ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 14, Nomor 2, Tahun 2025, Halaman 619 - 630
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

PEMILIHAN *BANDWIDTH* OPTIMAL PADA REGRESI NONPARAMETRIK *KERNEL* MENGGUNAKAN *UNBIASED RISK* (UBR)

Muchni Illahi Efendi[1*], I Nyoman Budiantara[2], Vita Ratnasari[3]
[1,2,3] Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember
***e-mail** : muchniillahiefendi@gmail.com

DOI: 10.14710/ j.gauss.14.2.619-630

Article Info:
Received: 2025-10-21
Accepted: 2025-12-11
Available Online: 2025-12-30

Keywords:
Bandwidth; Economic Growth Rate; Kernel; Nonparametric Regression; Unbiased Risk.

Abstract: Nonparametric regression is used when the relationship between response and predictor variables is not clearly specified. Among its various approaches, the kernel method is widely applied due to its flexibility in capturing complex data patterns. A key component of kernel regression is selecting the optimal bandwidth, as it strongly affects estimation accuracy. One method for determining this bandwidth is the Unbiased Risk (UBR) approach. This study examines the 2024 economic growth rate of Indonesian provinces using predictors such as average years of schooling, life expectancy, and per capita expenditure. These predictors exhibit non-monotonic relationships, meaning that increases in their values do not consistently lead to changes in the response variable. This irregularity indicates a complex structure that does not follow any specific functional form, making kernel based nonparametric regression an appropriate modeling strategy. The study aims to derive the mathematical formulation of the UBR method and evaluate its effectiveness in selecting optimal bandwidths for the dataset. The UBR method produces bandwidths of $h_1 = 0.091$, $h_2 = 0.011$, and $h_3 = 0.003$, with a minimum UBR value of 2.827 and an MSE of 2.198. These results show that kernel regression with UBR-selected bandwidths yields strong predictive performance with relatively low estimation error.

1. PENDAHULUAN

Analisis regresi adalah metode statistika yang memiliki fungsi untuk mengkaji pola keterkaitan variabel dependen dengan variabel independen dengan cara memodelkan hubungan tersebut melalui kurva regresi. Beberapa peneliti telah mengembangkan sejumlah pendekatan analisis regresi yang didasarkan pada bentuk kurvanya meliputi regresi parametrik, nonparametrik, dan semiparametrik (Hardle, 1994). Regresi nonparametrik digunakan apabila bentuk kurva regresinya tidak mengikuti pola fungsi tertentu seperti linier, kuadratik, ataupun bentuk lainnya dan dikenal sebagai pendekatan yang bersifat fleksibel dalam pemodelan data, karena estimasi yang dihasilkan dapat menyesuaikan dengan pola data yang ada tanpa dipengaruhi penilaian subjektif peneliti (Eubank, 1999). Terdapat berbagai estimator yang telah dikembangkan untuk pemodelan regresi nonparametrik meliputi *Spline, Wavelet, Kernel, Fourier Series* serta metode-metode lainnya. Namun, di antara metode tersebut, regresi nonparametrik *kernel* sering digunakan karena bentuk perhitungan matematisnya lebih sederahan dan juga mampu memodelkan data yang memiliki pola yang tidak terstruktur (Chacón and Duong, 2018). Penelitian sebelumnya telah menerapkan regresi nonparametrik *kernel* dalam bidang ekonomi, dan menunjukkan bahwa metode nonparametrik *kernel* memberikan fleksibilitas yang dibutuhkan untuk menangani hubungan yang rumit antarvariabel (Wulandary and Purnama, 2020; Pembargi, Hadijati and Fitriyani, 2023).

Estimator Nadaraya Watson digunakan untuk mengestimasi fungsi regresi nonparametrik *kernel*. Estimator tersebut bergantung pada dua parameter yaitu *bandwidth* dan fungsi *kernel*. Namun, yang terpenting dalam estimasi dengan pendekatan ini adalah pemilihan *bandwidth* optimal (Komang *et al.*, 2012). *Bandwidth* merupakan parameter pemulus yang memiliki peran krusial dalam mengontrol tingkat kemulusan kurva hasil estimasi. Berdasarkan definisi tersebut, untuk memperoleh hasil estimasi yang mampu mencapai keseimbangan antara tingkat kemulusan kurva dan ketepatan model dalam merepresentasikan pola sebenarnya dari data dengan melakukan pemilihan *bandwidth* optimal. Parameter pemulus optimal biasanya ditentukan menggunakan metode *Cross Validation* (CV), *Generalized Cross Validation* (GCV), dan *Unbiased Risk* (UBR). Studi ini berfokus untuk mengkaji penentuan *bandwidth* optimal dengan metode UBR, yaitu metode yang mempertimbangkan estimasi varians dari *error* (Wang, 2011). Metode CV, GCV, dan UBR ini awalnya dikembangkan dalam konteks spline, namun sejumlah penelitian selanjutnya memperluas penerapannya ke berbagai model, dengan melakukan perbandingan antara metode GCV dan UBR pada model regresi nonparametrik *spline* dan GCV memberikan hasil terbaik (Pratiwi, Meina Ayuningsih and Dwijayani, 2021). Perbandingan antara CV dan GCV pada regresi nonparametrik *kernel* maupun *deret fourier* dengan hasil terbaik diperoleh menggunakan metode GCV (Lamusu et al., 2021; Amrullah and Hariksa Amalia, 2022). Selanjutnya, penelitian yang menggunakan metode UBR untuk menentukan osilasi optimal pada model regresi nonparametrik *deret fourier* (Kurnia *et al.*, 2025).

Pada penelitian ini, pendekatan regresi nonparametrik *kernel* dimanfaatkan untuk mengkaji Pertumbuhan ekonomi Indonesia. Pertumbuhan ekonomi dapat didefinisikan sebagai proses peningkatan secara berkelanjutan terhadap total nilai produksi barang serta jasa dalam suatu negara atau wilayah pada periode waktu tertentu. Berdasarkan penelitian-penelitian sebelumnya, variabel-variabel yang mempengaruhi laju pertumbuhan ekonomi telah banyak dimodelkan menggunakan pendekatan parametrik diantaranya, mengkaji dimensi pendidikan dan kesehatan dengan pertumbuhan ekonomi menggunakan model *Autoregressive Distributed Lag* (Sari and Prasetyani, 2025), selanjutnya penelitian juga mengkaji dimensi kesehatan, pendidikan serta ekonomi menggunakan metode data panel regresi (Huda and Indahsari, 2021). Namun, pendekatan parametrik tersebut menetapkan bentuk fungsi regresi tertentu misalnya linear, kuadratik dan lainnya, sehingga kurang fleksibel ketika pola hubungan antarvariabel tidak mengikuti bentuk fungsi tersebut. Oleh karena itu, regresi nonparametrik kernel cocok digunakan karena fleksibel untuk menangkap pola hubungan tanpa perlu menentukan bentuk fungsi terlebih dahulu.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa penelitian terkait pemilihan *bandwidth* optimal menggunakan UBR pada regresi nonparametrik *kernel* masih belum ditemukan. Oleh karena itu, penelitian ini berfokus pada pengkajian formula UBR yang mengacu pada formula Craven dan Wahba (1979) yang awalnya dikembangkan untuk *spline smoothing*, namun diterapkan pada kasus regresi nonparametrik *kernel* untuk pemilihan *bandwidth* optimal. Selanjutnya, penelitian ini dilakukan dengan studi kasus laju pertumbuhan ekonomi Indonesia tahun 2024 menggunakan regresi nonparametrik *kernel* dengan penerapan metode UBR dalam penentuan *bandwidth* optimal.

2. TINJAUAN PUSTAKA

Regresi nonparametrik merupakan pendekatan yang berfungsi untuk mengidentifikasi pola hubungan antara variabel prediktor dan variabel respon ketika bentuk fungsi regresinya belum diketahui serta informasi sebelumnya tidak tersedia dengan lengkap. Ketika diberikan variabel prediktor sebanyak k dengan pasangan data $(x_{1i}, x_{2i}, \dots, x_{ki}, y_i)$ maka hubungan

$(x_{i1}, x_{i2}, \ldots, x_{ik})$ yang merupakan variabel prediktor dengan respon (y_i) dapat dimodelkan melalui regresi nonparametrik sebagai berikut:

$$y_i = f(x_{i1}, x_{i2}, \ldots, x_{ik}) + \varepsilon_i, \qquad i = 1,2,\ldots,n \tag{1}$$

dengan mengasumsikan fungsi regresi f merupakan model regresi *additif*, sehingga persamaan (1) menjadi:

$$\begin{aligned} y_i &= f_1(x_{i1}) + f_2(x_{i2}) + \cdots + f_k(x_{ik}) + \varepsilon_i \\ &= \sum_{l=1}^{k} f(x_{li}) + \varepsilon_i \end{aligned} \tag{2}$$

dimana $f(x_{il})$ merupakan fungsi regresi nonparametrik dan ε_i merupakan *error* acak dimana $\varepsilon_i \sim IIDN(0, \sigma^2)$ (Budiantara, 2019).

Regresi nonparametrik *kernel* merupakan metode yang berperan dalam mengestimasi hubungan variabel respon dengan variabel prediktor yang bersifat nonlinear dengan cara memberikan bobot yang sesuai pada setiap pengamatan. Dalam regresi nonparametrik, fungsi kernel sebagai fungsi pembobot yang menetapkan bobot yang lebih besar bagi pengamatan yang posisinya dekat dengan titik estimasi, dan bobot kecil bagi pengamatan yang berada lebih jauh dari titik tersebut. Bentuk umum fungsi *kernel* K yang bergantung pada *bandwidth h* didefinisikan (Hardle, 1994)

$$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right), i = 1,2,\cdots,n. \tag{3}$$

Serta memenuhi, $K(u) \geq 0$, semua u, $\int_{-\infty}^{\infty} K(u)\ du = 1$, $\int_{-\infty}^{\infty} uK(u)\ du = 0$ dan $\int_{-\infty}^{\infty} u^2 K(u)\ du = \sigma^2 > 0$.

Menurut Hardle (1994) terdapat berbagai jenis fungsi kernel, antara lain Uniform, Epanicnikov, Kuadrat, Gaussian, Triweight, Cosinus. Pada penelitian ini menggunakan fungsi kernel karena hasil estimasinya lebih halus dari kernel lainnya (Hidayat et al, 2020). Persamaan *Kernel* Gaussian dituliskan pada Persamaan (4)

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right); -\infty < x < \infty \tag{4}$$

Nadaraya dan Watson (1964) memperkenalkan suatu estimator yang dikenal sebagai dengan estimator Nadaraya Watson (Hardle, 1994). Estimator ini digunakan untuk melakukan estimasi terhadap fungsi regresi *kernel*. Bentuk umum dari estimator Nadaraya Watson dituliskan pada Persamaan (5) berikut:

$$\hat{f}_h(x) = n^{-1} \sum_{i=1}^{n} V_{hi}(x)\, y_i, i = 1,2,\ldots,n \tag{5}$$

dengan $V_{hi}(x) = \frac{K_h(x - X_i)}{n^{-1}\sum_{j=1}^{n} K_h(x - X_j)}$ dan $K_h(x - X_i) = \frac{1}{h} K\left(\frac{x - X_i}{h}\right)$

$\hat{f}_h(x)$ merupakan estimator kurva regresi *kernel* di titik *x*, dengan $V_{hi}(x)$ merupakan fungsi pembobot, $K_h(x - x_i)$ merupakan fungsi *kernel* yang bergantung pada *bandwidth h*, dan *h* merupakan parameter *bandwidth*. Dalam pemodelan regresi polinomial lokal, fungsi regresi nonparametrik pada Persamaan (2) dapat didekati dengan deret Taylor di sekitar *x* yang memuat polinomial berderajat *p*.

$$f(x_{il}) \approx \frac{1}{0!} f(x) + \frac{1}{1!} f'(x)(X_{il} - x_l) + \cdots + \frac{1}{p!} f^{(p)}(x)(X_{il} - x_l)^p \tag{6}$$

Apabila $f(x_l) = \beta_{0l}$, dan $f'(x_l) = \beta_{1l}$, dan seterusnya hingga $\frac{1}{p!} f^{(p)}(x) = \beta_{pl}$, maka $f(x_{li})$ dapat dituliskan sebagai Persamaan (7).

$$f(x_{il}) \approx \beta_{l0} - \beta_{l1}(X_{il} - x_l) - \cdots - \beta_{pl}(X_{il} - x_l)^p \tag{7}$$

Parameter yang berisi konstanta pada Persamaan (7) dapat dituliskan sebagai berikut:

$$\beta_0 = \beta_{01} + \beta_{02} + \cdots + \beta_{0k} \tag{8}$$

Sehingga persamaan y_i menjadi Persamaan (9)

$$\begin{aligned} y_i &= \sum_{l=1}^{k} f(x_{il}) + \varepsilon_i \\ &= \sum_{l=1}^{k} \left(\beta_0 + \beta_{1l}(X_{il} - x_l) + \cdots + \beta_{pl}(X_{il} - x_l)^p\right) + \varepsilon_i \end{aligned} \tag{9}$$

Persamaan (9) dapat dinyatakan dalam bentuk matriks, yaitu:

$$\mathbf{y}_{(n\times 1)} = \mathbf{X}_{(n\times(pk+1))}\boldsymbol{\beta}_{(pk+1)\times n} + \boldsymbol{\varepsilon}_{(n\times 1)} \tag{10}$$

dengan,

$\mathbf{y} = [y_1 \quad y_2 \quad \cdots \quad y_n]^T$;

$$\mathbf{X} = \begin{bmatrix} 1 & (X_{11} - x_1) & \cdots & (X_{11} - x_2)^p & \cdots & 1 & (X_{1k} - x_k) & \cdots & (X_{1k} - x_k)^p \\ 1 & (X_{21} - x_1) & \cdots & (X_{21} - x_2)^p & \cdots & 1 & (X_{2k} - x_k) & \cdots & (X_{2k} - x_k)^p \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & (X_{n1} - x_1) & \cdots & (X_{n1} - x_2)^p & \cdots & 1 & (X_{nk} - x_k) & \cdots & (X_{nk} - x_k)^p \end{bmatrix};$$

$\boldsymbol{\beta} = [\beta_0 \quad \beta_{11} \quad \cdots \quad \beta_{1k} \quad \beta_{21} \quad \cdots \quad \beta_{2k} \quad \cdots \quad \beta_{pl} \quad \cdots \quad \beta_{pk}]^T$; dan

$\boldsymbol{\varepsilon} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]^T$.

Penyelesaian optimasi dilakukan dengan meminimumkan error digunakan *Weighted Least Square* (WLS), yaitu menambahkan bobot *kernel*, maka diperoleh Persamaan (11)

$$\min_{\beta} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{l=1}^{k} \left(\beta_{1l}(X_{il} - x_l) - \cdots - \beta_{pl}(X_{il} - x_l)^p\right) \right)^2 K_h(x_l - X_{li}) \right\} \tag{11}$$

Dimana, K_h adalah fungsi *kernel*. Bobot *kernel* dapat ditulis dalam bentuk matriks,

$$\mathbf{W} = \begin{bmatrix} \sum_{l=1}^{k} K_{hl}(x_l - X_{1l}) & 0 & \cdots & 0 \\ 0 & \sum_{l=1}^{k} K_{hl}(x_l - X_{2l}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{l=1}^{k} K_{hl}(x_l - X_{nl}) \end{bmatrix} \tag{12}$$

Sehingga diperoleh estimator $\hat{\boldsymbol{\beta}}$ dengan menggunakan WLS, sebagai berikut

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{y} \tag{13}$$

Karena estimator Nadaraya Watson merupakan regresi polinomial lokal saat $p = 0$, didapatkan

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{il}) & \cdots & \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{il}) \\ \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{il}) & \cdots & \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{il}) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{il}) & \cdots & \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{il}) \end{bmatrix}_{k\times k}^{-1} \begin{bmatrix} \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{1l})\, y_i \\ \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{2l}) y_i \\ \vdots \\ \sum_{i=1}^{n}\sum_{l=1}^{k} K_{hl}(x_l - X_{nl})\, y_i \end{bmatrix}_{k\times 1} \tag{14}$$

Estimator Nadaraya Watson untuk model dengan beberapa variabel prediktor dalam bentuk skalar dapat pula ditulis sebagai Persamaan (15).

$$\hat{y}_i = \sum_{l=1}^{k} \hat{f}_h(x_{il}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{k}\left(\frac{K_{hl}(x_l - X_{il})}{n^{-1}\sum_{j=1}^{n} K_{hl}(x_l - X_{jl})}\right) y_i \tag{15}$$

atau jika ditulis dalam bentuk matriks menjadi

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{f}}_h = \mathbf{C}(\mathbf{h})\mathbf{y} \tag{16}$$

dengan $\mathbf{C}(\mathbf{h}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{k}\left(\frac{K_{hl}(x_l - X_{il})}{n^{-1}\sum_{j=1}^{n} K_{hl}(x_l - X_{jl})}\right)$ yang merupakan matriks pemulus untuk menaksir estimator dan $\hat{\boldsymbol{f}}_h = \left[\sum_{l=1}^{k}\hat{f}_h(x_{1l}), \dots, \sum_{l=1}^{k}\hat{f}_h(x_{nl})\right]^T$ yang merupakan vektor hasil estimasi fungsi regresi nonparametrik kernel multivariabel aditif untuk setiap observasi.

Bandwidth yang dilambangkan dengan h merupakan parameter pemulus (*smoothing*) mengatur kehalusan dari kurva hasil estimasi. Secara teoritis bandwidth didefinisikan untuk seluruh nilai $h > 0$ (Wand and Jones, 1995). Namun dalam implementasi praktis, pencarian h tidak dilakukan pada seluruh domain tersebut, melainkan dibatasi pada rentang yang proporsional terhadap skala data, agar pencarian efisien dan menghindari h yang terlalu kecil atau terlalu besar. Pemilihan *bandwidth* optimal sangat penting dilakukan karena akan berpengaruh pada bentuk kurva hasil estimasi. Salah satu metode yang dapat digunakan untuk memperoleh paramater pemulus secara optimal adalah dengan metode *Unbiased Risk* (UBR). Sebelum menentukan nilai optimal dari parameter pemulus dengan menggunakan metode UBR, perlu untuk melakukan estimasi varians *error.*

Craven dan Wahba (1979) memperkenalkan metode UBR untuk memilih parameter pemulus (k) dalam *spline smoothing.* Secara umum, estimasi *spline* dapat dinyatakan sebagai berikut

$$\hat{\mathbf{f}} = \mathbf{A}(\mathbf{K})\mathbf{y} \tag{17}$$

$\mathbf{A}(\mathbf{K})$ merupakan matriks pemulus *spline*. Craven dan Wahba membentuk rumus UBR sebagai berikut :

$$UBR(k) = \frac{1}{n}\|\mathbf{I} - \mathbf{A}(\mathbf{K})\,\mathbf{y}\|^2 - \frac{\hat{\sigma}^2}{n} tr(\mathbf{I} - \mathbf{A}(\mathbf{K})\,)^T(\mathbf{I} - \mathbf{A}(\mathbf{K})\,) + \frac{\hat{\sigma}^2}{n} tr\big(\mathbf{A}^T(\mathbf{K})\big)\big(\mathbf{A}(\mathbf{K})\big) \tag{18}$$

dengan $\mathbf{I}$ merupakan matriks identitas berukuran $n \times n$. Estimasi dari varians *error* $(\hat{\sigma}^2)$ sebagai berikut

$$\hat{\sigma}^2 = \frac{\|\mathbf{I} - \mathbf{A}(\mathbf{K})\,\mathbf{y}\|^2}{trace[\mathbf{I} - \mathbf{A}(\mathbf{K})\,]} \tag{19}$$

Pemilihan parameter pemulus optimal dilakukan dengan cara meminimalkan nilai kriteria UBR.

$$k_{optimal} = \arg\min_{k} UBR(k)$$

Jika dalam suatu model regresi terdapat multikolinearitas, maka akan sulit untuk memperkirakan pengaruh masing-masing prediktor secara akurat. Multikolinearitas merupakan adanya korelasi antar variabel prediktor dalam suatu model regresi (Otse et al., 2025). Untuk menguji adanya multikolinearitas bisa dilakukan dengan *Varian Inflation Factor* (VIF). Jika nilai VIF lebih dari 10 dan nilai *tolerance* kurang dari 0.05, maka hal ini menunjukkan adanya multikolinearitas. Rumus VIF dijabarkan pada persamaan (21),

$$VIF = \frac{1}{1 - R^2} \tag{20}$$

dimana nilai R^2 yang menunjukkan koefisien determinasi dari variabel prediktor terhadap variabel prediktor lainnya.

Mean Squared Error (MSE) merupakan salah satu ukuran untuk menilai ketepatan model dan selalu bernilai non-negatif. Nilai MSE menunjukkan rata-rata selisih kuadrat antara hasil estimasi dengan nilai sebenarnya. Semakin kecil nilai MSE yang diperoleh, maka semakin baik kemampuan model dalam mempresentasikan data sebenarnya. Secara matematis, rumus MSE dapat dituliskan pada Persamaan (21)

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i})^2, i = 1,2, \dots, n \tag{21}$$

dengan $\hat{\mathrm{y}}_\mathrm{i}$ merupakan nilai estimasi dari variabel respon.

3. METODE PENELITIAN

Penelitian ini menggunakan data sekunder yang bersumber dari publikasi Badan Pusat Statistik Indonesia tahun 2024. Unit observasi meliputi 38 Provinsi di Indonesia. Adapun variabel penelitan yang digunakan terdiri atas Laju pertumbuhan ekonomi (y) yang merupakan variabel respon dan variabel prediktor meliputi rata-rata lama sekolah (x_1), angka harapan hidup (x_2), dan pengeluaran per kapita (x_3).

Tahapan penelitian ini dilakukan melalui beberapa langkah berikut:

1) Mengkaji formula *Unbiased Risk* (UBR) dalam pemilihan *bandwidth* optimal pada regresi nonparametrik *kernel*.
2) Melakukan eksplorasi data dengan statistika deskriptif untuk mengetahui karakteristik pada variabel respon dan variabel prediktor.
3) Membuat scatterplot antara variabel respon dengan masing-masing variabel prediktor.
4) Mengujian multikolinearitas pada masing-masing variabel prediktor.
5) Memodelkan data Laju Pertumbuhan Ekonomi di Indonesia dengan menggunakan estimator regresi nonparametrik *kernel* pada Persamaan (16).
6) Menentukan batas nilai *bandiwdth* (h)
7) Membuat *grid bandwidth* (h)
8) Menghitung nilai UBR di setiap nilai *bandwidth* (h)
9) Mendapatkan *bandwidth* optimum dari nilai UBR yang minimum.
10) Menetapkan model terbaik dengan melihat nilai UBR terkecil.
11) Menghitung MSE untuk mendapatkan nilai kebaikan model berdasarkan *bandwidth* optimal.
12) Membuat kesimpulan.

4. HASIL DAN PEMBAHASAN

Tahap awal penelitian ini dilakukan dengan mengkaji formula UBR dalam pemilihan *bandwidth* optimal pada regresi nonparametrik *kernel*. Pengkajian formula ini mengacu pada formula Craven dan Wahba (1979) yang dikembangkan untuk *spline smoothing*. Metode ini mengukur pemilihan paramter penghalus optimal dengan menambahkan informasi mengenai varians *error.* Selain itu, juga untuk mengevaluasi dan membandingkan model menggunakan penilaian risiko yang tak bias. Oleh sebab itu, metode ini menilai kualitas model berdasarkan estimasi risiko yang menyimpang dari nilai sebenarnya. Pembentukan rumus UBR diawali dengan mendefinisikan fungsi *loss.* Fungsi *loss* untuk persamaan fungsi regresi multivariabel yang diasumsikan aditif yang terdapat pada persamaan (2) didefinisikan pada persamaan (22)

$$\begin{aligned} L(h) &= \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{l=1}^{k}\hat{f}_h(x_{il}) - \sum_{l=1}^{k} f_h(x_{il})\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{k}\left(\hat{f}_h(x_{il}) - f_h(x_{il})\right)^2 \end{aligned} \tag{22}$$

Kemudian, fungsi *risk* didefinisikan sebagai ekspektasi dari fungsi *loss* yang disajikan pada persamaan (23).

$$R(h) = E\big(L(h)\big) = E\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{k}\left(\hat{f}_h(x_{il}) - f_h(x_{il})\right)^2\right) \tag{23}$$

Berdasarkan pada persamaan (16), persamaan fungsi *risk* dijabarkan seperti Persamaan (24)

$$\begin{aligned} R(h) &= \frac{1}{n}E\|\mathbf{C}(\mathbf{h})(\mathbf{f}_{\mathrm{h}} + \boldsymbol{\varepsilon}) - \mathbf{f}_{\mathrm{h}}\|^2 \\ &= \frac{1}{n}E\|\mathbf{C}(\mathbf{h})\mathbf{f}_{\mathrm{h}} + \mathbf{C}(\mathbf{h})\boldsymbol{\varepsilon} - \mathbf{f}_{\mathrm{h}}\|^2 \\ &= \frac{1}{n}E\left\|-\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\mathbf{f}_{\mathrm{h}} + \mathbf{C}(\mathbf{h})\boldsymbol{\varepsilon}\right\|^2 \\ &= \frac{1}{n}\left\|\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\mathbf{f}_{\mathrm{h}}\right\|^2 + 0 + \frac{1}{n}E\big(\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\boldsymbol{\varepsilon}\big) \\ &= \frac{1}{n}\left\|\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\mathbf{f}_{\mathrm{h}}\right\|^2 + \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \end{aligned} \tag{24}$$

dimana, $\mathbf{f}_{\mathrm{h}} = \left[\sum_{l=1}^{k} f_h(x_{1l}), \sum_{l=1}^{k} f_h(x_{2l}), \dots, \sum_{l=1}^{k} f_h(x_{nl})\right]^T$ merupakan vektor dari fungsi regresi nonparametrik kernel multivariabel aditif untuk setiap observasi. Kemudian menunjukkan $E\left(\hat{R}(\mathrm{h})\right) = R(h)$ yang dapat didefinisikan pada Persamaan (26).

$$\begin{aligned} \hat{R}(\mathrm{h}) &= \frac{1}{n}\left\|\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\boldsymbol{y}\right\|^2 - \frac{\sigma^2}{n} trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right) \\ &\quad + \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \\ &= \frac{1}{n}\mathbf{y}^{\mathrm{T}}\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right)\mathbf{y} - \frac{\sigma^2}{n} trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right) \\ &\quad + \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \end{aligned} \tag{26}$$

dengan $\hat{R}(\mathrm{h})$ adalah kriteria UBR. Kemudian mencari nilai ekspektasi dari kriteria UBR.

$$\begin{aligned} E\left(\hat{R}(\mathrm{h})\right) = E\Bigg(&\frac{1}{n}\mathbf{y}^{\mathrm{T}}\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right)\mathbf{y} - \frac{\sigma^2}{n} trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right) \\ &+ \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right)\Bigg) \end{aligned} \tag{27}$$

Diberikan suatu Teorema

$$E(\mathbf{y}^{\mathrm{T}}\mathbf{U}\mathbf{y}) = \boldsymbol{\mu}^{\mathrm{T}}\mathbf{U}\boldsymbol{\mu} + trace(\mathbf{U}\boldsymbol{\Sigma})$$

dengan asumsi $\boldsymbol{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, dan berdasarkan teorema di atas didapatkan,

$$\begin{aligned} E\left(\hat{R}(\mathrm{h})\right) &= \frac{1}{n}\Bigg((\mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right)(\mathbf{X}\boldsymbol{\beta}) \\ &\quad + trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\sigma^2\mathbf{I}\right)\Bigg) \\ &\quad - \frac{\sigma^2}{n} trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right) + \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \\ &= \frac{1}{n}\mathbf{f}_{\mathrm{h}}{}^{\mathrm{T}}\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right)\mathbf{f}_{\mathrm{h}} + \frac{\sigma^2}{n} trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right) \\ &\quad - \frac{\sigma^2}{n} trace\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right) + \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \\ &= \frac{1}{n}\mathbf{f}_{\mathrm{h}}{}^{\mathrm{T}}\left(\left(\mathbf{I} - \mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\big(\mathbf{I} - \mathbf{C}(\mathbf{h})\big)\right)\mathbf{f}_{\mathrm{h}} + \frac{\sigma^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \end{aligned} \tag{28}$$

Berdasarkan Persamaan (25) dan (28), terbukti bahwa

$$E\left(\hat{R}(\mathrm{h})\right) = R(h)$$

Sehingga rumus UBR untuk mencari *bandiwidth* optimal seperti persamaan (29)

$$UBR(h) = \frac{1}{n}\mathbf{y}^{\mathrm{T}}\left(\left(\mathbf{I}-\mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\left(\mathbf{I}-\mathbf{C}(\mathbf{h})\right)\right)\mathbf{y} - \frac{\hat{\sigma}^2}{n} trace\left(\left(\mathbf{I}-\mathbf{C}^{\mathrm{T}}(\mathbf{h})\right)\left(\mathbf{I}-\mathbf{C}(\mathbf{h})\right)\right) + \frac{\hat{\sigma}^2}{n} trace\left(\mathbf{C}^{\mathrm{T}}(\mathbf{h})\mathbf{C}(\mathbf{h})\right) \quad (29)$$

dengan, $\mathbf{C}(\mathbf{h})$ merupakan matriks pemulus yang bergantung pada *bandwidth* h, $\mathbf{I}$ merupakan matriks identitas berukuran $n \times n$. Estimasi dari varians *error* ($\hat{\sigma}^2$) sebagai berikut

$$\hat{\sigma}^2 = \frac{\|\mathbf{I}-\mathbf{C}(\mathbf{h})\mathbf{y}\|^2}{trace[\mathbf{I}-\mathbf{C}(\mathbf{h})]} \quad (30)$$

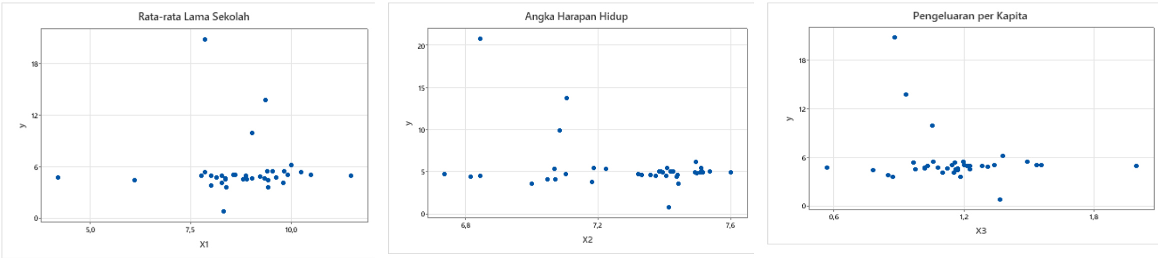
Setelah mendapatkan formula dari UBR, maka langkah selanjutnya adalah menentukan *bandwidth* optimal dengan menerapkan metode UBR pada regresi nonparametrik *kernel* menggunakan data laju pertumbuhan ekonomi setiap provinsi di Indonesia tahun 2024. Sebelum pemodelan dilakukan, karakteristik variabel respon dan variabel prediktor perlu diketahui melalui statistika deskriptif, sebagaimana ditampilkan pada Tabel 1.

Tabel 1. Statistik Deskriptif Variabel Penelitian

Variabel	Satuan	Rata-rata	Standar Deviasi	Minimum	Maksimum
Laju Pertumbuhan Ekonomi (y)	Persen	5.451	3.170	0.770	20.800
Rata-rata Lama Sekolah (x_1)	Tahun	8.841	1.228	4.210	11.490
Angka Harapan Hidup (x_2)	Tahun	7.287	0.230	6.739	7.599
Pengeluaran per Kapita(x_3)	Juta Rupiah	1.159	0.246	0.570	1.995

Berdasarkan Tabel 1, persentase terendah Laju Pertumbuhan Ekonomi (LPE) di Indonesia tahun 2024, yaitu sebesar 0.77%, sedangkan persentase tertingginya sebesar 20,80% dengan rata-rata LPE sebesar 5.45% yang menunjukkan pertumbuhan ekonomi indonesia tumbuh postif. Namun, standar deviasi 3.17% yang artinya data LPE tiap provinsi di Indonesia variasinya cukup tinggi yang menunjukkan tingkat pertumbuhan ekonomia belum merata di seluruh wilayah. Rata-rata Lama Sekolah (RLS) terendah di Indonesia adalah 4.21 tahun, sedangkan RLS tertinggi di Indonesia selama 11.49 tahun. Rata-rata lama sekolah selama 8.84 tahun, yang berarti bahwa mayoritas penduduk di Indonesia telah menempuh pendidikan hingga tingkat SMP. Standar deviasi 1.23 tahun, menunjukkan bahwa perbedaan RLS antar provinsi tidak terlalu besar. Angka Harapan Hidup (AHH) terendah di Indonesia adalah 6.739 atau 67.39 tahun, sedangkan AHH tertinggi di Indonesia selama 7.599 atau 75.99 tahun. Rata-rata AHH di Indonesia, yaitu selama 7.287 atau 72.87 tahun yang menunjukkan bahwa kondisi kesehatan penduduk di Indonesia cukup baik. Standar deviasi 0.230 atau 2.30 tahun yang artinya perbedaan antar provinsi tidak terlalu besar. Pengeluaran per Kapita (PPK) terendah di Indonesia sebesar 0.570 atau 5.70 juta rupiah, sedangkan yang tertinggi sebesar 1.995 atau 19.95 juta. Rata-rata PPK di Indonesia sebesar 1.159 atau 11.59 juta rupiah yang menunjukkan bahwa tingkat kesejahteraan penduduk di Indonesia masih tergolong pada kategori cukup baik. Standar deviasi sebesar 0.246 atau 2.46 juta rupiah menunjukkan bahwa perbedaan pengeluaran per kapita antar provinsi cukup besar.

Selanjutnya, menampilkan *scatterplot* untuk setiap variabel respon dengan variabel prediktor ditunjukkan pada Gambar 1. Berdasarkan Gambar 1, dapat diamati bahwa pola hubungan antara variabel respon, yaitu laju pertumbuhan ekonomi dengan variabel-variabel prediktor tidak mengikuti bentuk fungsi tertentu seperti liniear, kuadratik dan lainnya. Berdasarkan pola yang terlihat pada *scatterplot*, data akan dimodelkan menggunakan pemodelan regresi nonparametrik *kernel*. Karena metode ini bersifat fleksibel dalam pemodelan data dan estimasi yang dihasilkan dapat menyesuaikan dengan pola data yang ada tanpa dipengaruhi penilaian subjektif peneliti. Oleh karena itu, ketiga variabel prediktor akan dimasukkan ke dalam satu pemodelan, maka pemodelan yang dilakukan adalah multivariabel dengan sifat aditif.



Gambar 1. *Scatterplot* Laju Pertumbuhan Ekonomi terhadap RLS, AHH, dan PPK.

Sebelum membentuk model, hal yang perlu dilakukan adalah mendeteksi adanya multikolinearitas antar variabel prediktor. Dalam penelitian ini, *Variance Inflation Factor* (VIF) akan digunakan untuk pendeteksian multikolinearitas. Hasil perhitungan VIF ditampilkan pada Tabel 2.

Tabel 2. Hasil Uji VIF

Variabel	VIF	Keterangan
Rata-rata Lama Sekolah (x_1)	1,90	Tidak terjadi multikolinearitas
Angka Harapan Hidup (x_2)	2,74	Tidak terjadi multikolinearitas
Pengeluaran per Kapita(x_3)	3,44	Tidak terjadi multikolinearitas

Berdasarkan Tabel 2, seluruh nilai VIF berada di bawah batas kritis 10, yang berarti bahwa variabel prediktor tidak mengalami masalah multikolinearitas.

Setelah mendeteksi multikolinearitas antar variabel prediktor, tahapan selanjutnya adalah memodelkan laju pertumbuhan ekonomi di Indonesia dengan menggunakan estimator regresi nonparamaterik pada Persamaan (16). Karena pada penelitian ini hanya menggunakan tiga prediktor, maka bentuk pemodelannya ditunjukkan pada Persamaan (31)

$$\hat{y}_i = \sum_{l=1}^{3} \hat{f}_h(x_{il}) = \frac{1}{n}\sum_{i=1}^{34}\sum_{l=1}^{3}\left(\frac{K_{hl}(x_l - X_{il})}{n^{-1}\sum_{j=1}^{n} K_{hl}\left(x_l - X_{jl}\right)}\right) y_i \tag{31}$$

Sehingga,

$$\hat{y}_i = \frac{1}{n}\left[\sum_{i=1}^{n}\left(\frac{K_{h1}(x_1 - X_{i1})}{n^{-1}\sum_{j=1}^{n} K_{h1}\left(x_1 - X_{j1}\right)}\right) y_i + \sum_{i=1}^{n}\left(\frac{K_{h2}(x_2 - X_{i2})}{n^{-1}\sum_{j=1}^{n} K_{h2}\left(x_2 - X_{j2}\right)}\right) y_i + \sum_{i=1}^{n}\left(\frac{K_{h1}(x_3 - X_{i3})}{n^{-1}\sum_{j=1}^{n} K_{h3}\left(x_3 - X_{j3}\right)}\right) y_i\right] \tag{32}$$

Penelitian ini menggunakan fungsi *kernel* Gaussian. Sehingga persamaan (32) akan menjadi Persamaan (33)

$$\hat{y}_i = \frac{1}{n}\left[\sum_{i=1}^{n}\left(\frac{\frac{1}{h_1}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{x_1 - X_{i1}}{h_1}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{h_1}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{x_1 - X_{j1}}{h_1}\right)^2}\right)y_i + \sum_{i=1}^{n}\left(\frac{\frac{1}{h_2}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{x_2 - X_{i2}}{h_2}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{h_2}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{x_2 - X_{j2}}{h_2}\right)^2}\right)y_i + \sum_{i=1}^{n}\left(\frac{\frac{1}{h_3}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{x_3 - X_{i3}}{h_3}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{h_3}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{x_3 - X_{j3}}{h_3}\right)^2}\right)y_i\right] \tag{33}$$

Persamaan (33) disederhanakan menjadi Persamaan (34)

$$\hat{y}_i = \frac{1}{n}\left[\sum_{i=1}^{n}\left(\frac{\frac{1}{h_1}\exp\left(\frac{x_1 - X_{i1}}{h_1}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{h_1}\exp\left(\frac{x_1 - X_{j1}}{h_1}\right)^2}\right)y_i + \sum_{i=1}^{n}\left(\frac{\frac{1}{h_2}\exp\left(\frac{x_2 - X_{i2}}{h_2}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{h_2}\exp\left(\frac{x_2 - X_{j2}}{h_2}\right)^2}\right)y_i + \sum_{i=1}^{n}\left(\frac{\frac{1}{h_3}\exp\left(\frac{x_3 - X_{i3}}{h_3}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{h_3}\exp\left(\frac{x_3 - X_{j3}}{h_3}\right)^2}\right)y_i\right] \tag{34}$$

Setelah didapatkan model untuk estimator regresi nonparametrik *kernel* dengan 3 variabel prediktor, selanjutnya adalah pemilihan *bandwidth* yang optimal. Pemilihan *bandwidth* sebagai pembobot untuk setiap variabel prediktor secara teoritis nilainya leboh besar dari 0 memiliki batas tak berhingga. Pemilihan *bandwidth* optimal dilakukan dengan menggunakan metode UBR yang sudah di peroleh pada Persamaan (29). *Bandwidth* yang optimal ketika nilai UBR minimum. Pada penelitian ini, *bandwidth* yang dicobakan bervariasi dengan batasan mulai dari $h > 0$ hingga rentang masing-masing prediktor. Selanjutnya, pembentukan grid *bandwidth* dari batasan *bandwidth* tersebut dibagi menjadi 400 nilai terpilih sehingga menghasilkan kumpulan kandidat *bandwidth.* Selanjutnya, menghitung nilai UBR pada setiap kombinasi *bandwidth*. Nilai UBR diurutkan dan diambil *bandwidth* yang menghasilkan nilai UBR terendah. Nilai *bandwidth* yang optimal pada penelitian ini untuk setiap variabel disajikan pada Tabel 3.

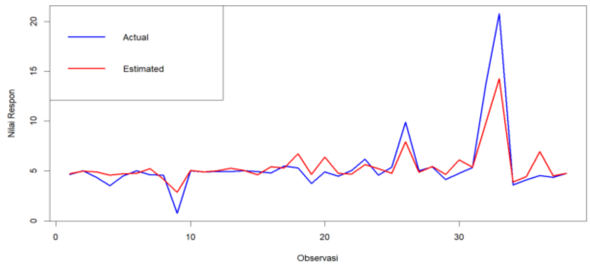
Tabel 3. Estimasi Nilai *Bandwidth*

Bandwidth	Estimasi
h_1	0.091
h_2	0.011
h_3	0.003

Nilai UBR yang dihasilkan pada nilai *bandwidth* yang ada pada Tabel 3. adalah 2.827. Apabila estimasi *bandwidth* disubstitusi ke Persamaan (34), maka akan diperoleh persamaan seperti Persaman (35).

$$
\begin{aligned}
\hat{y}_i = \frac{1}{34}\Bigg[&\sum_{i=1}^{34}\left(\frac{\frac{1}{0.091}\exp\left(\frac{x_1 - X_{i1}}{0.091}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{0.091}\exp\left(\frac{x_1 - X_{j1}}{0.091}\right)^2}\right)y_i \\
&+\sum_{i=1}^{34}\left(\frac{\frac{1}{0.011}\exp\left(\frac{x_2 - X_{i2}}{0.011}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{0.011}\exp\left(\frac{x_2 - X_{j2}}{0.011}\right)^2}\right)y_i \\
&+\sum_{i=1}^{34}\left(\frac{\frac{1}{0.003}\exp\left(\frac{x_3 - X_{i3}}{0.003}\right)^2}{n^{-1}\sum_{j=1}^{n}\frac{1}{0.003}\exp\left(\frac{x_3 - X_{j3}}{0.003}\right)^2}\right)y_i\Bigg]
\end{aligned} \tag{35}
$$

Selanjutnya, menghitung nilai *Mean Squared Error (MSE)* untuk mendapatkan nilai kebaikan model berdasarkan *bandwidth* optimal. Berdasarkan hasil perhitungan, diperoleh nilai MSE sebesar 2.198, yang berarti model memiliki tingkat kesalahan prediksi yang relatif rendah dan mampu mempresentasikan pola hubungan antara variabel respon dan variabel prediktor cukup baik. Selanjutnya, ditampilkan plot nilai aktual (y) dan nilai estimasi $(\hat{y})$ untuk melihat kesesuaian hasil estimasi dengan data aktual yang disajikan pada Gambar 2.



Gambar 2. Visualisasi Nilai Aktual (y) dan Nilai Estimasi $(\hat{y})$

Berdasarkan Gambar 2, hasil estimasi yang dihasilkan oleh model regresi nonparametrik *kernel* Nadaraya Watson cenderung mengikuti pola data aktual, dimana garis berwarna merah menunjukkan nilai estimasi $(\hat{y})$ dan garis berwaana biru menunjukkan nilai aktual (y), yaitu data Laju Pertumbuhan Ekonomi di Indonesia tahun 2024. Namun, ada sedikit perbedaan antara nilai aktual dengan nilai estimasi pada beberapa titik, dimana nilai aktual mengalami fluktuasi yang cukup tinggi sementara estimasi model belum sepenuhnya mengikuti perubahan tersebut. Secara keseluruhan model regresi nonparametrik dengan *bandwidth* optimal yang diperoleh dengan menggunakan metode UBR memiliki kemampuan prediksi yang baik dengan tingkat kesalahan estimasi yang relatif rendah.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa metode UBR yang diperoleh dapat digunakan dalam pemilihan *bandwidth* optimal pada regresi nonparametrik *kernel* pada data Laju Pertumbuhan Ekonomi di Indonesia tahun 2024. Hasil *bandwidth* optimal yang diperoleh dari nilai minimum UBR untuk masing-masing variable prediktor, yaitu $h_1 = 0.091, h_2 = 0{,}011$, dan $h_3 = 0{,}003$ dengan UBR minimum 2.827 dan MSE sebesar 2.198 yang berarti bahwa model regresi nonparametrik dengan *bandwidth* optimal yang diperoleh dengan menggunakan metode UBR memiliki kemampuan prediksi yang baik dengan tingkat kesalahan estimasi yang relatif rendah.

DAFTAR PUSTAKA

Amrullah, Mn. and Hariksa Amalia, S. (2022) ‘Comparison Of Generalized Cross Validation (Gcv) Methods With Cross Validation (Cv) To Determine Optimal Knots In Fourier Series Nonparametric Regression (Case Study: Poverty Rate in North Sumatra Province)’, | *Jurnal Litbang Edusaintech*, 3(1), pp. 2022–2023. Available at: https://doi.org/10.51402/jle.v3i1.5.

Budiantara, I.N. (2019) *Regresi nonparametrik spline truncated*. Surabaya: ITS Press.

Chacon, J. and Duong, T. (2018) *Multivariate kernel smoothing and its applications*. Monographs on Statistics and Applied Probability, vol. 160.

Eubank, R.L. (1999) *Nonparametric regression and spline smoothing*. 2nd edn. New York: Marcel Dekker.

Hardle, W. (1994) *Applied nonparametric regression*. Berlin: Humboldt Universitat zu Berlin.

Hidayat, R., et al. (2019) ‘Kernel-Spline estimation of additive nonparametric regression model’, *IOP Conference Series: Materials Science and Engineering*, 546, 052028.

Huda, N. and Indahsari, K. (2021) ‘Pengaruh Rata-Rata Lama Sekolah, Angka Harapan Hidup Dan Pengeluaran Perkapita Terhadap Pertumbuhan Ekonomi Provinsi Jawa Timur Tahun 2014-2018’, *Buletin Ekonomika Pembangunan*, 2(1), pp. 55–66. Available at: https://doi.org/10.21107/bep.v2i1.13849.

Komang, I. *et al.* (2012) ‘Estimator Kernel Dalam Model Regresi Nonparametrik’, *Jurnal Matematika*, 2(1), pp. 20–30.

Kurnia, D.A. *et al.* (2025) ‘Metode unbiased risk diterapkan untuk pemilihan osilasi optimal dalam regresi nonparametrik menggunakan pendekatan deret fourier’, 14, pp. 139–148. Available at: https://doi.org/10.14710/j.gauss.14.1.139-148.

Lamusu, F., Machmud, T. and Resmawan, R. (2021) ‘Estimator Nadaraya-Watson dengan Pendekatan Cross Validation dan Generalized Cross Validation untuk Mengestimasi Produksi Jagung’, *Indonesian Journal of Applied Statistics*, 3(2), p. 85. Available at: https://doi.org/10.13057/ijas.v3i2.42125.

Pratiwi, L.P.S., Meina Ayuningsih, N.P. and Dwijayani, N.M. (2021) ‘Perbandingan Gcv Dan Ubr Dalam Regresi Nonparametrik Multivariabel’, *Jurnal Matematika*, 11(1), p. 64. Available at: https://doi.org/10.24843/jmat.2021.v11.i01.p137.

Pembargi, J.A., Hadijati, M. and Fitriyani, N. (2023) ‘Kernel Nonparametric Regression for Forecasting Local Original Income’, *Jurnal Varian*, 6(2), pp. 119–126. Available at: https://doi.org/10.30812/varian.v6i2.2585.

Sari, V.K. and Prasetyani, D. (2025) ‘Does Human Capital Matter for Indonesia’s Economic Growth?’, *Jurnal Ekonomi Pembangunan*, 22(2), pp. 291–302. Available at: https://doi.org/10.29259/jep.v22i2.23186.

Wand, M.P. & Jones, M.C. (1995) *Kernel Smoothing.* London: Chapman & Hall.

Wang, Y. (2011) *Smoothing Splines Methods and Applications*. California: CRC Press.

Wulandary, S. And 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), pp. 89–97. Available at: https://doi.org/10.34312/jjps.v1i2.7501.