menggunakan Box Plot pada Gambar 1 dan Bar Plot pada Gambar 2.

Berdasarkan Gambar 1 dan 2, dilakukan analisis statistika deskriptif terhadap lima variabel utama. Variabel tersebut meliputi Produksi Beras, Jumlah Penduduk, Produktivitas Padi, Luas Panen Padi, dan Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Timur. Hasil analisis menunjukkan adanya ketimpangan di beberapa wilayah, khususnya antara kabupaten/kota besar dengan kabupaten/kota yang lebih kecil di provinsi tersebut.

Dalam hal Produksi Beras, terdapat konsentrasi yang jelas di kabupaten/kota besar seperti Kabupaten Lamongan dan Kabupaten Blitar, yang menunjukkan produksi beras yang jauh lebih tinggi dibandingkan dengan daerah lainnya. Analisis menggunakan Box Plot yang ditampilkan pada Gambar 1 menunjukkan bah-



Gambar 1. Box Plot untuk setiap variabel penelitian



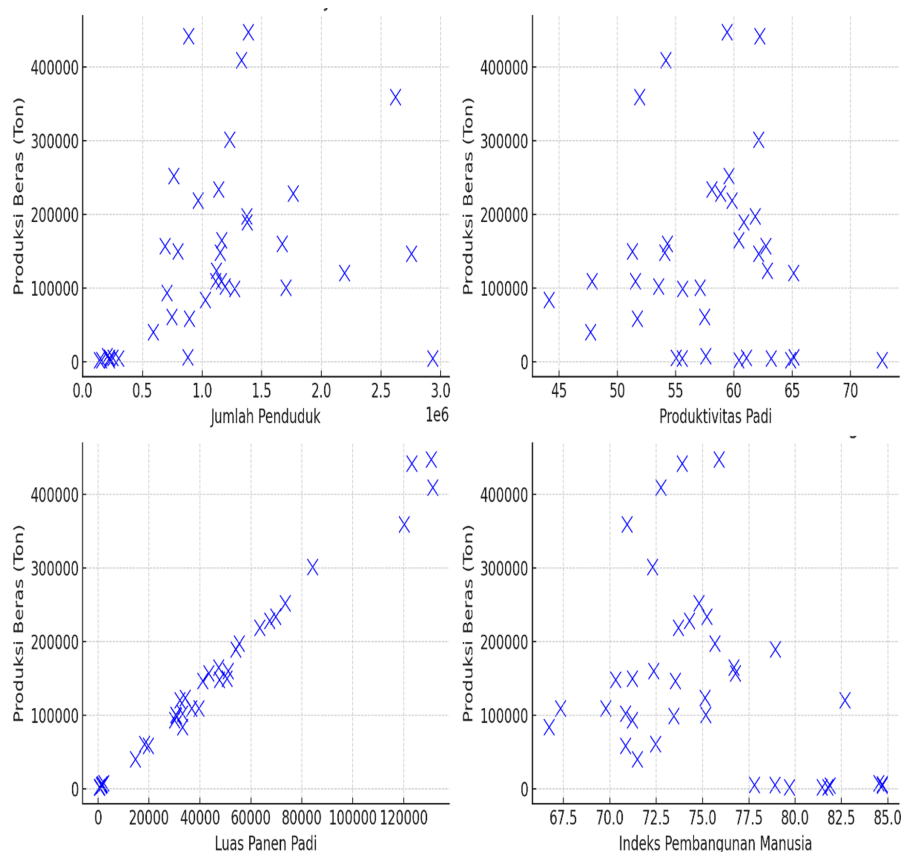
Gambar 2. Bar Plot untuk setiap variabel penelitian dengan nilai batas Median

wa sebagian besar kabupaten/kota memiliki produksi beras yang lebih rendah, dengan outlier yang mencerminkan kabupaten dengan kapasitas produksi yang sangat tinggi. Ketimpangan produksi ini mengindikasikan adanya kesenjangan dalam kapasitas produksi antar kabupaten/kota di Provinsi Jawa Timur.

Pada variabel Jumlah Penduduk, Kota Surabaya menjadi kabupaten/kota dengan jumlah penduduk tertinggi, yang jauh melampaui kabupaten/kota lainnya, seperti yang terlihat dalam Bar Plot pada Gambar 2. Hal ini sejalan dengan fakta bahwa kota-kota besar, seperti Surabaya dan Malang, memiliki konsentrasi pendu-

duk yang lebih padat dibandingkan dengan daerah lainnya. Box plot pada Gambar 1 menunjukkan bahwa sebagian besar kabupaten/kota memiliki jumlah penduduk yang lebih rendah dari rata-rata, dengan beberapa kabupaten besar yang menjadi outlier di sisi kanan, mencerminkan konsentrasi penduduk yang tinggi di kota-kota besar.

Mengenai Produktivitas Padi, terlihat bahwa Kota Batu dan Kabupaten Blitar memiliki produktivitas padi yang lebih tinggi dibandingkan kabupaten/kota lainnya. Box Plot pada Gambar 1 menunjukkan distribusi yang cukup sempit, mengindikasikan bah-



Gambar 3. Diagram Pencar antara variabel respon terhadap masing-masing variabel prediktor

wa sebagian besar kabupaten/kota memiliki produktivitas yang seragam, dengan sedikit variasi. Walaupun terdapat kabupaten-kabupaten yang memiliki produktivitas padi di bawah rata-rata, perbedaan antar kabupaten/kota dalam hal ini relatif lebih kecil dibandingkan dengan produksi beras atau jumlah penduduk.

Untuk variabel Luas Panen Padi, terlihat bahwa kabupaten seperti Banyuwangi dan Lamongan memiliki luas panen padi yang jauh lebih besar dibandingkan kabupaten/kota lainnya. Box plot pada Gambar 1 menunjukkan distribusi yang cukup lebar, dengan beberapa kabupaten/kota memiliki luas panen yang sangat besar (*outlier*), sementara sebagian besar kabupaten/kota lainnya memiliki luas panen yang lebih kecil. Ketimpangan dalam distribusi luas panen padi ini mencerminkan ketidakmerataan dalam penggunaan lahan pertanian untuk padi di Provinsi Jawa Timur. Hal ini selaras dengan faktor geografis yang berbeda-beda antar Kabupaten/Kota di Provinsi Jawa Timur.

Selanjutnya, Indeks Pembangunan Manusia (IPM) menunjukkan distribusi yang lebih merata di Provinsi Jawa Timur. Box Plot pada Gambar 1 dan Bar Plot pada Gambar 2 menunjukkan bahwa sebagian besar kabupaten/kota memiliki nilai IPM yang berkisar antara 70 hingga 80, dengan Surabaya dan Malang memiliki IPM yang sedikit lebih tinggi. Tidak ada outlier yang signifikan dalam distribusi IPM, yang menunjukkan bahwa tingkat kesejahteraan dan kualitas hidup di Provinsi Jawa Timur relatif seragam dibandingkan dengan variabel lainnya.

Berdasarkan temuan yang menunjukkan ketimpangan antara kabupaten/kota di Provinsi Jawa Timur, terutama dalam hal Produksi Beras, Jumlah Penduduk, dan Luas Panen Padi, penelitian ini menyoroti pentingnya penggunaan model yang lebih aku-

rat untuk memperkirakan produksi beras. Dengan adanya ketimpangan ini, estimasi yang lebih tepat sangat diperlukan untuk merumuskan kebijakan yang lebih efektif. Pendekatan estimator campuran menjadi inovatif serta krusial dalam penelitian ini, karena dapat mengakomodasi interaksi kompleks antar variabel-variabel prediktor, seperti Jumlah Penduduk, Produktivitas Padi, Luas Panen Padi, dan IPM. Estimator campuran ini diharapkan dapat memberikan estimasi yang lebih akurat dalam memperkirakan produksi beras, sehingga strategi kebijakan pertanian yang lebih efisien dan berbasis data dapat dirancang. Pendekatan ini juga dapat meningkatkan ketahanan pangan dan keberlanjutan pertanian di Jawa Timur, serta memberikan kontribusi terhadap pengembangan metode regresi nonparametrik dalam berbagai bidang.

3.2. Identifikasi Pola Hubungan dengan Diagram Pencar

Identifikasi pola hubungan dalam analisis regresi, khususnya dalam regresi nonparametrik dengan estimator campuran, sangat penting untuk memahami dan menangkap pola yang lebih kompleks antara variabel-variabel prediktor terhadap variabel respon. Diagram pencar ditampilkan pada Gambar 3.

Berdasarkan Gambar 3, pola hubungan antara Produksi Beras (Ton) dengan berbagai variabel prediktor lainnya menunjukkan dinamika yang cukup kompleks. Jumlah Penduduk menunjukkan adanya hubungan non-linier dengan Produksi Beras. Hal ini mengindikasikan bahwa meskipun jumlah penduduk mungkin memengaruhi produksi beras, pengaruhnya tidak bersifat langsung atau linear, dan ada faktor lain yang lebih dominan seperti kebijakan pertanian, ketersediaan sumber daya, serta faktor so-

Tabel 2. Hasil Pemodelan untuk setiap Kombinasi Variabel terhadap Estimator

Spline	Kernel	Kode Model	Jumlah Titik Knot	GCV	RMSE	Koefisien Determinasi
X_1	$X_2 ; X_3 ; X_4$	M1	1	6930816789	354536,10	66,72314
		M2	2	5661878280	353768,40	66,49865
		M3	3	5152013829	363705	64,37818
X_2	$X_1 ; X_3 ; X_4$	M4	1	5725340946	329978,50	70,849
		M5	2	5312337780	320694,50	74,53874
		M6	3	4894873642	325561,40	72,57043
X_3	$X_1 ; X_2 ; X_4$	M7	1	1440852280	162091,20	95,53628
		M8	2	973279535	180158,50	94,31164
		M9	3	883558584	151539,30	96,28271
X_4	$X_1 ; X_2 ; X_3$	M10	1	6728930000	346859,80	67,74623
		M11	2	5322398566	319931,60	76,15557
		M12	3	4813094152	320600,10	75,96256
$X_1 ; X_2$	$X_3 ; X_4$	M13	1	3987932462	271139,50	82,74302
		M14	2	3523328411	265958,60	84,10687
		M15	3	3270733508	257278,90	85,29399
$X_1 ; X_3$	$X_2 ; X_4$	M16	1	594223135	353966,10	79,11646
		M17	2	502229471	196762	94,06367
		M18	3	428609010	192401,60	94,22986
$X_1 ; X_4$	$X_2 ; X_3$	M19	1	5147401713	300342,80	75,67152
		M20	2	3591229735	260715,70	84,49
		M21	3	3229457654	260559	84,83969
$X_2 ; X_3$	$X_1 ; X_4$	M22	1	170158556	282799,90	88,59547
		M23	2	132074171	267551	89,50007
		M24	3	85504949	242723,60	91,24336
$X_2 ; X_4$	$X_1 ; X_3$	M25	1	3273105334	252969,90	85,49708
		M26	2	2475173787	238275,80	88,07041
		M27	3	2217921019	235538,20	88,43215
$X_3 ; X_4$	$X_1 ; X_2$	M28	1	266283390	220681,60	92,51657
		M29	2	170960318	186428,60	94,50331
		M30	3	158834555	174798,20	95,16502
$X_1 ; X_2 ; X_3$	X_4	M31	1	585721255	330835,50	85,90229
		M32	2	512830444	246074	91,41346
		M33	3	433781740	230312,50	92,28802
$X_1 ; X_3 ; X_4$	X_2	M34	1	556267334	251858,70	90,88705
		M35	2	471762996	244119,80	91,32406
		M36	3	404658851	239535,20	91,592
$X_1 ; X_2 ; X_4$	X_3	M37	1	299944141	74106,30	99,01965
		M38	2	225006299	70698,23	99,12281
		M39	3	201507358	66921,21	99,21745
$X_2 ; X_3 ; X_4$	X_1	M40	1	163688255	338357,80	85,28098
		M41	2	127976737	291224,60	88,3838
		M42	3	86125592	296319,40	88,08823

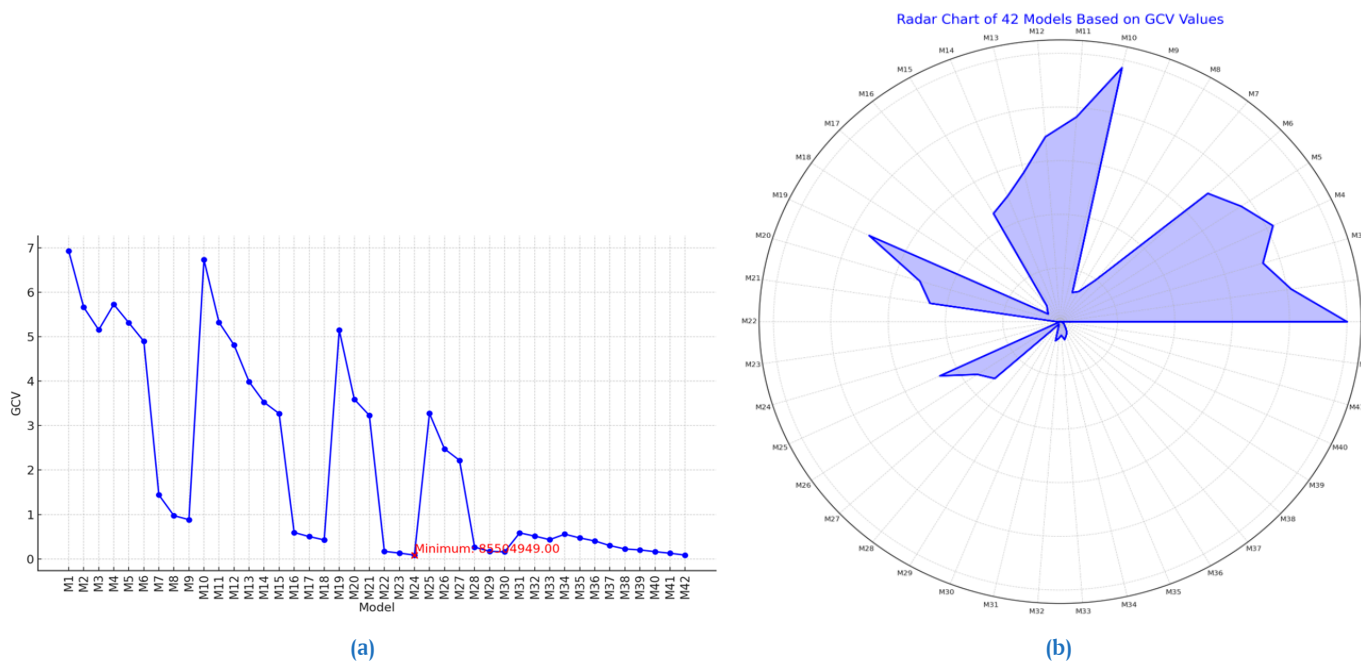
sial ekonomi lainnya. Sementara itu, Produktivitas Tanaman Padi (Ku/ha) juga menunjukkan hubungan yang bersifat non-linier dengan Produksi Beras, yang menunjukkan bahwa peningkatan produktivitas padi tidak selalu berbanding lurus dengan peningkatan hasil produksi beras. Hal ini menunjukkan bahwa meskipun produktivitas tanaman padi merupakan faktor penting, ia bukan satu-satunya faktor yang menentukan hasil akhir produksi beras yang optimal. Namun, Luas Panen Tanaman Padi (Ha) menunjukkan korelasi yang sangat kuat dengan Produksi Beras, yang menunjukkan hubungan yang lebih linear dan jelas. Hal ini mengindikasikan bahwa semakin luas area yang digunakan untuk bertani padi, semakin besar pula hasil produksi beras yang dihasilkan. Terakhir, meskipun terdapat kecenderungan positif antara Indeks Pembangunan Manusia (IPM) dan Produksi Beras, hubungan tersebut tidak sekuat hubungan dengan variabel-variabel lain, yang

semakin menunjukkan kompleksitas hubungan antara variabel-variabel ini. Dalam menghadapi hubungan-hubungan non-linier tersebut, penggunaan estimator campuran, seperti kombinasi antara spline truncated dan kernel, sangat relevan.

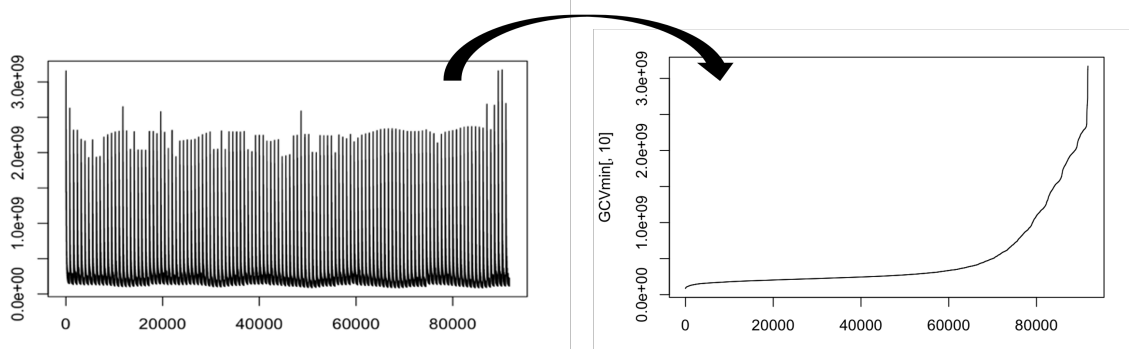
3.3. Pemodelan Produksi Beras dengan Estimator Campuran Spline Truncated – Kernel

Hasil pemodelan produksi beras dengan estimator campuran spline truncated–kernel dilakukan untuk kombinasi setiap variabel prediktor yang berubah-ubah. Setiap variabel ditentukan sebagai komponen spline truncated dan kernel. Perhitungan nilai GCV, RMSE, dan koefisien determinasi dilakukan menggunakan pers. (15) hingga pers. (17) secara berurutan, yang hasil selengkapnya ditampilkan pada Tabel 2.

Berdasarkan Tabel 2, diperoleh 14 kombinasi variabel yang



Gambar 4. Perubahan nilai GCV: (a) Diagram Garis (b) Radar Chart



Gambar 5. Visualisasi perubahan nilai GCV sebelum diurutkan dan sesudah diurutkan dari minimum ke maksimum

ditentukan sebagai komponen spline truncated dan kernel dalam model estimator campuran spline truncated-kernel. Pemilihan kombinasi variabel terbaik, jumlah titik knot, dan bandwidth optimal berdasarkan nilai GCV minimum. Visualisasi perubahan nilai GCV berdasarkan Tabel 2 ditampilkan pada Gambar 4.

Berdasarkan Gambar 4, diketahui model terbaik berdasarkan nilai GCV minimum terdapat pada variabel X_2 dan X_3 yang ditetapkan sebagai komponen spline truncated dengan tiga titik knot, sedangkan variabel X_1 dan X_4 komponen Kernel (M24). Berdasarkan Tabel 2, kita ketahui proses penentuan titik knot dan bandwidth memainkan peran penting dalam membangun model estimasi yang kompleks dengan menggunakan estimator spline truncated dan kernel.

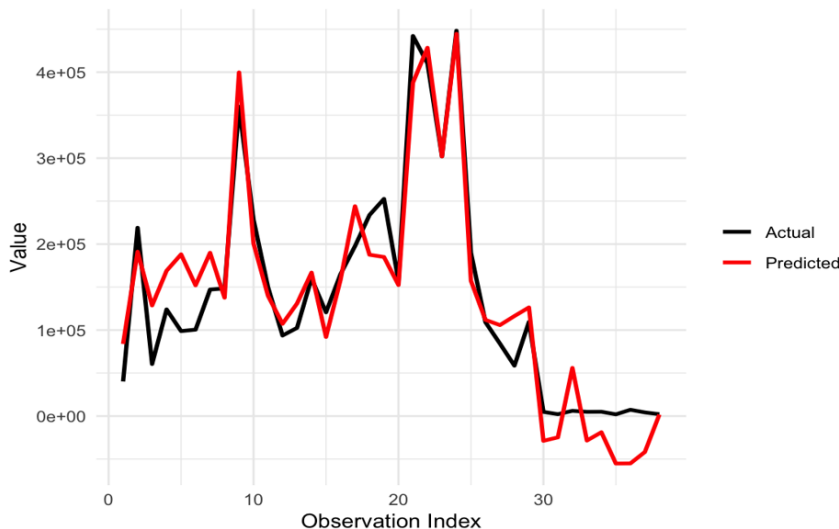
Titik knot digunakan dalam fungsi spline truncated untuk membagi rentang data menjadi beberapa segmen, di mana dalam setiap segmen, hubungan antara variabel prediktor dan variabel respon dapat diperlakukan secara terpisah. Titik-titik knot dihasilkan dengan cara menentukan sejumlah nilai dalam rentang minimum hingga maksimum untuk setiap variabel prediktor. Selanjutnya, bandwidth dihitung untuk fungsi Kernel Gaussian, yang mengontrol lebar pengaruh data pada model. Bandwidth menen-

tukan seberapa besar pengaruh setiap titik data terhadap estimasi model, dengan nilai bandwidth yang lebih besar mengindikasikan pengaruh yang lebih luas. Proses penentuan bandwidth dilakukan dengan menghitung urutan nilai dari 0 hingga selisih nilai maksimum dan minimum pada setiap variabel kernel yang ada. Dengan menggunakan titik knot dan bandwidth yang telah ditentukan, maka model akan mencakup berbagai kombinasi parameter yang akan diuji. Proses ini menghasilkan kombinasi yang sangat banyak, dalam penelitian ini diperoleh 91.728 kombinasi atau kemungkinan, sehingga memungkinkan untuk mengeksplorasi berbagai kemungkinan hubungan non-linier antara variabel prediktor dan variabel dependen.

Kombinasi sebanyak 91.728 diperoleh dari pencarian parameter optimal melalui grid search terhadap posisi titik knot untuk komponen spline dan parameter bandwidth untuk komponen kernel. Untuk spline, setiap variabel dibagi menjadi 30 titik kandidat, kemudian titik awal dan akhir dihilangkan sehingga tersisa 28 kandidat. Dengan $kn = 3$, jumlah kombinasi titik knot untuk dua variabel spline adalah $\binom{28}{3} = 3.276$ kemungkinan. Untuk kernel, setiap variabel memiliki 28 kandidat bandwidth, se-

Tabel 3. Titik Knot dan Bandwidth Optimal pada Model Estimator Campuran

Model Estimator Campuran Spline Truncated dan Kernel Terbaik	Titik Knot		Bandwidth	
	X_2	X_3	X_1	X_4
M24	60.88966	77130.63862	9.631028e+04	6.196552e-01
	63.84724	90653.11897		
	68.77655	113190.58621		



Gambar 6. Visualisasi perbandingan data aktual dan estimasi dari model estimator campuran spline truncated - kernel

hingga untuk dua variabel kernel terdapat $28 \times 28 = 784$ kemungkinan kombinasi. Perkalian kedua komponen ini menghasilkan total $3.276 \times 784 = 91.728$ kombinasi parameter yang dievaluasi. Nilai ini mencerminkan seluruh ruang pencarian parameter yang digunakan dalam prosedur estimasi berdasarkan pers. (11) hingga pers. (14).

Kombinasi berbagai titik knot dan bandwidth ini memperkenalkan kompleksitas pada model, karena setiap kombinasi menghasilkan fungsi spline truncated yang berbeda dan mempengaruhi cara kernel Gaussian beroperasi pada data. Model yang dihasilkan dari kombinasi yang berbeda ini diuji untuk menilai seberapa baik mereka mengestimasi variabel dependen melalui penghitungan GCV (Generalized Cross Validation). Setelah menghitung GCV untuk setiap kombinasi titik knot dan bandwidth, model dengan nilai GCV terendah akan dipilih sebagai model yang paling optimal. GCV adalah ukuran yang digunakan untuk menilai keseimbangan antara akurasi estimasi dan kompleksitas model. Terdapat 91.728 kombinasi untuk model estimator campuran spline truncated dan kernel yang ilustrasikan melalui time series plot nilai GCV pada Gambar 5.

Rincian hasil pemodelan dengan estimator campuran spline truncated dan kernel terbaik dengan GCV minimum ditampilkan pada Tabel 3. Selanjutnya, model estimasi produksi beras di Provinsi Jawa Timur dari model terbaik estimator campuran berdasarkan pers. (9) dapat dituliskan sebagai berikut:

$$\hat{y}_i = \{-1042,97 - 1065,86X_2 + 11455,68(X_2 - 60.8897)_+ - 6786,28(X_2 - 63.8472)_+ - 5512,99(X_2 - 68.7766)_+ + 0.951X_3 + 11,814(X_3 - 77130.64)_+ - 13,89(X_3 - 90653.12)_+ - 3,097(X_3 - 113190.59)_+\}$$

$$+ \left\{ \left(\frac{1}{38} \right) \sum_{i=1}^n \left(\frac{\frac{1}{9.63103e+04} K\left(\frac{X_1 - X_{1i}}{9.63103e+04}\right)}{\frac{1}{38} \sum_{i=1}^n \left(\frac{1}{9.63103e+04} K\left(\frac{X_1 - X_{1i}}{9.63103e+04}\right)\right)} \right) y_i \right\} + \left(\frac{1}{38} \right) \sum_{i=1}^n \left(\frac{\frac{1}{6.196552e-01} K\left(\frac{X_4 - X_{4i}}{6.19655e-01}\right)}{\frac{1}{38} \sum_{i=1}^n \left(\frac{1}{6.19655e-01} K\left(\frac{X_4 - X_{4i}}{6.19655e-01}\right)\right)} \right) y_i.$$

Visualisasi hasil estimasi dengan data aktual dari model estimator campuran spline truncated dan kernel terbaik ditampilkan pada Gambar 6.

Dalam grafik yang ditampilkan pada Gambar 6, garis merah mewakili estimasi nilai dari variabel respon (\hat{y}) yang dihasilkan oleh model estimator campuran, yang menggabungkan pendekatan spline dan kernel untuk menangkap hubungan yang lebih kompleks antar variabel. Secara umum, model ini mampu mengikuti pola data aktual (ditunjukkan oleh garis hitam) dengan baik, terutama pada sebagian besar titik data. Meskipun demikian, terdapat beberapa perbedaan signifikan antara hasil estimasi dengan nilai aktual, terutama pada titik-titik tertentu di mana terjadi lonjakan tajam pada indeks observasi sekitar ke-10 dan ke-20. Perbedaan ini menunjukkan bahwa model estimator campuran menghadapi kesulitan dalam menangkap fluktuasi besar atau perubahan mendadak dalam data. Secara keseluruhan, model estimator campuran menunjukkan kinerja yang baik, namun masih perlu penyempurnaan untuk menangani fluktuasi besar dalam data dengan lebih akurat.

3.4. Pembahasan

Pendekatan estimator campuran spline truncated dan kernel menawarkan kebaruan yang signifikan dalam memodelkan hubungan antar variabel prediktor dan respon, khususnya dalam

estimasi produksi beras. Dengan menggabungkan dua teknik estimasi nonparametrik, yaitu spline truncated dan kernel, model ini dapat menangkap pola hubungan yang lebih kompleks dan non-linier antara variabel-variabel prediktor seperti Jumlah Penduduk, Produktivitas Tanaman Padi, Luas Panen Tanaman Padi, dan Indeks Pembangunan Manusia dengan Produksi Beras. Kebaruan utama dari pendekatan ini adalah kemampuannya untuk mengoptimalkan model dengan menguji berbagai kombinasi variabel dan memilih titik knot dan bandwidth yang paling tepat menggunakan teknik Generalized Cross-Validation (GCV), yang memastikan model yang lebih akurat dan fleksibel.

Model terbaik estimator campuran diperoleh berdasarkan nilai GCV minimum dengan variabel X_2 dan X_3 yang ditetapkan sebagai komponen spline truncated dengan tiga titik knot, sedangkan variabel X_1 dan X_4 komponen kernel. Model terbaik pada penelitian ini mampu menangkap pola hubungan yang lebih kompleks dan non-linier dengan nilai GCV minimum sebesar 85.504.949. Nilai ini mencerminkan kemampuan model dalam menghindari overfitting dan memberikan estimasi yang lebih stabil serta lebih dapat digeneralisasi. Selain itu, nilai RMSE untuk model sebesar 242.723,60 dengan Koefisien Determinasi (R^2) mencapai 91,24%, yang menandakan bahwa model memiliki kemampuan yang sangat baik dalam menjelaskan hubungan antara variabel yang dianalisis. Secara keseluruhan, hasil yang ditunjukkan oleh model estimator campuran spline truncated - kernel ini memberikan wawasan yang lebih mendalam mengenai faktor-faktor yang mempengaruhi produksi beras dan pentingnya pendekatan fleksibel dalam memodelkan hubungan non-linier antara variabel-variabel prediktor. Keberhasilan model dalam menangkap pola-pola kompleks ini memperkuat relevansi penggunaan estimator campuran dalam analisis data yang memiliki karakteristik yang bervariasi dan tidak mengikuti pola linier sederhana.

4. Kesimpulan

Berdasarkan hasil penelitian yang menggunakan model estimator campuran spline truncated dan kernel, dapat disimpulkan bahwa model ini berhasil memberikan pendekatan yang lebih fleksibel dan akurat dalam mengestimasi Produksi Beras di Provinsi Jawa Timur. Kombinasi variabel prediktor yang diuji, yakni Luas Panen Tanaman Padi (Ha) dan Produktivitas Tanaman Padi (Ku/ha) pada komponen Spline Truncated serta Jumlah Penduduk dan Indeks Pembangunan Manusia (IPM) pada komponen Kernel, telah menghasilkan model dengan 91.728 kemungkinan yang diuji melalui perhitungan Generalized Cross Validation (GCV). Model terbaik yang ditemukan memiliki nilai GCV minimum sebesar 85.504.949, menunjukkan kemampuannya dalam menghindari overfitting dan menghasilkan estimasi yang stabil serta dapat digeneralisasi. Nilai RMSE sebesar 242.723,60 dan R^2 sebesar 91,24% menegaskan bahwa model ini memiliki kemampuan yang sangat baik dalam menjelaskan hubungan antara variabel-variabel yang dianalisis.

Namun, penelitian ini juga mengakui adanya beberapa keterbatasan, terutama terkait dengan kompleksitas komputasi yang tinggi akibat banyaknya kombinasi titik knot dan bandwidth yang dicoba. Selain itu, model ini sangat bergantung pada pemilihan parameter yang tepat, yang dapat mempengaruhi hasil estimasi. Untuk penelitian selanjutnya, disarankan untuk melakukan validasi silang dengan data yang lebih bervariasi dan jumlah sam-

pel yang lebih besar guna meningkatkan ketahanan dan aplikasi model dalam skala yang lebih luas.

Kontribusi Penulis. A. T. R. Dani: Konseptualisasi, Metodologi, Analisis Formal, Penulisan – Persiapan Draft Asli, Visualisasi. F. B. Putra: Kurasi Data, Perangkat Lunak, Penulisan – Tinjauan dan Penyuntingan, Visualisasi. I. N. Budiantara: Metodologi, Validasi, Supervisi. V. Ratnasari: Metodologi, Validasi, Supervisi. Semua penulis telah membaca dan menyetujui versi manuskrip yang diterbitkan.

Ucapan Terima Kasih. Para penulis mengucapkan terima kasih kepada Universitas Mulawarman dan Institut Teknologi Sepuluh Nopember yang telah memfasilitasi kolaborasi dalam penulisan dan diseminasi penulisan Karya Ilmiah. Penulis juga mengucapkan terima kasih kepada editor dan reviewer atas pembacaan yang teliti, masukan yang membangun, dan saran-saran yang membantu dalam menyempurnakan tulisan ini.

Pembiayaan. Penelitian ini tidak menerima pembiayaan eksternal.

Konflik Kepentingan. Para penulis menyatakan tidak ada konflik kepentingan yang terkait dengan artikel ini.

Ketersediaan Data. Seluruh data bersumber dari Badan Pusat Statistik (BPS) Jawa Timur. Data publik tersedia melalui situs resmi BPS, sedangkan data detail atau granular diperoleh melalui permintaan resmi kepada BPS sesuai prosedur layanan data yang berlaku.

Referensi

- [1] X. Zhang, J. Ru, and C. Wu, "A Nonparametric Regression-Based Multi-Scale Gradient Correlation Filtering Method for Infrared Small Target Detection," *Electronics (Switzerland)*, vol. 12, no. 7, Apr. 2023, doi: [10.3390/electronics12071562](https://doi.org/10.3390/electronics12071562).
- [2] V. Ratnasari, S. H. Utama, and A. T. R. Dani, "Toward Sustainable Development Goals (SDGs) with Statistical Modeling: Recursive Bivariate Binary Probit," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 8, pp. 1515–1521, 2024.
- [3] I. N. Budiantara et al., "Modeling Percentage of Poor People in Indonesia Using Kernel and Fourier Series Mixed Estimator in Nonparametric Regression," *Revista Investigacion Operacional*, vol. 40, no. 4, pp. 538–550, 2019.
- [4] P. Čížek and S. Sadıkoğlu, "Robust nonparametric regression: A review," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 12, no. 3, pp. 1–16, May 2020, doi: [10.1002/wics.1492](https://doi.org/10.1002/wics.1492).
- [5] A. T. R. Dani, V. Ratnasari, and I. N. Budiantara, "Optimal Knots Point and Bandwidth Selection in Modeling Mixed Estimator Nonparametric Regression," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1115, no. 1, p. 012020, Mar. 2021, doi: [10.1088/1757-899x/1115/1/012020](https://doi.org/10.1088/1757-899x/1115/1/012020).
- [6] V. Ratnasari, Purhadi, I. C. Aviantholib, and A. T. R. Dani, "Parameter Estimation and Hypothesis Testing The Second Order of Bivariate Binary Logistic Regression (S-BBLR) Model With Berndt Hall-Hall-Hausman (BHHH) Iterations," *Commun. Math. Biol. Neurosci.*, vol. 2022, 2022, doi: [10.28919/cmbn/7258](https://doi.org/10.28919/cmbn/7258).
- [7] E. A. Souza-Rodrigues, "Nonparametric regression with common shocks," *Econometrics*, vol. 4, no. 3, Sep. 2016, doi: [10.3390/econometrics4030036](https://doi.org/10.3390/econometrics4030036).
- [8] L. R. Cheruiyot, "Local linear regression estimator on the boundary correction in nonparametric regression estimation," *J. Stat. Theory Appl.*, vol. 19, no. 3, pp. 460–471, Sep. 2020, doi: [10.2991/jsta.d.201016.001](https://doi.org/10.2991/jsta.d.201016.001).
- [9] E. Seijo and B. Sen, "Nonparametric least squares estimation of a multivariate convex regression function," *Ann. Stat.*, vol. 39, no. 3, pp. 1633–1657, 2011, doi: [10.1214/10-AOS852](https://doi.org/10.1214/10-AOS852).
- [10] S. Sifriyani, A. T. R. Dani, M. Fauziyah, M. N. Hayati, S. Wahyuningsih, and S. Prangga, "Spline and Kernel Mixed Estimators In Multivariable Nonparametric Regression For Dengue Hemorrhagic Fever Model," *Commun. Math. Biol. Neurosci.*, vol. 2023, pp. 1–15, 2023.
- [11] S. Sifriyani, A. T. R. Dani, M. Fauziyah, and I. N. Budiantara, "Statistical Modeling: A New Regression Curve Approximation using Mixed Estimators Truncated Spline and Epanechnikov Kernel," *Engineering Letters*, vol. 31, no. 4, pp. 1–7, 2023.

- [12] R. Rahmania, S. Sifriyani, and A. T. R. Dani, "Modeling Open Unemployment Rate in Kalimantan Island Using Nonparametric Regression With Fourier Series Estimator," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 1, pp. 0245–0254, Mar. 2024, doi: [10.30598/barekengvol18iss1pp0245-0254](https://doi.org/10.30598/barekengvol18iss1pp0245-0254).
- [13] M. Rifada, N. Chamidah, R. A. Ningrum, and L. Muniroh, "Stunting Determinants Among Toddlers in Probolinggo District of Indonesia Using Parametric and Nonparametric Ordinal Logistic Regression Models," *Communications in Mathematical Biology and Neuroscience*, vol. 2023, 2023, doi: [10.28919/cmbn/6690](https://doi.org/10.28919/cmbn/6690).
- [14] I. N. Budiantara, V. Ratnasari, M. Ratna, and I. Zain, "The Combination of Spline and Kernel Estimator for Nonparametric Regression and its Properties," *Applied Mathematical Sciences*, vol. 9, no. 122, pp. 6083–6094, 2015, doi: [10.12988/ams.2015.58517](https://doi.org/10.12988/ams.2015.58517).
- [15] I. Wayan Sudiarsa, I. Nyoman Budiantara, S. Suhartono, and S. W. Purnami, "Combined estimator Fourier series and spline truncated in multivariable nonparametric regression," *Applied Mathematical Sciences*, vol. 9, no. 97–100, pp. 4997–5010, 2015, doi: [10.12988/ams.2015.55394](https://doi.org/10.12988/ams.2015.55394).
- [16] V. Ratnasari, N. Budiantara, M. Ratna, and I. Zain, "Estimation of Nonparametric Regression Curve using Mixed Estimator of Multivariable Truncated Spline and Multivariable Kernel," *Global Journal of Pure and Applied Mathematics*, vol. 12, no. 6, pp. 5047–5057, 2016.
- [17] N. Chamidah, B. Lestari, I. N. Budiantara, and D. Aydin, "Estimation of Multiresponse Multipredictor Nonparametric Regression Model Using Mixed Estimator," *Symmetry (Basel)*, vol. 16, no. 4, pp. 1–25, Apr. 2024, doi: [10.3390/sym16040386](https://doi.org/10.3390/sym16040386).
- [18] M. A. D. Octavanny, I. N. Budiantara, H. Kuswanto, and D. P. Rahmawati, "A New Mixed Estimator in Nonparametric Regression for Longitudinal Data," *Journal of Mathematics*, vol. 2021, 2021, doi: [10.1155/2021/3909401](https://doi.org/10.1155/2021/3909401).
- [19] D. P. Rahmawati, I. N. Budiantara, D. D. Prastyo, and M. A. D. Octavanny, "Mixed Spline Smoothing and Kernel Estimator in Biresponse Nonparametric Regression," *International Journal of Mathematics and Mathematical Sciences*, vol. 2021, 2021, doi: [10.1155/2021/6611084](https://doi.org/10.1155/2021/6611084).
- [20] N. P. A. M. Mariati, I. N. Budiantara, and V. Ratnasari, "The Application of Mixed Smoothing Spline and Fourier Series Model in Nonparametric Regression," *Symmetry (Basel)*, vol. 13, no. 11, Nov. 2021, doi: [10.3390/sym13112094](https://doi.org/10.3390/sym13112094).
- [21] V. Ratnasari, N. Budiantara, and A. T. R. Dani, "Nonparametric Regression Mixed Estimators of Truncated Spline and Gaussian Kernel based on Cross-Validation (CV), Generalized Cross-Validation (GCV), and Unbiased Risk (UBR) Methods," *International Journal on Advanced Science Engineering Information Technology*, vol. 11, no. 6, pp. 2400–2406, 2021.
- [22] H. Husain, I. N. Budiantara, and I. Zain, "Mixed estimator of spline truncated, Fourier series, and kernel in biresponse semiparametric regression model," in *IOP Conference Series: Earth and Environmental Science*, vol. 880, Nov. 2021, doi: [10.1088/1755-1315/880/1/012046](https://doi.org/10.1088/1755-1315/880/1/012046).
- [23] A. T. R. Dani, L. Ni'matuzzahroh, V. Ratnasari, and I. N. Budiantara, "Pemodelan Regresi Nonparametrik Spline Truncated pada Data Longitudinal," *Inferensi*, vol. 4, no. 1, p. 47, Mar. 2021, doi: [10.12962/j27213862.v4i1.8737](https://doi.org/10.12962/j27213862.v4i1.8737).
- [24] M. D. Regier and R. D. Parker, "Smoothing using fractional polynomials: An alternative to polynomials and splines in applied research," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 7, no. 4, pp. 275–283, Jul. 2015, doi: [10.1002/wics.1355](https://doi.org/10.1002/wics.1355).
- [25] L. N. Berry and N. E. Helwig, "Cross-Validation, Information Theory, or Maximum Likelihood? A Comparison of Tuning Methods for Penalized Splines," *Stats (Basel)*, vol. 4, no. 3, pp. 701–724, Sep. 2021, doi: [10.3390/stats4030042](https://doi.org/10.3390/stats4030042).
- [26] N. Y. Adrianingsih, I. N. Budiantara, and J. D. T. Purnomo, "Modeling with Mixed Kernel, Spline Truncated and Fourier Series on Human Development Index in East Java," *IOP Conference Series: Materials Science and Engineering*, vol. 1115, no. 1, p. 012024, Mar. 2021, doi: [10.1088/1757-899X/1115/1/012024](https://doi.org/10.1088/1757-899X/1115/1/012024).
- [27] W. Cui and M. Wei, "Strong Consistency of Kernel Regression Estimate," *Open Journal of Statistics*, vol. 03, no. 03, pp. 179–182, 2013, doi: [10.4236/ojs.2013.33020](https://doi.org/10.4236/ojs.2013.33020).