

# Comparison of Kernel and Spline Nonparametric Regression (Case Study: Food Security Index of Jambi Province 2023)

Rosa Salsabila Azarine, Septrina Kiki Arisandi, Fadhilah Fitri\*, dan Yenni Kurniawati

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

\*Corresponding author: [fadhilahfitri@fmipa.unp.ac.id](mailto:fadhilahfitri@fmipa.unp.ac.id)

Submitted : 25 Juni 2025  
Revised : 26 Agustus 2025  
Accepted : 27 Agustus 2025

## ABSTRACT

*Food security is one of the issues that play an important role in national development, especially in regions with varying levels of economic welfare such as Jambi Province. One of the main factors affecting food security is food expenditure, which reflects the economic capacity of households to access food. The complex and non-linear relationship between Food Security Index (FSI) and Food Expenditure requires a flexible modeling approach in the analysis. This study aims to compare the performance of nonparametric regression Kernel and Spline regression methods, namely the Nadaraya-Watson Estimator (NWE) and Local Polynomial Estimator (LPE) for Kernel Regression as well as Smoothing Spline and B-Spline for Spline Regression. The analysis was conducted using secondary data obtained from the Food Security and Vulnerability Map (FSVA) of 2023, with a total of 141 subdistricts in Jambi Province. The response variable is the Food Security Index (FSI), while the predictor variable is Food Expenditure. Model evaluation was conducted using the Mean Squared Error (MSE) and the coefficient of determination ( $R^2$ ). The results showed that the NWE method had the best performance with the smallest MSE value of 24.47690 and the highest  $R^2$  value of 0.3332, meaning that approximately 33.32% of the variation in FSI could be explained by Food Expenditure. The LPE method showed nearly comparable performance, while Smoothing Spline and B-Spline exhibited higher prediction error rates. Therefore, the NWE method can be recommended as an effective nonparametric regression approach for modeling the relationship between food expenditure and food security.*

**Keywords:** Food Security Index, Kernel Regression, Nonparametric Regression, Spline Regression



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

## I. PENDAHULUAN

Analisis regresi adalah salah satu metode dalam statistika yang digunakan untuk mempelajari pola hubungan fungsional antara variabel respons dengan variabel prediktor, dengan tujuan untuk menentukan bentuk estimasi dari kurva regresi (Sihombing & Famalika, 2022). Menurut Salam (2022), analisis regresi dapat dilakukan melalui pendekatan parametrik maupun nonparametrik. Metode parametrik digunakan ketika data mengikuti pola tertentu, sedangkan nonparametrik digunakan ketika data tidak mengikuti pola tertentu. Regresi nonparametrik memiliki keunggulan dalam penyesuaian pola data secara adaptif dan cocok untuk kasus yang memiliki pola hubungan yang kompleks atau tidak terdefinisi secara jelas.

Kernel dan *Spline* merupakan metode regresi nonparametrik yang umum digunakan. Penelitian oleh Wulandari dan Purnama (2020) yang membandingkan regresi nonparametrik kernel NWE (*Nadaraya-Watson Estimator*) dan *B-spline* pada pemodelan rata-rata lama sekolah serta pengeluaran perkapita di Indonesia, menunjukkan bahwa metode *B-spline* lebih sesuai dalam memodelkan kedua variabel tersebut. Sementara itu, penelitian Pratiwi dan Wijaya (2022) yang juga membandingkan estimator spline dan kernel dalam model regresi nonparametrik menggunakan data PDRB di Indonesia, menemukan bahwa estimator kernel memberikan performa terbaik karena menghasilkan nilai GCV terkecil. Perbedaan hasil temuan dari kedua penelitian ini menunjukkan bahwa performa metode regresi nonparametrik dapat bervariasi tergantung pada karakteristik data yang digunakan.

Pada penelitian ini akan dilakukan untuk membandingkan antara kedua metode tersebut, dimana digunakan pendekatan *Nadaraya-Watson Estimator* (NWE) dan *Local Polynomial Estimator* (LPE) dari regresi kernel dengan pendekatan *Smoothing Spline* dan *B-spline* dari regresi *spline*. Perbandingan ini bertujuan untuk evaluasi keakuratan masing-masing metode dalam memodelkan data, dimana data yang digunakan adalah data Indeks Ketahanan Pangan (IKP) Provinsi Jambi tahun 2023 sebagai studi kasus. Pemilihan data ketahanan pangan ini karena pola hubungan antar

variabel cenderung tidak mengikuti distribusi tertentu. Melalui penelitian ini diharapkan diperoleh evaluasi keakuratan masing-masing metode dalam memodelkan data yang memiliki karakteristik serupa dengan data Indeks Ketahanan Pangan.

Melalui situs resmi Kementerian Sosial Republik Indonesia, Widodo (2022) menulis bahwa pengertian ketahanan pangan menurut *Food and Agriculture Organizations (FAO)* adalah kondisi dimana setiap individu, setiap waktu, memiliki akses secara fisik maupun ekonomi terhadap pangan yang memadai, aman dan bergizi, guna mencukupi kebutuhan harian sesuai preferensinya masing-masing. Menurut Rumawas dkk. (2021), ketahanan pangan memiliki peranan strategis dalam mewujudkan pembangunan ekonomi nasional. Indeks Ketahanan Pangan dapat menjadi indikator yang menghasilkan skor komposit untuk menilai ketahanan pangan suatu wilayah. Oleh karena itu, metode yang tepat dalam melakukan analisis terhadap IKP menjadi penting untuk penerapan intervensi program penguatan ketahanan pangan.

Penelitian oleh Rahayu dan Sihombing (2021) pada *Return On Assets (ROA)* Bank Syariah Indonesia tahun 2018 menghasilkan metode *spline* dengan pendekatan *B-Splines* memberikan performa terbaik dengan tingkat akurasi sebesar 74,92%. Meski demikian, belum dapat dipastikan apakah metode tersebut tetap efektif jika diterapkan pada studi kasus yang berbeda. Oleh karena itu, dilakukan penelitian serupa menggunakan data IKP Provinsi Jambi 2023 untuk menguji efektivitas masing-masing metode.

Penelitian ini bertujuan untuk mengidentifikasi metode regresi nonparamterik yang paling optimal antara metode kernel dan *spline* pada data IKP. Temuan dari penelitian ini diharapkan dapat menjadi referensi dalam menentukan metode yang paling sesuai untuk permasalahan serupa, baik di bidang ketahanan pangan atau sektor lain yang memiliki pola data sejenis.

## II. METODE PENELITIAN

### A. Data dan Variabel Penelitian

Analisis pada penelitian ini menggunakan data sekunder yang bersumber dari Peta Ketahanan dan Kerentanan Pangan atau *Food Security and Vulnerability Atlas (FSVA)* yang disusun oleh Badan Pangan Nasional. FSVA adalah peta tematik yang menyajikan gambaran geografis dari hasil analisis berbagai indikator yang menunjukkan tingkat kerentanan terhadap pangan. Objek pengamatan pada penelitian ini merupakan data 141 kecamatan di Provinsi Jambi tahun 2023. Variabel yang dianalisis terdiri atas satu variabel respon ( $Y$ ) yaitu Indeks Ketahanan Pangan dan satu variabel prediktor ( $X$ ) yaitu Pengeluaran Pangan.

### B. Tahapan Analisis Data

Penelitian ini dilakukan dengan menggunakan analisis regresi nonparametrik. Fokus dari penelitian ini adalah membandingkan dua metode dalam regresi nonparametrik yaitu metode kernel dan metode *spline*. Pendekatan *Nadaraya-Watson Estimator (NWE)* dan *Local Polynomial Eestimator (LPE)* digunakan untuk metode kernel, sedangkan pendekatan *smoothing spline* dan *b-spline* untuk metode *spline*. Analisis data dilakukan dengan menggunakan bantuan *software R*. Adapun tahapan analisis dalam penelitian ini dilakukan melalui langkah-langkah sebagai berikut.

#### 1. Eksplorasi data dan Visualisasi Data

Eksplorasi dan visualisasi data perlu dilakukan sebelum data masuk ke tahap analisis. Statistika deskriptif dapat digunakan dalam eksplorasi data untuk mengetahui bagaimana gambaran data yang akan digunakan dalam penelitian. Selain itu, *Scatter Plot* digunakan sebagai alat visualisasi data untuk melihat hubungan antara variabel respon dan variabel prediktor.

#### 2. Regresi Kernel dengan NWE dan LPE

Sebelum dilakukan analisis regresi kernel, perlu menentukan fungsi densitas kernel terlebih dahulu. Fungsi penaksir kernel adalah salah satu metode yang digunakan untuk memperkirakan kurva regresi nonparametrik. Pendekatan ini banyak dipilih karena sifatnya yang fleksibel dan perhitungan matematisnya relatif sederhana. Penaksir kernel adalah bentuk penyempurnaan dari estimator histogram diperkenalkan oleh Rosenblatt pada tahun 1956 dan Parzen tahun 1962, sehingga sering dikenal sebagai penaksir densitas kernel Rosenblatt-Parzen (Hardle, 1992).

Fungsi penaksir kernel  $K$  dengan bandwidth  $h$  secara umum didefinisikan sebagai berikut:

$$K_h = \frac{1}{h} K\left(\frac{x}{h}\right), \quad \text{untuk } -\infty < x < \infty, \quad h > 0 \quad (1)$$

Penaksir densitas kernel untuk fungsi densitas  $f(x)$  didefinisikan sebagai berikut :

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{(x - x_i)}{h}\right) \quad (2)$$

Persamaan (2) di atas memperlihatkan bahwa nilai  $\hat{f}_h(x)$  ditentukan oleh fungsi kernel  $K$  dan parameter  $h$ . Fungsi kernel  $K$  berperan dalam menentukan bentuk bobot kernel, sementara parameter pemulus  $h$  (yang dikenal sebagai bandwidth) mengatur besaran bobot tersebut. Hardle (1992) mengidentifikasi beberapa macam fungsi kernel, di antaranya:

Tabel 1. Fungsi Kernel

Nama	$k$	$h$	Kernel
Uniform	2	0	$\frac{1}{2} 1_{[-1,1]}(x)$
Triangle	2	0	$(1 -  x ) 1_{[-1,1]}(x)$
Normal	2	$\infty$	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$
Epanechnikov	2	1	$\frac{3}{4}(1 - x^2) 1_{[-1,1]}(x)$
Bisquare	2	2	$\frac{15}{6}(1 - 2x^2 + x^4) 1_{[-1,1]}(x)$
Triweight	2	3	$\frac{35}{32}(1 - 3x^2 + 3x^4) 1_{[-1,1]}(x)$
Cauchy	2	0	$[\pi(1 + x^2)]^{-1}$
Picard	2	$\infty$	$\frac{1}{2} \exp(- x )$

(Sumber: Hardle, 1994).

### Pemilihan *Bandwith* Optimal

Masalah kunci dalam estimasi kernel terletak pada penentuan bandwidth yang tepat, bukan pada pemilihan fungsi kernel. Hardle (1992) menjelaskan bahwa bandwidth merepresentasikan bobot kernel yang didefinisikan melalui parameter  $u$ , dimana besaran bobot ini dikontrol oleh nilai  $h$  (lihat Tabel 1). Untuk memperoleh bandwidth optimal, kriteria yang digunakan adalah meminimalkan tingkat kesalahan estimasi (Rahayu & Sihombing, 2021). Metode *Cross Validation* (CV) merupakan salah satu metode yang digunakan dalam pemilihan bandwidth optimal (Sihombing & Famalika, 2022). Berikut adalah rumus dari CV:

$$CV(h) = n^{-1} \sum_{j=1}^n [Y_j - \hat{m}_{h,j}(X_j)]^2 w(X_j) \quad (3)$$

Nilai  $CV(h)$  dihitung sebagai rata-rata kesalahan kuadrat antara nilai aktual  $Y_j$  dan prediksi  $\hat{m}_{h,j}(X_j)$  pada titik  $X_j$  dengan prosedur *leave-one-out*, serta diberi bobot  $w(X_j)$ . *Bandwidth* terbaik dipilih dengan cara meminimalkan nilai  $CV(h)$  (Altman & Leger, 1994).

### Regresi Kernel dengan Metode *Nadaraya-Watson* (NWE)

NWE merupakan salah satu metode regresi nonparametrik yang digunakan untuk mengestimasi fungsi regresi  $m(X)$ . Metode ini dikembangkan secara independen oleh *Nadaraya* dan *Watson* pada tahun 1964. Kedua peneliti tersebut mempublikasikan teknik estimasi  $m(X)$  yang kemudian dikenal sebagai metode *Nadaraya-Watson* (NWE). Rumusan metode ini dapat dinyatakan sebagai berikut (Rahayu & Sihombing, 2021) :

$$\hat{m}(X) = \frac{\sum_{i=1}^n K_h(X_i - x) Y_i}{\sum_{i=1}^n K_h(X_i - x)} \quad (4)$$

Untuk  $X$  tetap, penaksir  $\hat{\theta}$  yang meminimumkan

$$\sum_{i=1}^n (Y_i - \theta)^2 K_h(X_i - x) \quad (5)$$

Memiliki bentuk  $\sum_{i=1}^n a_i - Y_i$  NWE merupakan minimizir persamaan (5) dimana :

$$a_i = \frac{(X_i - x)}{\sum_{i=1}^n K_h(X_i - x)} \quad (6)$$

### Regresi Kernel dengan Metode Polinomial Lokal (LPE)

LPE merupakan metode alternatif dalam regresi nonparametrik, dimana estimasi fungsi regresi  $m(X)$  dilakukan dengan memanfaatkan bentuk polinomial (Fitri et al., 2017). Metode ini menerapkan pendekatan polinomial dalam mengestimasi kurva regresi. Secara matematis, model regresi polinomial lokal dapat dirumuskan sebagai berikut:

$$Y_i = \beta_0 + \beta_1(X_i - x) + \beta_2(X_i - x)^2 + \dots + \beta_p(X_i - x)^p + \varepsilon_i \quad (7)$$

Dengan  $\beta_0, \beta_1, \dots, \beta_p$  adalah koefisien polinomial yang diestimasi secara lokal berdasarkan titik  $x$ . Persamaan 7 memodelkan  $Y_i$  sebagai fungsi polinomial berorde  $p$  dari selisih  $(X_i - x)$ , untuk menangkap pola hubungan lokal antara variabel (Louvain, 1996).

### 3. Regresi *Smoothing Spline*

Fungsi pemulus yang ideal adalah penduga yang mampu merepresentasikan data dengan akurat sekaligus memiliki variansi error yang minimal. Berdasarkan  $n$  observasi, estimasi  $m(X_i)$  dapat diperoleh melalui optimasi fungsi *Penalized Least Square (PLS)*, yaitu:

$$PLS = \sum_{i=1}^n (Y_i - m(X_i))^2 + \lambda \int [m''(x)]^2 dx \quad (8)$$

Fungsi PLS digunakan error prediksi sekaligus mengontrol kelicinan kurva *spline* melalui parameter  $\lambda$ .

### 4. Regresi *B-Splines*

*B-Splines* adalah salah satu pendekatan yang digunakan dalam regresi nonparametrik untuk mengestimasi fungsi regresi secara fleksibel. Metode ini termasuk dalam kategori fungsi polinomial *piecewise* yang memiliki dukungan lokal sesuai derajat polinomial. *B-Splines* ke- $j$  berderajat  $v$ , dikonstruksi berdasarkan urutan *knot*  $t_0, \dots, t_u$  untuk  $j = 1, \dots, v + u$  dinotasikan dengan:

$$B_j(x; v) = \frac{x - t_j}{t_{j+v-1} - t_j} B_j(x; v - 1) + \left( 1 - \frac{x - t_{j+1}}{t_{j+v} - t_{j+1}} B_{j+1}(x; v - 1) \right) \quad (9)$$

Fungsi basis *B-Spline* dibentuk secara rekursif dari fungsi basis orde yang lebih rendah. Variabel  $t_j$  (*knot*) mengontrol letak potongan *spline*, sedangkan  $v$  mengontrol orde polinomialnya.

### Pemilihan *Knot* dalam Regresi *Spline*

Menurut Eubank yang dikutip dalam Yulianti dkk (2020), kinerja suatu estimasi kurva regresi dapat dievaluasi menggunakan beberapa indikator, antara lain rataan kuadrat sisaan (MSE), fungsi *loss*, fungsi risiko dan *Generalized Cross-Validation (CGV)*. Rumus MSE:

$$MSE(\lambda) = \frac{1}{n} (Y - f)^T (Y - f) \quad (10)$$

CGV merupakan modifikasi dari *Cross-Validation (CV)*, rumusnya:

$$CGV(\lambda) = \frac{MSE(\lambda)}{[n^{-1}tr(1 - H)]^2} \quad (11)$$

Cantonie & Hastie dalam penelitian Famalika dkk (2022) menyatakan bahwa faktor pemulus dapat dipilih melalui derajat bebas dalam kasus regresi nonparametrik. Derajat bebas pada *smoothing spline* didefinisikan sebagai berikut:

$$df = tr(s) = \sum_{i=1}^n \frac{1}{1 + \lambda d_i} \quad (12)$$

### 5. Perbandingan Regresi Kernel dan Regresi *Spline*

*Mean Square Error (MSE)* merupakan salah satu ukuran untuk mengetahui tingkat kesalahan suatu estimator. Semakin kecil tingkat kesalahannya semakin baik estimasinya (Sukarsa & Srinadi, 2012). *Mean Squared Error (MSE)* adalah evaluasi kinerja model regresi yang digunakan untuk menghitung rata-rata kuadrat selisih antara nilai yang diprediksi oleh model dan nilai sebenarnya. Persamaan untuk menghitung MSE dapat ditulis sebagai berikut :

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

Model terbaik juga dapat ditentukan dengan melihat nilai koefisien determinasi model. Nilai koefisien determinasi menunjukkan seberapa besar variabel prediktor dapat menjelaskan persentase keragaman dari variabel respon. Berikut ini adalah salah rumus koefisien determinasi:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

Ketika nilai  $R^2$  berada di antara 0 dan 1, hal ini menunjukkan bahwa variabel prediktor mampu menjelaskan variabel respon dalam suatu tingkat tertentu. Semakin mendekati 1, semakin besar proporsi variabilitas variabel respon yang dapat dijelaskan oleh variabel prediktor. Sebaliknya, semakin mendekati 0, semakin kecil proporsi tersebut. Berdasarkan kriteria dari Hair dkk. (2017), nilai  $R^2$  dapat diklasifikasikan menjadi tiga kategori: kuat ( $\geq 0.75$ ), moderat ( $\geq 0.50$ ) dan lemah ( $\geq 0.25$ ).

### III. HASIL DAN PEMBAHASAN

#### A. Eksplorasi dan Visualisasi Data

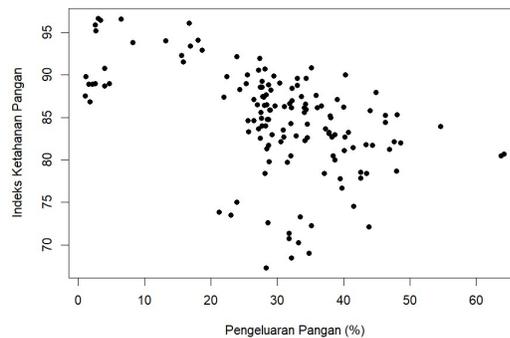
Sebelum data dianalisis, eksplorasi data perlu dilakukan untuk mengetahui gambaran dari data yang digunakan dalam penelitian. Data yang digunakan terdiri atas dua variabel yaitu Indeks Ketahanan Pangan (IKP) dan Pengeluaran Pangan. Statistika deskriptif digunakan untuk eksplorasi data guna mengetahui gambaran data sebelum dilakukan penelitian.

Tabel 2. Statistika Deskriptif

Variabel	Minimum	Maksimum	Median	Mean	Std. Deviasi
IKP (Y)	67.27	96.68	85.59	84.69	6.08
Pengeluaran Pangan (X)	1.09	64.13	30.87	29.89	12.33

Berdasarkan Tabel 2 tersebut diperoleh informasi bahwa rata-rata IKP kecamatan di Provinsi Jambi tahun 2024 adalah 84.69 dan standar deviasi sebesar 6.08. Artinya, dapat disimpulkan bahwa Provinsi Jambi secara umum memiliki ketahanan pangan yang tinggi dan merata. Sebagian besar kecamatan berada dalam kategori ketahanan pangan tangguh, dengan variasi antar kecamatan yang relatif kecil, hal ini menunjukkan konsistensi dan pemerataan kondisi ketahanan pangan di Provinsi Jambi. Sedangkan, pengeluaran pangan kecamatan di Provinsi Jambi tahun 2024 memiliki rata-rata pengeluaran pangan sebesar 29.89%. Artinya, 29.89% dari total pengeluaran rumah tangga kecamatan di Provinsi Jambi digunakan untuk kebutuhan pangan. Hal ini menunjukkan bahwa mayoritas penghasilan atau pengeluaran tidak terlalu terbebani oleh kebutuhan pangan, karena persentase pengeluaran pangan yang tergolong rendah. Ini juga dapat digunakan sebagai indikator bahwa ketahanan pangan di Provinsi Jambi cukup tinggi. Namun, dengan standar deviasi sebesar 12.33 artinya terdapat perbedaan persentase pengeluaran pangan yang tinggi antar kecamatan. Hal ini mengindikasikan bahwa terjadi ketimpangan dalam kondisi akses dan kemampuan konsumsi pangan di Provinsi Jambi.

Selanjutnya, perlu diamati hubungan antara variabel prediktor dan variabel respon melalui visualisasi data untuk memahami pola keterkaitan antara keduanya. Pada gambar 1 tampak bahwa titik-titik data menyebar secara acak yang mengindikasikan tidak adanya pola hubungan antara IKP dan Pengeluaran Pangan. Dengan demikian, pendekatan regresi nonparametrik dianggap lebih tepat digunakan untuk menggambarkan hubungan antara kedua variabel tersebut.



Gambar 1. Scatter Plot IKP dan Pengeluaran Pangan

Untuk mengetahui keeratn hubungan antara IKP dan Pengeluaran Pangan dapat digunakan korelasi pearson diperoleh -0.491379, artinya terdapat hubungan negatif dengan kekuatan sedang antara IKP dengan Pengeluaran Pangan. Artinya, semakin rendah nilai IKP maka semakin tinggi Pengeluaran, begitupun sebaliknya. Hal ini

mengindikasikan bahwa wilayah dengan pengeluaran pangan yang tinggi cenderung memiliki tingkat ketahanan pangan yang lebih rendah.

### B. Regresi Kernel menggunakan Metode NWE dan LPE

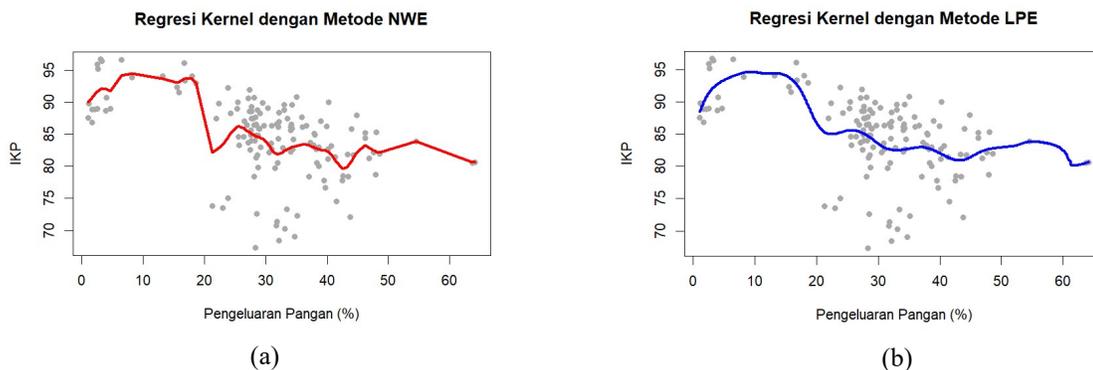
Estimasi densitas pada regresi kernel untuk data IKP dan Pengeluaran pangan dilakukan menggunakan beberapa fungsi kernel, yaitu gaussian, epanechnikov, dan rectangular. Bandwidth optimal digunakan dalam estimasi densitas untuk setiap fungsi kernel tersebut. Selanjutnya, nilai CV yang dihasilkan oleh masing-masing fungsi kernel dengan menggunakan bandwidth optimal dalam estimasi densitas disajikan pada Tabel 3.

**Tabel 3.** Nilai *Bandwidth* (h) dan Fungsi Kernel

Fungsi Kernel	Nilai <i>Bandwidth</i> (h)	Cross Validation
Gaussian	2.403416	26.86349
Epanechnikov	2.683298	27.00450
Rectangular	10.946014	27.51758

Tabel 3 menunjukkan bahwa fungsi gaussian dengan bandwidth optimal sebesar 2.403416 memiliki nilai CV terkecil yaitu 26.86349. Hal ini berarti bahwa bandwidth optimal sebesar 2.403416 dengan fungsi gaussian dapat digunakan dalam estimasi kernel dengan metode NWE dan LPE karena memiliki tingkat kesalahan estimasi paling kecil.

Pada analisis regresi kernel, pemilihan *bandwidth* (h) optimal dan fungsi densitas kernel yang digunakan dapat mempengaruhi hasil estimasi yang didapatkan. Regresi kernel dilakukan menggunakan dua pendekatan dalam regresi kernel menggunakan fungsi densitas kernel yang memiliki bandwidth optimal yaitu fungsi gaussian dengan bandwidth sebesar 2.403416.



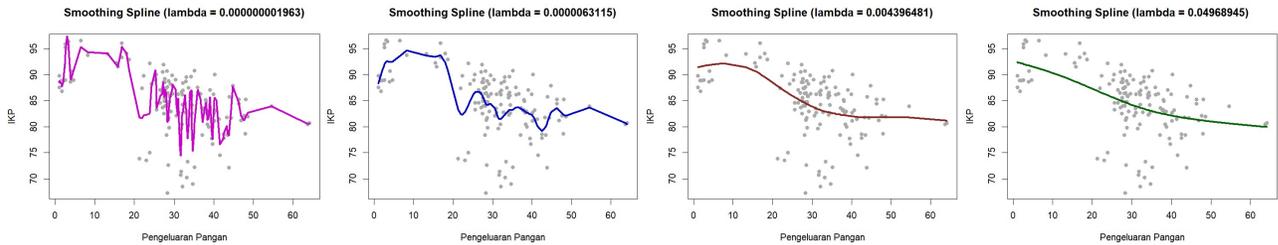
**Gambar 2.** a) Kurva Regresi Kernel dengan Metode NWE, b) Kurva Regresi Kernel dengan Metode LPE

Pada gambar 2 terlihat bahwa kurva regresi kernel menggunakan metode LPE mempunyai bentuk yang lebih halus (*smooth*) dibandingkan dengan kurva metode NWE. Akan tetapi, kurva regresi metode NWE terlihat lebih mengikuti sebaran data dibandingkan dengan LPE. Sehingga berdasarkan kedua kurva regresi tersebut, metode NWE lebih baik dalam memodelkan IKP dan pengeluaran pangan karena mempunyai kurva dengan bentuk yang lebih halus (*smooth*) serta grafik yang cenderung mengikuti plot data.

### C. Regresi Smoothing Spline

Pada regresi *smoothing spline*, nilai dari parameter pemulus ( $\lambda$ ) sangat mempengaruhi kurva regresi yang terbentuk. Semakin besar nilai  $\lambda$  maka bentuk kurva regresi akan semakin mulus. Akan tetapi bentuk kurva yang mulus tidak selalu memberikan hasil yang baik. Kurva regresi yang baik adalah kurva regresi yang halus dan cenderung mendekati bentuk dari plot data dengan bias yang kecil.

Pada analisis dengan *smoothing spline* ini digunakan beberapa nilai  $\lambda$  untuk diujicobakan. Sehingga didapatkan nilai  $\lambda$  optimal sebesar 0.004396481 dengan nilai GCV terkecil yaitu 27.97855. Beberapa kurva regresi dengan nilai  $\lambda$  yang berbeda disajikan pada gambar 4.



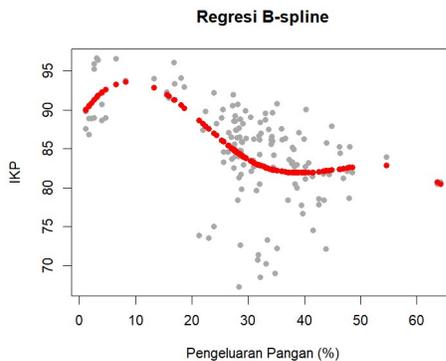
**Gambar 3.** Kurva Regresi *Smoothing Spline* berdasarkan nilai parameter yang diujicoba

Gambar 3 diatas menunjukkan bahwa semakin kecil nilai  $\lambda$  maka kurva regresi yang dihasilkan akan semakin mulus. Berdasarkan kurva di atas maka kurva dengan parameter pemulus sebesar 0.004396481 merupakan model regresi *smoothing spline* terbaik karena memiliki kurva yang halus (*smooth*) dan lebih dekat dengan sebaran data.

#### D. Regresi *B-Splines*

Ada tiga kriteria yang harus diperhatikan dalam membangun model regresi *B-Spline*, yaitu menentukan orde (degree) untuk model, banyak *knot* dan penempatan *knot*. *Knot* merupakan titik-titik yang menandai peralihan pola atau karakteristik data antarinterval. Model *B-Spline* terbaik dipilih berdasarkan penempatan *knot* yang optimal. *Knot* optimal dipilih berdasarkan nilai GCV. Setelah dianalisis diperoleh model *B-Spline* optimal dengan nilai GCV terkecil yaitu model dengan degree 3, jumlah *knot* sebanyak 1 dan 30.87 sebagai *knot* optimalnya.

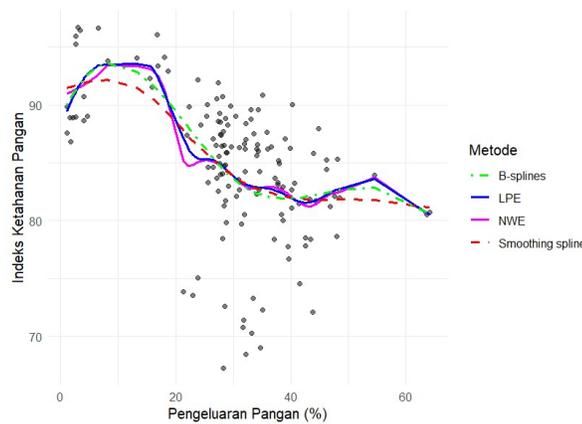
Kurva regresi dari model *B-Spline* terbaik memiliki bentuk yang halus (*smooth*) seperti yang disajikan pada gambar 4 berikut.



**Gambar 4.** Kurva Regresi *B-Splines* dengan *degree* 3 dan jumlah *knot* 1

#### E. Perbandingan Regresi Kernel dan Regresi *Spline*

Berdasarkan analisis sebelumnya, beberapa pendekatan regresi nonparametrik seperti NWE dan LPE untuk regresi kernel serta *smoothing spline* dan *B-spline* untuk regresi *spline* telah digunakan untuk memodelkan pengaruh Persentase Pengeluaran Pangan terhadap Indeks Ketahanan Pangan. Gambar 5 menunjukkan perbandingan kurva regresi dari keempat metode tersebut.



**Gambar 5.** Perbandingan Kurva Regresi NWE, LPE, *Smoothing Spline* dan *B-Splines*

Untuk mengevaluasi kinerja dari keempat metode regresi nonparametrik, yaitu NWE, LPE, *Smoothing Spline*, dan *B-Spline* dalam memodelkan hubungan antara Pengeluaran Pangan sebagai variabel prediktor dan Indeks Ketahanan Pangan sebagai variabel respon, maka digunakan nilai MSE. Nilai MSE yang semakin kecil menunjukkan kinerja model yang semakin baik dalam menjelaskan hubungan antar variabel.

Tabel berikut menunjukkan nilai MSE dari masing-masing metode:

**Tabel 4.** Nilai MSE dan  $R^2$

Metode	MSE	$R^2$
NWE	24.47690	0.3332
LPE	24.83375	0.3235
<i>Smoothing Spline</i>	25.92747	0.2937
<i>B-Spline</i>	25.73724	0.2989

Tabel 4 menunjukkan bahwa metode NWE menghasilkan nilai MSE terkecil, yaitu sebesar 24.47690, dan nilai  $R^2$  tertinggi, yaitu sebesar 0.3332. Hal ini menunjukkan bahwa metode NWE memberikan hasil prediksi yang paling mendekati nilai aktual dari IKP. Selain itu, nilai  $R^2$  sebesar 0.3332 menunjukkan bahwa sebesar 33.32% variasi dalam IKP dapat dijelaskan oleh variabel Pengeluaran Pangan.

Metode LPE menunjukkan performa yang hampir sebanding, dengan nilai MSE sebesar 24.83375 dan  $R^2$  sebesar 0.3235, sedikit lebih rendah dari NWE. Sementara itu, metode *Smoothing Spline* dan *B-Spline* menunjukkan performa prediksi yang lebih rendah, dengan MSE masing-masing sebesar 25.92747 dan 25.73724, serta nilai  $R^2$  yang juga lebih kecil, yaitu 0.2937 dan 0.2989.

#### IV. KESIMPULAN

Berdasarkan hasil evaluasi terhadap metode NWE dan LPE pada regresi kernel serta *Smoothing Spline* dan *B-Spline* pada regresi *spline* dapat disimpulkan bahwa metode *Nadaraya-Watson Estimator* (NWE) adalah metode regresi nonparametrik yang paling efektif dalam memodelkan hubungan antara Pengeluaran Pangan ( $X$ ) dan Indeks Ketahanan Pangan ( $Y$ ) di Provinsi Jambi tahun 2023. Metode NWE menghasilkan nilai MSE terkecil sebesar 24.47690 dengan koefisien determinasi ( $R^2$ ) sebesar 0.3332. Hal ini menunjukkan bahwa metode NWE lebih unggul dalam mengidentifikasi pola hubungan nonlinear antar variabel secara fleksibel dan presisi dibandingkan metode lainnya serta mampu memprediksi nilai Indeks Ketahanan Pangan dengan lebih akurat. Selain itu, nilai  $R^2$  sebesar 0.3332 menunjukkan bahwa sebesar 33.32% variasi dalam IKP dapat dijelaskan oleh variabel Pengeluaran Pangan. Sementara itu, metode LPE memiliki performa yang hampir sama, dengan MSE sebesar 24.83375 dan  $R^2$  sebesar 0.3235. Metode *Smoothing Spline* dan *B-Spline* menunjukkan performa yang relatif lebih rendah, baik dari segi nilai MSE maupun  $R^2$ .

Dengan demikian, NWE dapat dijadikan sebagai metode regresi nonparametrik yang paling direkomendasikan untuk kasus serupa yang melibatkan hubungan nonlinear, khususnya dalam konteks pemodelan ketahanan pangan. Hasil penelitian ini juga dapat menjadi acuan dalam pemilihan metode regresi nonparametrik yang tepat pada sektor pangan maupun bidang lain dengan karakteristik data yang serupa. Penelitian selanjutnya disarankan untuk menggunakan variabel independen yang lebih beragam dan relevan, membandingkan metode NWE dengan pendekatan nonparametrik

lain maupun metode *machine learning* modern, serta menerapkan analisis pada periode waktu dan wilayah berbeda sehingga diperoleh hasil yang lebih umum.

## DAFTAR PUSTAKA

- Altman, N., & Leger, C. (1994). *Bandwith Selection for Kernel Distribution Function Estimation Estimation*. July.
- Fitri, F., Pratama, J. A., Gamayanti, N. F., Andriyana, Y., Program, M., Statistika, M., & Padjadjaran, U. (2017). *Penerapan Regresi Kernel untuk Pemodelan Hasil Tangkap Ikan di Provinsi Jawa Barat*. 2017(Sns Vi).
- FSVA. (2023). *Peta Ketahanan dan Kerentanan Pangan*. <https://fsva.badanpangan.go.id/>
- Hair, J. F., Hult, G. T., Ringle, C., & Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) - Joseph F. Hair, Jr., G. Tomas M. Hult, Christian Ringle, Marko Sarstedt. In *Sage*.
- Hardle, W. (1992). Applied Nonparametric Methods. *Handbook of Econometrics, IV*(26), 1–35.
- Louvain, U. C. De. (1996). *A Study Of Variable Bandwidth Selection For*. 6, 113–127.
- Pratiwi, L. P. S., & Informasi, S. S. (2022). *Perbandingan Estimator Spline Dan Kernel Dalam Model Regresi Nonparametrik Menggunakan Data Studi Sistem Informasi , ITB STIKOM Bali*. 19, 144–152.
- Rahayu, P. I., & Sihombing, P. R. (2021). Penerapan Regresi Nonparametrik Kernel Dan Spline Dalam Memodelkan Return on Asset (RoA) Bank Syariah Di Indonesia. *Jurnal Matematika Murni Dan Terapan Epsilon*, 14(2), 115.
- Rumawas, V. V., Nayoan, H., & Kumayas, N. (2021). Peran Pemerintah Dalam Mewujudkan Ketahanan Pangan di Kabupaten Minahasa Selatan (Studi Dinas Ketahanan Pangan Minahasa Selatan). *Governance*, 1(1), 1–12.
- Salam, N., Sukmawaty, Y., & Halida, A. (2022). Estimasi Model Regresi Nonparametrik dengan Metode B-Spline. *Jurnal Bina Wakya*, 3(7), 6843–6848.
- Sihombing, P. R., & Famalika, A. (2022). Penerapan Analisis Regresi Nonparametrik dengan Pendekatan Regresi Kernel dan Spline. *Jurnal Ekonomi Dan Statistik Indonesia*, 2(2), 172–181.
- Sukarsa, I. K. G., & Srinadi, I. G. A. M. (2012). Estimator Kernel Dalam Model Regresi Nonparametrik. *Jurnal Matematika*, 2(1), 19–30.
- Widodo. (2022). *Kementerian Sosial Republik Indonesia*. Ketahanan Pangan Dan Hak Asasi Manusia. <https://kemensos.go.id/jurnal-dan-artikel/sekretariat-jenderal/ketahanan-pangan-dan-hak-asasi-manusia>
- Wulandary, S., & Purnama, D. I. (2020). Perbandingan Regresi Nonparametrik Kernel Dan B-Splines Pada Pemodelan Rata-Rata Lama Sekolah Dan Pengeluaran Perkapita Di Indonesia. *Jambura Journal of Probability and Statistics*, 1(2), 89–97.