

Multiresponse Nonparametric Regression Model with Mixed Estimator of Truncated Spline and Kernel for Poverty Indicators Analysis in Nusa Tenggara

Model Regresi Nonparametrik Multirespon dengan Estimator Campuran Spline *Truncated* dan Kernel untuk Analisis Indikator Kemiskinan di Nusa Tenggara

Aris Aswadi^{1*}, I Nyoman Budiantara², Ismaini Zain³

^{1,2,3} Department of Statistics, Institut Teknologi Sepuluh Nopember, Indonesia

Email: ¹ariesaswadie@gmail.com, ²i_nyoman_b@statistika.its.ac.id, ³ismaini_z@statistika.its.ac.id

*Corresponding author

Abstract

Regression analysis is a statistical method used to describe the causal relationship between response variables and predictor variables. Regression analysis can be classified into parametric regression, nonparametric regression, and semiparametric regression, depending on whether the regression curve is fully known, unknown, or partially known. This study aims to obtain a multiresponse nonparametric regression model with a mixed spline truncated and kernel estimator. The model obtained is applied to Poverty Indicators in 32 districts/cities in Nusa Tenggara. The response variables include the Percentage of Poor People, Poverty Depth Index, and Poverty Severity Index, while the predictors are the Human Development Index, Open Unemployment Rate, and GRDP per capita. The estimation method used in this research is Weighted Least Square (WLS). The result shows that the Human Development Index predictor variable can be approximated by a truncated spline function, while the Open Unemployment Rate and Gross Regional Domestic Product (GRDP) per capita predictor variables can be approximated by a kernel function. The multiresponse nonparametric regression model with a mixture of truncated spline and kernel estimators can be used to model Poverty Indicators in Nusa Tenggara. Results show that the Human Development Index aligns with the spline function, while the other predictors align with the kernel function. The best model is a model with one knot point and two bandwidths where the model produces an R^2 value of 89.86% based on the smallest GCV value.

Keywords: Regression Analysis, Nonparametric Regression, Poverty, Truncated Spline, Kernel, Weighted Least Square

Abstrak

Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menggambarkan hubungan sebab akibat antara variabel respon dengan variabel prediktor. Analisis regresi dapat diklasifikasikan ke dalam regresi parametrik, nonparametrik, dan semiparametrik, tergantung pada



bentuk kurva regresinya diketahui sepenuhnya, tidak diketahui, atau sebagian diketahui. Penelitian ini bertujuan untuk memperoleh model regresi nonparametrik multirespon dengan estimator gabungan spline *truncated* dan kernel. Model yang diperoleh diterapkan pada indikator kemiskinan pada 32 kabupaten/kota di wilayah Nusa Tenggara. Variabel respon yang digunakan meliputi Persentase Penduduk Miskin, Indeks Kedalaman Kemiskinan, dan Indeks Keparahan Kemiskinan, sedangkan variabel prediktor meliputi Indeks Pembangunan Manusia, Tingkat Pengangguran Terbuka, dan Produk Domestik Regional Bruto (PDRB) per kapita. Metode estimasi yang digunakan dalam penelitian ini adalah *Weighted Least Square* (WLS). Hasil penelitian menunjukkan bahwa model regresi nonparametrik multirespon dengan estimator gabungan spline *truncated* dan kernel dapat diestimasi menggunakan metode WLS. Model ini dapat digunakan untuk memodelkan indikator kemiskinan di wilayah Nusa Tenggara. Hasil menunjukkan bahwa variabel prediktor Indeks Pembangunan Manusia dapat didekatkan dengan fungsi spline *truncated*, sedangkan variabel prediktor Tingkat Pengangguran Terbuka dan Produk Domestik Regional Bruto (PDRB) per kapita dapat didekatkan dengan fungsi kernel. Model terbaik merupakan model dengan satu titik knot dan dua *bandwidth* dimana model tersebut menghasilkan nilai R^2 sebesar 89,86% berdasarkan nilai GCV terkecil.

Kata kunci: Analisis Regresi, Regresi Nonparametrik, Kemiskinan, Spline *Truncated*, Kernel, *Weighted Least Square*

1. PENDAHULUAN

Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menggambarkan hubungan sebab akibat antara variabel respon dengan variabel prediktor [13]. Secara umum, analisis regresi dapat diklasifikasikan menjadi tiga berdasarkan diketahui atau tidaknya bentuk hubungan antara variabel respon dan variabel prediktor, yaitu regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Regresi parametrik harus memenuhi asumsi dan bersifat kaku, karena digunakan ketika bentuk kurva regresi harus diketahui. Sebaliknya, regresi nonparametrik digunakan ketika bentuk kurva regresinya diasumsikan tidak diketahui. Adapun regresi semiparametrik merupakan kombinasi antara regresi parametrik dan nonparametrik, dan digunakan ketika hanya sebagian bentuk kurva regresinya diketahui dan sebagian lagi tidak diketahui [7]. Bentuk/pola yang dimaksud adalah seperti berbentuk linear, polinomial, eksponensial, atau lainnya. Dalam situasi di mana bentuk/pola hubungan variabel tidak diketahui bentuk kurva regresinya maka pendekatan regresi nonparametrik lebih direkomendasikan. Beberapa model regresi nonparametrik yang sering mendapat perhatian dari para peneliti antara lain kernel, deret fourier, polinomial lokal, dan spline [9].

Diantara pendekatan yang digunakan pada regresi nonparametrik yang telah disebutkan sebelumnya, spline dikenal sebagai salah satu model/pendekatan yang unggul karena mempunyai interpretasi statistik dan interpretasi visual yang sangat baik [9]. Spline *truncated* merupakan potongan polinomial yang bersifat tersegmen atau terpotong. Kelebihan dari model spline *truncated* adalah sifat fleksibilitas yang dikarenakan spline *truncated* memiliki titik knot yang menunjukkan terjadinya perubahan pola data [14]. Selain itu, spline efektif dalam menangani karakter data/fungsi yang bersifat mulus (*smooth*). Spline juga memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu [5]. Dalam membangun model regresi nonparametrik spline *truncated*, terdapat beberapa hal penting yang perlu diperhatikan. Pertama, penentuan orde model spline *truncated*, yang umumnya ditentukan dengan pola umum berdasarkan pola yang terlihat pada data. Kedua, menentukan jumlah knot, yang bergantung pada seberapa sering pola data berubah pada sub-sub interval. Setiap kali terjadi perubahan pada bentuk/pola kurva, dapat diasumsikan adanya sebuah titik knot pada lokasi perubahan tersebut, sehingga jumlah knot dapat ditentukan berdasarkan banyaknya perubahan yang terjadi. Ketiga, penetapan lokasi knot, yang dilakukan dengan mengidentifikasi area pada kurva yang menunjukkan

adanya perubahan pola kurva. Estimator lain yang sering digunakan dalam regresi nonparametrik yaitu fungsi kernel. Ada beberapa jenis fungsi kernel, antara lain kernel *uniform*, *triangle*, *epanechnikov*, *gaussian*, *kuartik* dan *cosinus* [11]. Estimator kernel adalah estimator linear sama seperti estimator lainnya, tetapi pendekatan dengan Kernel lebih khusus dengan adanya *bandwidth* [9]. Estimator kernel mempunyai kemampuan yang sangat baik dalam memodelkan data yang tidak memiliki pola atau bentuk atau bahkan seolah terlihat acak [11].

Beberapa penelitian yang telah dilakukan sebelumnya menggunakan estimator tunggal untuk semua prediktornya. Namun, pada kenyataannya setiap variabel prediktor memiliki pola data atau kurva regresi terhadap variabel respon yang berbeda satu dengan yang lainnya. Jumlah titik knot dan lebar *bandwidth* juga seringkali diasumsikan sama untuk setiap variabel prediktornya, pada kenyataannya setiap variabel prediktor memiliki pola yang berbeda, sehingga jumlah titik knot maupun lebar *bandwidth* dari setiap variabel prediktor tersebut akan berbeda satu dengan yang lainnya. Pada penelitian ini digunakan model campuran antara dua estimator regresi nonparametrik yang berbeda yaitu estimator spline *truncated* dan kernel. Kemudian pada penelitian ini akan dikembangkan model yang tidak hanya terbatas pada respon tunggal. Pada kenyataannya terdapat kasus yang mana melibatkan lebih dari satu variabel respon untuk dianalisis secara bersamaan sehingga diperlukan analisis regresi multirespon dengan estimator campuran. Sebelumnya penelitian untuk melakukan pemodelan regresi nonparametrik dengan estimator campuran spline *truncated* dan kernel masih terbatas pada birespon oleh [8], dimana untuk mengestimasi modelnya digunakan metode *Weighted Least Square* (WLS). Oleh karena itu, penelitian ini diperoleh model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel. Untuk mengestimasi model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel digunakan metode *Weighted Least Square* (WLS) dua tahap. Metode WLS digunakan untuk mengakomodasi adanya korelasi antar variabel respon. Kemudian, model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel dapat digunakan untuk memodelkan Indikator Kemiskinan di Nusa Tenggara Barat dan Nusa Tenggara Timur.

2. TINJAUAN PUSTAKA

2.1 Regresi Parameter

Regresi parametrik merupakan metode regresi yang digunakan untuk memodelkan pola hubungan antara variabel respon dengan variabel prediktor dengan menggunakan persamaan matematika, di mana bentuk kurva atau fungsi regresi diketahui. *Scatterplot* adalah salah satu alat bantu visual yang digunakan untuk mengetahui apakah bentuk pola kurva regresi diketahui atau tidak. Melalui *scatterplot*, dapat diperoleh informasi apakah suatu fungsi dari kurva regresi mengikuti pola linier, kuadratik, polinomial atau tidak mengikuti suatu pola tertentu. Model regresi parametrik mengasumsikan bahwa bentuk/pola kurva regresinya diketahui [9]. Jika kurva regresinya diketahui yaitu kurva regresi linear berganda, maka persamaan yang terbentuk [2] adalah persamaan (2.1) sebagai berikut.

$$\begin{aligned} y_i &= g(x_1, x_2, \dots, x_p) \\ &= \beta_0 + \sum_{j=1}^p \beta_j x_{ji} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i \end{aligned} \quad 2.1$$

di mana y_i variabel respon pada pengamatan ke- i , x_{ji} adalah variabel prediktor ke- j dan observasi ke- i , β_0 intersep, β_j merupakan koefisien regresi ke- j , dan ε_i merupakan *error* random yang diasumsikan berdistribusi normal, identik, serta independen dengan rata-rata 0 dan varians σ^2 .

2.2 Regresi Nonparametrik

Metode regresi nonparametrik merupakan metode regresi yang dapat mengatasi kelemahan dari metode regresi parameter di mana bentuk/pola dari kurva regresi g yang harus diketahui [9]. Oleh

karena itu, metode regresi nonparametrik digunakan ketika kurva regresi antara variabel respon dan variabel prediktor tidak diketahui pola atau bentuknya. Pendekatan regresi nonparameterik yang menyatakan hubungan antara variabel respon dengan p variabel prediktor yang melibatkan n observasi dapat disajikan dalam bentuk persamaan (2.2) berikut.

$$y_i = \sum_{j=1}^p g(x_{ji}) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad 2.2$$

dimana y_i adalah variabel respon pada observasi ke- i , g merupakan kurva regresi nonparametrik yang menghubungkan prediktor dengan variabel respon, ε_i adalah error random yang independen dan berdistribusi normal dengan rata-rata nol dan varians σ^2 pada observasi ke- i , x_{ji} adalah variabel prediktor ke- j dan observasi ke- i , dan n adalah banyaknya observasi [15].

2.3 Spline *Truncated* dalam Regresi Nonparametrik

Regresi nonparametrik dengan estimator spline *truncated* merupakan pendekatan regresi nonparametrik yang sesuai untuk digunakan dalam menangani data yang memiliki pola atau perilaku berubah-ubah pada sub-sub interval tertentu. Model regresi nonparametrik spline *truncated* dapat disajikan seperti persamaan (2.2) dimana $g(x_{ji})$ merupakan kurva regresi yang didekati dengan fungsi spline *truncated* yang memiliki orde/tingkatan M . Spline *truncated* sangat bergantung pada titik knot, baik itu jumlah maupun letak titik knot. Titik knot merupakan titik perubahan pola pada data [9]. Adapun spline *truncated* derajat M dengan titik-titik knot $K_1, K_2, \dots, K_u$ ($K_1 \leq K_2 \leq \dots \leq K_u$) dapat dituliskan seperti persamaan (2.2) dimana $g(x_i)$ diberikan oleh persamaan (2.3) berikut [3].

$$g(x_i) = \sum_{m=0}^M \alpha_m x_i^m + \sum_{s=1}^u \beta_s (x_i - K_s)_+^m \quad 2.3$$

dimana fungsi *truncated* (potongan) diberikan oleh persamaan berikut.

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

2.4 Spline *Truncated* dalam Regresi Nonparametrik Multirespon

Dalam praktiknya, peneliti kerap dihadapkan pada situasi di mana dua atau lebih variabel respon yang diamati perlu dianalisis secara bersamaan berdasarkan sejumlah variabel prediktor. Untuk kondisi seperti ini, model regresi nonparameterik multirespon merupakan metode yang paling tepat untuk memodelkan fungsi dari variabel-variabel tersebut. Diberikan data multirespon berpasangan $(y_{1i}, y_{2i}, \dots, y_{qi}, x_{1i}, x_{2i}, \dots, x_{pi})$ dengan q variabel respon dan p variabel prediktor. Model regresi nonparametrik multirespon dapat dinyatakan seperti persamaan (2.4) berikut.

$$\begin{aligned} y_{1i} &= g_{11}(x_{1i}) + g_{12}(x_{2i}) + \dots + g_{1p}(x_{pi}) + \varepsilon_{1i} \\ y_{2i} &= g_{21}(x_{1i}) + g_{22}(x_{2i}) + \dots + g_{2p}(x_{pi}) + \varepsilon_{2i} \\ &\vdots \\ y_{qi} &= g_{q1}(x_{1i}) + g_{q2}(x_{2i}) + \dots + g_{qp}(x_{pi}) + \varepsilon_{qi} \end{aligned} \quad 2.4$$

dimana $g(x)$ diasumsikan bentuk regresinya tidak diketahui sehingga didekati dengan fungsi spline *truncated* pada persamaan (2.5) berikut.

$$g_{hj}(x_{ji}) = \sum_{j=1}^p \left[\sum_{m=0}^M \alpha_{hjm} x_{ji}^m + \sum_{s=1}^u \beta_{hjs} (x_{ji} - K_{hjs})_+^m \right] \quad 2.5$$

dimana

$$(x_{ji} - K_{hjs})_+^m = \begin{cases} (x_{ji} - K_{hjs})^m & , x_{ji} \geq K_{hjs} \\ 0 & , x_{ji} < K_{hjs} \end{cases}$$

Model regresi nonparametrik termasuk dalam model linear dalam parameter sehingga variabel respon $\tilde{y}$ juga berdistribusi normal multivariat $\tilde{y} \sim N(\mathbf{g}, \mathbf{M})$. Matriks varians-kovarians $\mathbf{M}$ [16] dinyatakan sebagai berikut.

$$\mathbf{M} = \begin{bmatrix} \sigma_1^2 & \cdots & 0 & \rho\sigma_1\sigma_2 & \cdots & 0 & \vdots & \rho\sigma_1\sigma_r & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_1^2 & 0 & \cdots & \rho\sigma_1\sigma_2 & \cdots & 0 & \cdots & \rho\sigma_1\sigma_r \\ \rho\sigma_2\sigma_1 & \cdots & 0 & \sigma_2^2 & \cdots & 0 & \vdots & \rho\sigma_2\sigma_r & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \rho\sigma_2\sigma_1 & 0 & \cdots & \sigma_2^2 & \cdots & 0 & \cdots & \rho\sigma_2\sigma_r \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ \rho\sigma_r\sigma_1 & \cdots & 0 & \rho\sigma_r\sigma_2 & \cdots & 0 & \cdots & \sigma_r^2 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \rho\sigma_r\sigma_1 & 0 & \cdots & \rho\sigma_r\sigma_2 & \cdots & 0 & \cdots & \sigma_r^2 \end{bmatrix}_{(r+1) \times (r+1)}$$

Matriks $\mathbf{M}$ dapat disederhanakan menurut persamaan (2.6) berikut.

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \cdots & \mathbf{M}_{1r} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \cdots & \mathbf{M}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{r1} & \mathbf{M}_{r2} & \cdots & \mathbf{M}_{rr} \end{bmatrix} \quad 2.6$$

2.5 Kernel dalam Regresi Nonparametrik

Pendekatan kernel sangat bergantung pada *bandwidth* α , yang berguna untuk mengendalikan *smoothness* dari kurva estimasi. Pemilihan *bandwidth* terbaik dilihat dari nilai *bandwidth* yang optimum yang bersesuaian dengan fungsi kernel yang digunakan [12]. Estimator kernel merupakan estimator regresi nonparametrik yang digunakan untuk menangani kasus di mana data tidak memiliki pola/bentuk atau seolah terlihat acak. Kurva regresi yang dihipotesiskan fungsi kernel, dapat disajikan dalam bentuk persamaan (2.7) berikut.

$$\hat{f}_\alpha(t) = n^{-1} \sum_{i=1}^n \left[\frac{K_\alpha(t-t_i)}{n^{-1} \sum_{j=1}^n K_\alpha(t-t_j)} \right] y_i = n^{-1} \sum_{i=1}^n W_{\alpha i}(t) y_i \quad 2.7$$

dimana

$$W_{\alpha i}(t) = \frac{K_\alpha(t-t_i)}{n^{-1} \sum_{j=1}^n K_\alpha(t-t_j)}; \quad K_\alpha(t-t_i) = \frac{1}{\alpha} K\left(\frac{t-t_i}{\alpha}\right)$$

2.6 Kernel Multirespon dalam Regresi Nonparametrik

Jika diterapkan pada kasus multirespon model yang diperoleh dapat didefinisikan sebagai persamaan (2.8) berikut.

$$y_{hi} = \sum_{j=1}^q f_{hj}(t_{ji}) + \varepsilon_{hi}; h = 1, 2, \dots, r; i = 1, 2, \dots, n \quad 2.8$$

dengan y_{hi} adalah respon ke- h pada observasi ke- i , f_{hj} adalah kurva regresi yang menghubungkan prediktor ke- j dengan respon ke- h , t_{ji} adalah prediktor ke- j pada pengamatan ke- i , ε_{hi} adalah *error*

random pada observasi ke- i dan respon ke- h , n adalah banyaknya observasi dan q adalah banyaknya prediktor. Regresi nonparametrik Kernel dapat dibentuk dengan menghampiri kurva regresi $f_{hj}(t_{ji})$ dengan estimator Kernel. Fungsi $f_{hj}(t_{ji})$ untuk $h = 1, 2, \dots, r$ dapat didekati dengan deret Taylor dengan t_j di sekitar t_{j0} [4]. Kemudian $f_{hj}(t_{ji})$ jika dijabarkan serupa dengan persamaan (2.7).

2.7 Generalized Cross Validation (GCV)

Pemilihan banyak serta letak titik knot dan parameter *bandwidth* yang tepat berfungsi untuk mendapatkan model yang optimal serta mengatur kemulusan kurva regresi yang diestimasi. Parameter penghalus pada model regresi nonparametrik spline *truncated* adalah titik knot K . Sedangkan pada model regresi nonparametrik dengan estimator Kernel adalah *bandwidth* (α). Oleh karena itu, akan dipilih titik knot K maupun *bandwidth* yang optimal sehingga akan diperoleh estimator yang paling sesuai untuk memodelkan data.

Salah satu metode yang dapat digunakan untuk menentukan parameter penghalus yang optimal adalah *Generalized Cross Validation* (GCV) [6]. Keunggulan dari metode GCV jika dibandingkan dengan metode lain seperti *Cross Validation* (CV) dan metode *Unbiased Risk* (UBR) ataupun *Generalized Maximum Likelihood* (GML), terletak pada karakteristik dari metode GCV yang memiliki sifat optimal asimtotik [15]. Fungsi GCV untuk model regresi nonparametrik multirespon dapat didefinisikan sebagai persamaan (2.9) berikut.

$$GCV(K, \alpha) = \frac{N^{-1} \|y_i - \hat{\mu}_{\alpha, \gamma}(x_i, t_i)\|^2}{(N^{-1} \text{trace}[\mathbf{I} - \mathbf{A}(\mathbf{K}, \alpha) - \mathbf{D}(\alpha)])^2}; N = nr \quad 2.9$$

2.8 Kemiskinan dan Indikator Kemiskinan

Untuk mengukur kemiskinan di Indonesia, Badan Pusat Statistik (BPS) menggunakan konsep kemampuan memenuhi kebutuhan dasar (*Basic Needs Approach*). Menurut pendekatan ini, kemiskinan dianggap sebagai ketidakmampuan secara ekonomi untuk memenuhi kebutuhan pokok makanan dan non-makanan, yang pengukurannya didasarkan pada besarnya pengeluaran. Masyarakat miskin adalah mereka yang rata-rata pengeluaran per kapita bulannya berada di bawah garis kemiskinan. Garis Kemiskinan (GK) terdiri atas dua komponen utama, yaitu Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non Makanan (GKNM). Menurut [1] garis kemiskinan didefinisikan sebagai hasil penjumlahan dari garis kemiskinan makanan dan garis kemiskinan non-makanan. Garis kemiskinan makanan merupakan tingkat pengeluaran pangan terendah atau setara dengan 2.100 kilokalori per orang per hari. Sementara itu, garis kemiskinan non-makanan kini mencakup kebutuhan minimum akan perumahan, sandang, pendidikan dan kesehatan. Berdasarkan Garis Kemiskinan tersebut terdapat beberapa dimensi untuk mengukur kemiskinan di Indonesia yaitu jumlah penduduk miskin, persentase penduduk miskin, Indeks Kedalaman Kemiskinan dan Indeks Keparahan Kemiskinan.

Menurut [10], setidaknya terdapat dua faktor utama yang perlu diperhatikan dalam upaya penanggulangan kemiskinan, yaitu pertumbuhan ekonomi dan distribusi pendapatan. Selain itu, dijelaskan bahwa pengeluaran pemerintah memiliki peran penting dalam mendorong pertumbuhan ekonomi, yang pada akhirnya dapat meningkatkan pendapatan masyarakat dan menurunkan jumlah penduduk yang berada di bawah garis kemiskinan. Oleh karena itu, diperlukan langkah-langkah nyata dalam bentuk program dan kebijakan dari pemerintah daerah yang mampu membuka peluang lebih luas bagi para pelaku ekonomi lokal, khususnya di wilayah yang membutuhkan perhatian khusus melalui kebijakan yang mendukung dan memudahkan aktivitas ekonomi. Dampaknya diharapkan mampu meningkatkan kapasitas ekonomi masyarakat miskin. Dengan demikian, strategi pembangunan ke depan perlu difokuskan pada penciptaan serta pengembangan potensi ekonomi daerah guna mendorong pertumbuhan ekonomi yang inklusif serta mengurangi kesenjangan distribusi pendapatan.

3. METODE PENELITIAN

3.1 Data Penelitian

Data yang digunakan pada penelitian ini merupakan data sekunder yang bersumber dari berbagai publikasi BPS Nusa Tenggara Barat dan Nusa Tenggara Timur yang terdiri atas 32 kab/kota di Nusa Tenggara. Variabel komposit yang digunakan dalam penelitian ini adalah Indikator Kemiskinan yang terdiri atas Persentase Penduduk Miskin (P_0), Indeks Kedalaman Kemiskinan (P_1), dan Indeks Keparahan Kemiskinan (P_2). Adapun variabel prediktor yang diduga berpengaruh terhadap variabel respon adalah Indeks Pembangunan Manusia, Tingkat Pengangguran Terbuka, dan laju PDRB per kapita.

3.2 Langkah-langkah Analisis

Bagian ini merupakan langkah-langkah aplikasi menerapkan model estimator campuran spline *truncated* dan kernel dalam regresi nonparametrik multirespon pada Persentase Penduduk Miskin, Indeks Kedalaman Kemiskinan, dan Indeks Keparahan Kemiskinan di kabupaten/kota di Nusa Tenggara. Untuk menjawab tujuan tersebut, dilakukan langkah-langkah sebagai berikut.

- Melakukan eksplorasi terhadap variabel respon dan variabel prediktor untuk memperoleh statistik deskriptif dari setiap variabel penelitian.
- Melakukan uji korelasi antar variabel respon dengan uji Pearson.
- Membuat scatterplot secara parsial antara variabel respon dengan masing-masing variabel prediktor.
- Menentukan variabel prediktor yang merupakan komponen Spline *truncated* dan komponen Kernel berdasarkan hasil scatterplot.
- Memodelkan Persentase Penduduk Miskin, Indeks Kedalaman Kemiskinan, dan Indeks Keparahan Kemiskinan menggunakan estimator campuran Spline *truncated* dan Kernel multirespon.
- Memilih titik knot dan parameter *bandwidth* optimal dalam estimator campuran spline *truncated* dan kernel multirespon dengan menggunakan metode GCV.
- Membuat berbagai macam kombinasi untuk masing-masing variabel prediktor yang merupakan komponen spline *truncated* dan kernel.
- Memilih model terbaik dengan GCV paling kecil.
- Menentukan nilai $\hat{\beta}$ dan MSE dari model terbaik untuk campuran spline *truncated* dan kernel dalam regresi nonparametrik multirespon.
- Membuat kesimpulan.

4. HASIL DAN PEMBAHASAN

4.1 Analisis Deskripsi

Berikut merupakan statistika deskriptif dari masing-masing variabel penelitian yang disajikan pada Tabel 4.1.

Tabel 4. 1. Statistika Deskriptif Variabel Penelitian

Variabel	Jumlah	Min	Maks	Mean	Median	Standar Deviasi
y_1	32	8,61	31,78	18,54	16,23	6,90
y_2	32	1,07	8,16	3,50	3,20	1,96
y_3	32	0,18	3,03	0,98	0,77	0,73
x_1	32	58,89	81,15	67,96	66,66	5,22
x_2	32	1,40	5,69	3,08	2,79	1,04

x_3	32	-10,37	5,77	3,30	3,66	2,62
-------	----	--------	------	------	------	------

4.2 Uji Korelasi

Analisis regresi multirespon diperlukan antar variabel respon memiliki korelasi tinggi. Oleh karena itu, selanjutnya dilakukan perhitungan nilai korelasi Pearson antar variabel respon, untuk melihat korelasi dari variabel respon. Adapun nilai korelasi antar variabel respon disajikan pada Tabel 4.2 berikut.

Tabel 4. 2. Nilai Korelasi Pearson

Variabel	y_1	y_2	y_3
y_1	1	0,9303 (0,000)	0,8647 (0,000)
y_2	0,9303 (0,000)	1	0,9811 (0,000)
y_3	0,8647 (0,000)	0,9811 (0,000)	1

Keterangan tabel: angka yang berada di dalam kurung menunjukkan nilai p -value

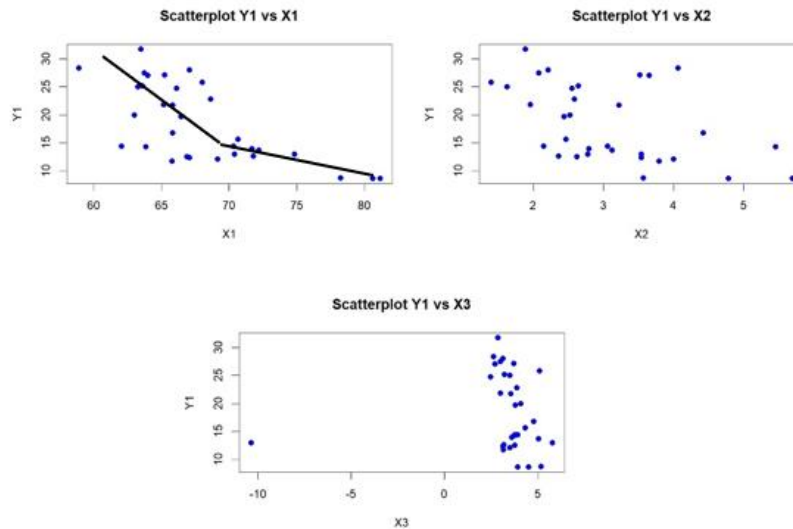
Berdasarkan Tabel 4.2 di atas, terlihat bahwa korelasi antar variabel respon termasuk sangat tinggi. Hal ini ditunjukkan oleh nilai koefisien korelasi yang tinggi, seperti antara y_1 dan y_2 sebesar 0,9303, antara y_1 dan y_3 sebesar 0,8647, serta antara y_2 dan y_3 sebesar 0,9811 yang juga menunjukkan korelasi kuat. Berdasarkan nilai p -value yang diperoleh untuk masing-masing pasangan variabel respon, seluruhnya lebih kecil dari taraf signifikansi 5%, sehingga Keputusan pengujian adalah menolak H_0 . Dengan demikian, dapat disimpulkan bahwa terdapat hubungan atau korelasi yang signifikan antara ketiga variabel respon tersebut.

Berdasarkan nilai korelasi yang telah dihitung, dilakukan penyusunan matriks varian-kovarians sebagai bobot untuk estimasi. Metode estimasi yang digunakan pada penelitian ini adalah *Weighted Least Square* (WLS), sehingga membutuhkan suatu bobot untuk melakukan estimasi. Bobot yang digunakan adalah matriks varian-kovarians (**M**). Matriks varian-kovarians ini digunakan untuk mengakomodasi korelasi antar variabel respon. Matriks varian-kovarians yang diperoleh dapat disajikan pada persamaan (4.1) berikut.

$$\mathbf{M} = \begin{bmatrix} \begin{vmatrix} 47,57 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 47,57 \end{vmatrix} & \begin{vmatrix} 12,58 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 12,58 \end{vmatrix} & \begin{vmatrix} 4,34 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 4,34 \end{vmatrix} \\ \begin{vmatrix} 12,58 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 12,58 \end{vmatrix} & \begin{vmatrix} 3,84 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 3,84 \end{vmatrix} & \begin{vmatrix} 1,40 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1,40 \end{vmatrix} \\ \begin{vmatrix} 4,34 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 4,34 \end{vmatrix} & \begin{vmatrix} 1,40 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1,40 \end{vmatrix} & \begin{vmatrix} 0,53 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0,53 \end{vmatrix} \end{bmatrix}_{96 \times 96} \quad 4.1$$

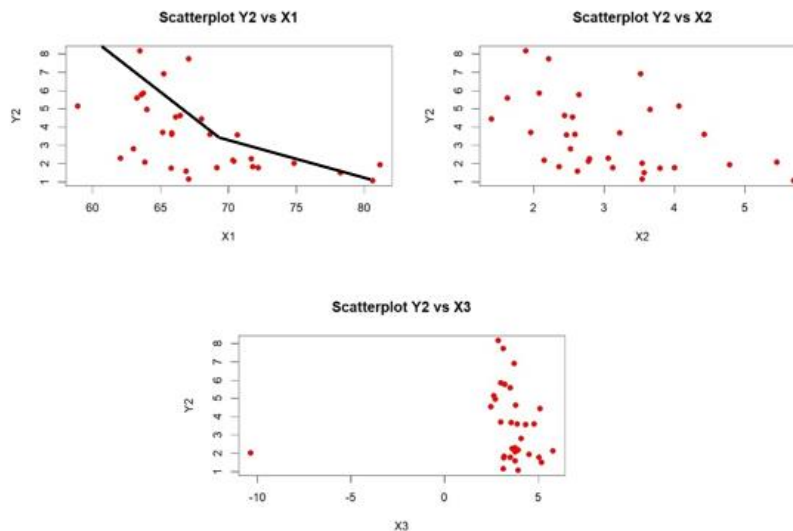
4.3 Model Regresi Nonparametrik Multirespon dengan Estimator Campuran Spline *Truncated* dan Kernel

Untuk melihat pola hubungan antara ketiga variabel respon dengan masing-masing variabel prediktor dapat dilihat dari grafik *scatter plot*. Adapun *scatter plot* variabel penelitian dapat dilihat seperti Gambar 4.1 berikut.



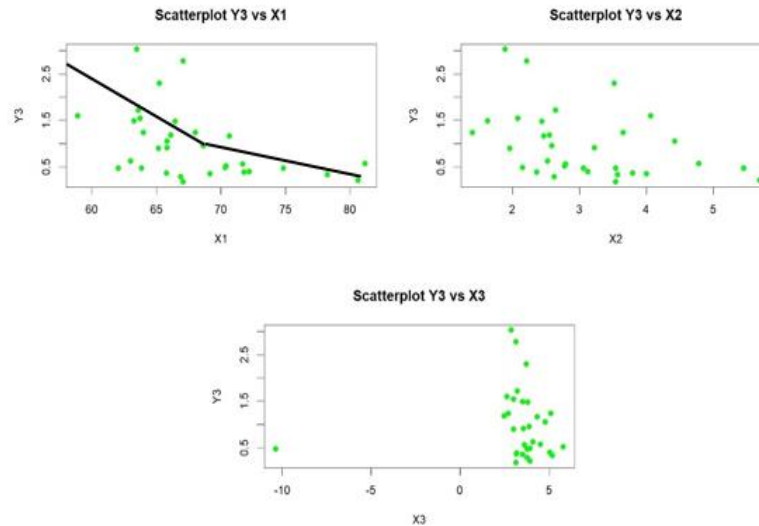
Gambar 4. 1. *Scatter Plot* antar Variabel Prediktor dengan Variabel Respon y_1

Pada Gambar 4.1 terlihat bahwa bentuk pola hubungan antara variabel respon Persentase Penduduk Miskin (y_1) dan variabel prediktor IPM (x_1), TPT (x_2), dan laju PDRB per kapita (x_3).



Gambar 4. 2. *Scatter Plot* antar Variabel Prediktor dengan Variabel Respon y_2

Pada Gambar 4.2 terlihat bahwa bentuk pola hubungan antara variabel respon Indeks Kedalaman Kemiskinan (y_2) dan variabel prediktor IPM (x_1), TPT (x_2), dan laju PDRB per kapita (x_3).



Gambar 4.3. Scatter Plot antar Variabel Prediktor dengan Variabel Respon y_3

Pada Gambar 4.3 terlihat bahwa bentuk pola hubungan antara variabel respon Indeks Keparahan Kemiskinan (y_3) dan variabel prediktor IPM (x_1), TPT (x_2), dan laju PDRB per kapita (x_3). Berdasarkan Gambar 4.1, Gambar 4.2, dan Gambar 4.3 terlihat bahwa bentuk pola hubungan antara ketiga variabel respon (y_1, y_2, y_3) dan variabel prediktor IPM (x_1) tidak mengikuti pola tertentu dan cenderung mengalami perubahan pada sub interval tertentu. Pada kabupaten/kota dengan IPM di bawah 70, kenaikan IPM berpengaruh pada penurunan semua variabel respon dengan pergerakan yang tajam. Sebaliknya pada interval IPM di atas 70 pergerakannya melambat. Pada data yang memiliki pola demikian identifikasi awal bahwa data dapat dimodelkan dengan pendekatan fungsi spline *truncated*. Selanjutnya, sama halnya dengan bentuk pola hubungan antara ketiga variabel respon (y_1, y_2, y_3) dengan variabel prediktor TPT (x_2) yang cenderung memiliki pola yang acak. Tidak terlihat adanya pola data, sehingga data yang memiliki pola acak seperti ini identifikasi awal bahwa data dapat dimodelkan dengan pendekatan fungsi kernel. Kemudian, bentuk pola hubungan antara ketiga variabel respon (y_1, y_2, y_3) dengan variabel prediktor laju PDRB/kapita (x_3), juga memiliki pola yang acak, tidak terdapat adanya pola data. Oleh karena itu, identifikasi awal bahwa data dapat dimodelkan dengan pendekatan fungsi kernel. Jika data dimodelkan dengan pendekatan fungsi spline *truncated*, maka variabel prediktor akan diwakili dengan variabel x . Sedangkan jika data dapat dimodelkan dengan pendekatan fungsi kernel maka variabel prediktor akan diwakili dengan variabel t . Berikut merupakan model populasi yang disajikan berdasarkan persamaan (4.2) untuk memodelkan Indikator Kemiskinan di provinsi Nusa Tenggara Barat dan Nusa Tenggara Timur, berdasarkan asumsi awal yang telah dipaparkan sebelumnya.

$$\begin{aligned}
 y_{hi} &= \mu_{hi}(x_{1i}, x_{2i}, \dots, x_{pi}, t_{1i}, t_{2i}, \dots, t_{qi}) + \varepsilon_{hi}, h = 1, 2, 3; i = 1, 2, \dots, n \\
 &= g_h(x_i) + \sum_{k=1}^q f_{hk}(t_{ki}) + \varepsilon_{hi} \\
 &= \alpha_{h0} + \alpha_{hi}x_i + \sum_{s=1}^u \beta_{hs}(x_i - K_{hs})_+ + f_{h1}(t_{1i}) + f_{h2}(t_{2i}) + \varepsilon_{hi}
 \end{aligned} \tag{4.2}$$

Kemudian jika dijabarkan berdasarkan persamaan (4.2), dapat disajikan seperti persamaan (4.3), persamaan (4.4), dan persamaan (4.5).

$$\begin{aligned}
 y_{1i} &= g_1(x_1) + f_{11}(t_1) + f_{12}(t_2) + \varepsilon_{1i} \\
 &= \alpha_{10} + \alpha_{1i}x_{1i} + \sum_{s=1}^u \beta_{1s}(x_{1i} - K_{1s})_+ + f_{11}(t_{1i}) + f_{12}(t_{2i}) + \varepsilon_{1i}
 \end{aligned} \tag{4.3}$$

$$y_{2i} = g_2(x_i) + f_{21}(t_1) + f_{22}(t_2) + \varepsilon_{2i}$$

$$= \alpha_{20} + \alpha_{2i}x_{1i} + \sum_{s=1}^u \beta_{2s}(x_{1i} - K_{2s})_+ + f_{21}(t_{1i}) + f_{22}(t_{2i}) + \varepsilon_{2i} \quad 4.4$$

$$y_{3i} = g_3(x_i) + f_{31}(t_1) + f_{32}(t_2) + \varepsilon_{3i}$$

$$= \alpha_{30} + \alpha_{3i}x_{1i} + \sum_{s=1}^u \beta_{3s}(x_{1i} - K_{3s})_+ + f_{31}(t_{1i}) + f_{32}(t_{2i}) + \varepsilon_{3i} \quad 4.5$$

Selanjutnya akan dilakukan pemodelan untuk memodelkan Indikator Kemiskinan dengan satu variabel prediktor yang didekatkan dengan fungsi spline *truncated* yaitu variabel IPM (x_1), dan dua variabel prediktor yang didekatkan dengan fungsi kernel, yaitu variabel TPT (t_1) dan variabel laju PDRB/kapita (t_2). Kemudian berdasarkan model pada persamaan (4.3), persamaan (4.4), dan persamaan (4.5), diberikan dua kombinasi jumlah titik knot (jumlah titik knot = 1 dan 2). Dengan demikian, model tersebut akan diperoleh kombinasi dua model, yaitu model yang terdiri dari satu titik knot dan model yang terdiri dari dua titik knot.

4.4 Model dengan Satu Titik Knot

Pemodelan untuk satu variabel prediktor yang didekatkan dengan fungsi spline *truncated* yaitu variabel prediktor x_1 , dan dua variabel prediktor yang didekatkan dengan fungsi kernel, yaitu variabel prediktor t_1 dan t_2 . Berikut merupakan model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* linier dan kernel dengan menggunakan satu titik knot dan banyaknya *bandwidth* yang digunakan adalah dua *bandwidth* disajikan pada persamaan (4.6).

$$y_{hi} = \alpha_{h0} + \alpha_{hi}x_{1i} + \beta_{h1}(x_{1i} - K_{h1})_+ + f_{h1}(t_{1i}) + f_{h2}(t_{2i}) + \varepsilon_{hi} \quad 4.6$$

Berikut merupakan letak titik knot untuk satu titik knot, *bandwidth* optimum, dan GCV yang diperoleh, dapat disajikan pada Tabel 4.3 berikut.

Tabel 4.3. Letak Titik Knot, Bandwidth, GCV untuk Model Satu Titik Knot

Parameter	Titik Knot dan <i>Bandwidth</i>						
K_{111}	59,11	59,11	...	59,11	...	80,93	80,93
K_{211}	59,61	59,61	...	59,61	...	80,93	80,93
K_{311}	58,62	58,62	...	58,62	...	80,92	80,92
α_{11}	0,48	0,95	...	1,91	...	3,34	3,81
α_{12}	1,79	3,59	...	7,17	...	12,55	14,35
α_{21}	0,39	0,78	...	1,56	...	2,73	3,12
α_{22}	1,47	2,93	...	5,87	...	10,27	11,74
α_{31}	0,33	0,66	...	1,32	...	2,31	2,64
α_{32}	1,24	2,48	...	4,97	...	8,69	9,93
GCV	11,55	11,18	...	10,97*)	...	11,02	11,02

*) GCV Minimum

Dengan K_{111} merupakan titik knot untuk variabel y_1 , K_{211} merupakan titik knot untuk variabel y_2 , K_{311} merupakan titik knot untuk variabel y_3 . α_{11} adalah *bandwidth* untuk variabel t_1 pada y_1 , α_{12} adalah *bandwidth* untuk variabel t_2 pada y_1 , α_{21} adalah *bandwidth* untuk variabel t_1 pada y_2 , α_{22} adalah *bandwidth* untuk variabel t_2 pada y_2 . Sedangkan α_{31} adalah *bandwidth* untuk variabel t_1 pada y_3 , dan α_{32} adalah *bandwidth* untuk variabel t_2 pada y_3 .

4.5 Model dengan Dua Titik Knot

Berikut merupakan model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* linier dan kernel dengan menggunakan dua titik knot dan banyaknya *bandwidth* yang digunakan adalah dua *bandwidth* disajikan pada persamaan (4.7).

$$y_{hi} = \alpha_{h0} + \alpha_{hi} x_{1i} + \beta_{h1} (x_{1i} - K_{h1})_+ + \beta_{h2} (x_{1i} - K_{h2})_+ + f_{h1}(t_{1i}) + f_{h2}(t_{2i}) + \varepsilon_{hi} \quad 4.7$$

Berikut merupakan letak titik knot untuk satu titik knot, *bandwidth* optimum, dan GCV yang diperoleh, dapat disajikan pada Tabel 4.4 berikut.

Tabel 4. 4. Letak Titik Knot, *Bandwidth*, GCV untuk Model Dua Titik Knot

Parameter	Titik Knot dan <i>Bandwidth</i>						
K_{111}	60,06	60,06	...	67,09	...	67,09	67,09
K_{112}	68,26	68,26	...	75,29	...	76,46	76,46
K_{211}	70,61	70,61	...	65,92	...	67,09	67,09
K_{212}	77,63	77,63	...	77,64	...	69,43	69,43
K_{311}	68,26	68,26	...	75,29	...	69,43	69,43
K_{312}	70,61	70,61	...	77,64	...	76,46	76,46
α_{11}	0,48	0,95	...	0,48	...	3,34	3,81
α_{12}	1,79	3,59	...	1,79	...	12,55	14,35
α_{21}	0,39	0,78	...	0,39	...	2,73	3,12
α_{22}	1,47	2,93	...	1,47	...	10,27	11,74
α_{31}	0,33	0,66	...	0,33	...	2,31	2,64
α_{32}	1,24	2,48	...	1,24	...	8,69	9,93
GCV	43,66	49,80	...	18,06*)	...	30,89	31,00

*) GCV Minimum

Dengan K_{111} merupakan titik knot pertama untuk variabel y_1 , K_{112} merupakan titik knot kedua untuk variabel y_1 , K_{211} merupakan titik knot pertama untuk variabel y_2 , K_{212} merupakan titik knot kedua untuk variabel y_2 , K_{311} merupakan titik knot pertama untuk variabel y_3 , K_{312} merupakan titik knot kedua untuk variabel y_3 . α_{11} adalah *bandwidth* untuk variabel t_1 pada y_1 , α_{12} adalah *bandwidth* untuk variabel t_2 pada y_1 , α_{21} adalah *bandwidth* untuk variabel t_1 pada y_2 , α_{22} adalah *bandwidth* untuk variabel t_2 pada y_2 . Sedangkan α_{31} adalah *bandwidth* untuk variabel t_1 pada y_3 , dan α_{32} adalah *bandwidth* untuk variabel t_2 pada y_3 .

4.6 Model dengan Titik Knot dan *Bandwidth* Optimal

Model terbaik diperoleh berdasarkan kombinasi titik knot dan *bandwidth* dengan nilai GCV terkecil. Secara ringkas GCV terkecil dari masing-masing kombinasi model terlihat pada Tabel 4.5 berikut.

Tabel 4. 5. Perbandingan GCV

Jumlah Titik Knot		Jumlah <i>Bandwidth</i>	GCV Terkecil
$K_{111} = 59,11$		$\alpha_{11} = 1,91$ $\alpha_{12} = 7,17$	10,97*)
$K_{211} = 59,61$		$\alpha_{21} = 1,56$ $\alpha_{22} = 5,87$	
$K_{311} = 58,62$		$\alpha_{31} = 1,32$ $\alpha_{32} = 4,97$	
$K_{111} = 67,09$	$K_{112} = 75,29$	$\alpha_{11} = 0,48$ $\alpha_{12} = 1,79$	18,06
$K_{211} = 65,92$	$K_{212} = 77,64$	$\alpha_{21} = 0,39$ $\alpha_{22} = 1,47$	
$K_{311} = 75,29$	$K_{312} = 77,64$	$\alpha_{31} = 0,33$ $\alpha_{32} = 1,24$	

*) GCV Minimum

Berdasarkan Tabel 4.5 di atas, dapat disimpulkan bahwa GCV terkecil berasal dari kombinasi satu titik knot dan dua *bandwidth*. Selanjutnya, hasil yang diperoleh akan digunakan untuk pemodelan Indikator Kemiskinan di Nusa Tenggara. Berikut merupakan estimasi parameter untuk model regresi

nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel, berdasarkan GCV terkecil yaitu dengan satu titik knot dan dua *bandwidth*, yang disajikan pada Tabel 4.6 berikut.

Tabel 4. 6. Estimasi Parameter pada Tiga Indikator Kemiskinan

Variabel Respon	Variabel Prediktor	Parameter	Estimasi
y_1	x_1	α_{10}	0,391
		α_1	-0,058
		β_{11}	-0,113
y_2	x_1	α_{20}	-1,956
		α_2	-0,182
		β_{21}	0,681
y_3	x_1	α_{30}	-0,017
		α_3	-0,155
		β_{31}	-0,028

Berdasarkan Tabel 4.5 dan Tabel 4.6 di atas, model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel untuk ketiga Indikator Kemiskinan dapat disajikan pada persamaan berikut.

$$\hat{y}_1 = 0,391 - 0,058x_{1i} - 0,113(x_{1i} - 59.12)_+ + \sum_{i=1}^{32} W_{11}y_{1i} + \sum_{i=1}^{32} W_{12}y_{2i} \quad 4.8$$

$$\text{dimana } W_{11} = \left[\frac{K\left(\frac{t_1 - t_{1i}}{1,91}\right)}{\sum_{i=1}^{32} K\left(\frac{t_1 - t_{1i}}{1,91}\right)} \right] \text{ dan } W_{12} = \left[\frac{K\left(\frac{t_2 - t_{2i}}{7,17}\right)}{\sum_{i=1}^{32} K\left(\frac{t_2 - t_{2i}}{7,17}\right)} \right]$$

$$\hat{y}_2 = -1,956 - 0,182x_{1i} + 0,681(x_{1i} - 59.61)_+ + \sum_{i=1}^{32} W_{21}y_{2i} + \sum_{i=1}^{32} W_{22}y_{2i} \quad 4.9$$

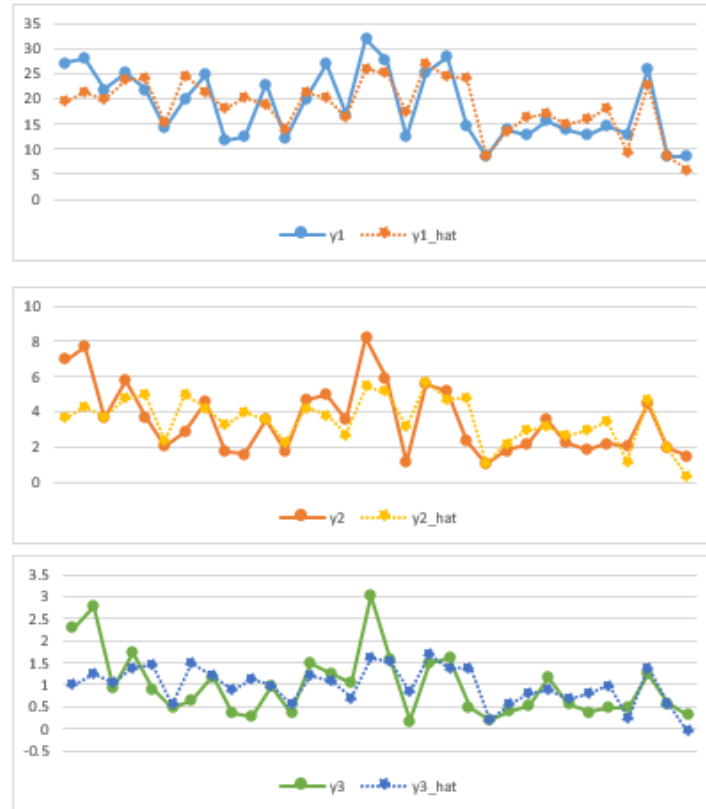
$$\text{dimana } W_{21} = \left[\frac{K\left(\frac{t_1 - t_{1i}}{1,56}\right)}{\sum_{i=1}^{32} K\left(\frac{t_1 - t_{1i}}{1,56}\right)} \right] \text{ dan } W_{22} = \left[\frac{K\left(\frac{t_2 - t_{2i}}{5,87}\right)}{\sum_{i=1}^{32} K\left(\frac{t_2 - t_{2i}}{5,87}\right)} \right]$$

$$\hat{y}_3 = -0,017 - 0,155x_{1i} - 0,028(x_{1i} - 58.62)_+ + \sum_{i=1}^{32} W_{31}y_{3i} + \sum_{i=1}^{32} W_{32}y_{3i} \quad 4.10$$

$$\text{dimana } W_{31} = \left[\frac{K\left(\frac{t_1 - t_{1i}}{1,32}\right)}{\sum_{i=1}^{32} K\left(\frac{t_1 - t_{1i}}{1,32}\right)} \right] \text{ dan } W_{32} = \left[\frac{K\left(\frac{t_2 - t_{2i}}{4,97}\right)}{\sum_{i=1}^{32} K\left(\frac{t_2 - t_{2i}}{4,97}\right)} \right]$$

Model yang telah diperoleh juga dapat digunakan untuk menjelaskan variabel-variabel yang mempengaruhi ketiga indikator kemiskinan, dalam hal ini Persentase Penduduk Miskin, Indeks Kedalaman Kemiskinan, dan Indeks Keparahan Kemiskinan di Provinsi Nusa Tenggara Barat dan Nusa Tenggara Timur. Namun, sebelum dilakukan interpretasi pengaruh variabel prediktor terhadap

variabel respon, dilihat secara visual selisih antara y dan $\hat{y}$. Berdasarkan model tersebut, selisih antara y dan $\hat{y}$ dapat juga disajikan dalam grafik perbandingan pada Gambar 4.4.



Gambar 4. 4. Grafik Perbandingan y dan $\hat{y}$ pada Data Indikator Kemiskinan

Pada model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel dengan satu titik knot dan dua *bandwidth*, diperoleh nilai $R^2 = 89,86\%$ dan $MSE = 2,84$ yang mengandung makna bahwa variabel prediktor dapat menjelaskan variasi dari variabel respon dengan baik sebesar 89,86% sisanya dipengaruhi oleh variabel prediktor lain. Kemudian, akan dilakukan interpretasi untuk melihat pengaruh variabel prediktor terhadap variabel respon. Namun, hanya hanya variabel prediktor yang mengikuti fungsi spline *truncated* yang dapat diinterpretasikan, sedangkan untuk variabel prediktor yang mengikuti fungsi kernel tidak dapat diinterpretasikan.

Pada penelitian ini, interpretasi model hanya dilakukan pada variabel prediktor Indeks Pembangunan Manusia (IPM), karena variabel tersebut merupakan variabel prediktor yang didekati dengan fungsi spline *truncated*. Hal pertama yang dilakukan adalah melakukan interpretasi model untuk variabel respon pertama yaitu Persentase Penduduk Miskin. Akan dilihat pengaruh dari variabel prediktor yang didekati dengan fungsi spline *truncated*, dan variabel prediktor yang didekati dengan fungsi kernel dianggap konstan. Berdasarkan persamaan (4.8), fungsi spline *truncated* linier yang terbentuk untuk variabel IPM terhadap Indikator Kemiskinan dapat disajikan sebagai berikut.

a. Variabel Respon: Persentase Penduduk Miskin/ P_0 (y_1)

Apabila variabel t_1 dan t_2 dianggap konstan, maka fungsi spline *truncated* pengaruh Indeks Pembangunan Manusia (x_1) terhadap Persentase Penduduk Miskin (y_1) adalah sebagaimana disajikan pada persamaan (4.11).

$$\hat{y}_1 = -0,058x_{1i} - 0,113(x_{1i} - 59,12)_+ + c_{1i} = \begin{cases} -0,058x_{1i} + c_{1i} & ; 0 \leq x_{1i} < 59,12 \\ -0,171x_{1i} + 6,68 + c_{1i} & ; x_{1i} \geq 59,12 \end{cases} \quad 4.11$$

Berdasarkan persamaan (4.11), jika suatu kabupaten/kota memiliki nilai Indeks Pembangunan Manusia (IPM) dibawah 59,12% dan nilai IPM naik satu satuan, maka akan menurunkan Persentase Penduduk Miskin di kabupaten/kota tersebut sebesar 0,058%.

- b. Variabel Respon: Indeks Kedalaman Kemiskinan/ P_1 (y_2)

Apabila variabel t_1 dan t_2 dianggap konstan, maka fungsi spline truncated pengaruh Indeks Pembangunan Manusia (x_1) terhadap Indeks Kedalaman Kemiskinan (y_2) adalah sebagaimana disajikan pada persamaan (4.12).

$$\hat{y}_2 = -0,182x_{2i} + 0,681(x_{2i} - 59,61)_+ + c_{2i} = \begin{cases} -0,182x_{2i} + c_{2i} & ; 0 \leq x_{2i} < 59,61 \\ 0,499x_{2i} - 40,59 + c_{2i} & ; x_{2i} \geq 59,61 \end{cases} \quad 4.12$$

Berdasarkan persamaan (4.12), jika suatu kabupaten/kota memiliki nilai Indeks Pembangunan Manusia (IPM) dibawah 59,61% dan nilai IPM naik satu satuan, maka akan menurunkan Indeks Kedalaman Kemiskinan di kabupaten/kota tersebut sebesar 0,182. Kemudian jika suatu kabupaten/kota memiliki nilai IPM lebih dari 59,61%, dan nilai IPM naik satu satuan, maka akan menurunkan Indeks Kedalaman Kemiskinan di kabupaten/kota tersebut.

- c. Variabel Respon: Indeks Keparahan Kemiskinan/ P_2 (y_3)

Apabila variabel t_1 dan t_2 dianggap konstan, maka fungsi spline *truncated* pengaruh Indeks Pembangunan Manusia (x_1) terhadap Indeks Keparahan Kemiskinan (y_3) adalah sebagaimana disajikan pada persamaan (4.13).

$$\hat{y}_3 = -0,155x_{3i} - 0,028(x_{3i} - 58,62)_+ + c_{3i} = \begin{cases} -0,155x_{3i} + c_{3i} & ; 0 \leq x_{3i} < 58,62 \\ -0,183x_{3i} + 1,64 + c_{3i} & ; x_{3i} \geq 58,62 \end{cases} \quad 4.13$$

Berdasarkan persamaan (4.13), jika suatu kabupaten/kota memiliki nilai Indeks Pembangunan Manusia (IPM) dibawah 58,62% dan nilai IPM naik satu satuan, maka akan menurunkan Indeks Keparahan Kemiskinan di kabupaten/kota tersebut sebesar 0,155. Kemudian jika suatu kabupaten/kota memiliki nilai IPM lebih dari 58,62%, dan nilai IPM naik satu satuan, maka akan menurunkan Indeks Keparahan Kemiskinan di kabupaten/kota tersebut.

Hasil penelitian ini menunjukkan bahwa model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel memberikan hasil yang sangat baik. Terbukti dengan nilai koefisien determinasi $R^2 = 89,86\%$ yang diperoleh yang termasuk sangat tinggi. Nilai ini menunjukkan bahwa model memiliki kemampuan yang tinggi dalam menjelaskan variasi Indikator Kemiskinan di provinsi Nusa Tenggara Barat dan Nusa Tenggara Timur. Oleh karena itu, dapat disimpulkan bahwa dengan pendekatan estimator campuran spline *truncated* dan kernel, mampu menangkap karakteristik data yang berbeda antar variabel prediktor secara lebih adaptif. Jika dibandingkan dengan estimator tunggal, penggunaan estimator campuran akan memberikan kesempatan bagi data yang memiliki pola/bentuk kurva yang berbeda antar variabel prediktor untuk diperlakukan dengan pendekatan yang berbeda juga. Jika tetap menggunakan satu jenis estimator untuk semua variabel prediktor, maka setiap variabel akan dipaksa mengikuti pendekatan estimator yang sama. Seharusnya setiap variabel prediktor diperlakukan sesuai dengan karakteristik atau pola kurva regresinya masing-masing, agar diperoleh model yang optimal dan andal.

Model regresi nonparametrik multirespon dengan estimator campuran spline *truncated* dan kernel merupakan pengembangan dari model regresi nonparametrik birespon dengan estimator campuran yang sama. Sebelumnya [8] menerapkan hanya pada kasus birespon. Dengan memperluas ke regresi nonparametrik multirespon, penelitian ini memberikan kontribusi baru pada literatur

regresi nonparametrik, khususnya untuk data Indikator Kemiskinan yang kompleks dan multidimensi. Selain memberikan hasil estimasi yang akurat, pendekatan ini juga memiliki potensi yang luas dalam pemodelan data sosial ekonomi yang lain dengan struktur variabel yang kompleks.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan maka dapat disimpulkan bahwa Model terbaik yang dihasilkan berdasarkan titik knot dan *bandwidth* optimum dengan memperhatikan GCV terkecil adalah sebagai berikut.

$$\hat{y}_1 = 0,391 - 0,058x_{1i} - 0,113(x_{1i} - 59.12)_+ + \sum_{i=1}^{32} W_{11}y_{1i} + \sum_{i=1}^{32} W_{12}y_{1i}$$

$$\text{dimana } W_{11} = \left[\frac{K\left(\frac{t_1 - t_{1i}}{1,91}\right)}{\sum_{i=1}^{32} K\left(\frac{t_1 - t_{1i}}{1,91}\right)} \right] \text{ dan } W_{12} = \left[\frac{K\left(\frac{t_2 - t_{2i}}{7,17}\right)}{\sum_{i=1}^{32} K\left(\frac{t_2 - t_{2i}}{7,17}\right)} \right]$$

$$\hat{y}_2 = -1,956 - 0,182x_{1i} + 0,681(x_{1i} - 59.61)_+ + \sum_{i=1}^{32} W_{21}y_{2i} + \sum_{i=1}^{32} W_{22}y_{2i}$$

$$\text{dimana } W_{21} = \left[\frac{K\left(\frac{t_1 - t_{1i}}{1,56}\right)}{\sum_{i=1}^{32} K\left(\frac{t_1 - t_{1i}}{1,56}\right)} \right] \text{ dan } W_{22} = \left[\frac{K\left(\frac{t_2 - t_{2i}}{5,87}\right)}{\sum_{i=1}^{32} K\left(\frac{t_2 - t_{2i}}{5,87}\right)} \right]$$

$$\hat{y}_3 = -0,017 - 0,155x_{1i} - 0,028(x_{1i} - 58.62)_+ + \sum_{i=1}^{32} W_{31}y_{3i} + \sum_{i=1}^{32} W_{32}y_{3i}$$

$$\text{dimana } W_{31} = \left[\frac{K\left(\frac{t_1 - t_{1i}}{1,32}\right)}{\sum_{i=1}^{32} K\left(\frac{t_1 - t_{1i}}{1,32}\right)} \right] \text{ dan } W_{32} = \left[\frac{K\left(\frac{t_2 - t_{2i}}{4,97}\right)}{\sum_{i=1}^{32} K\left(\frac{t_2 - t_{2i}}{4,97}\right)} \right]$$

Model regresi di atas memiliki koefisien determinasi $R^2 = 89,86\%$ dan $MSE = 2,84$. Dengan demikian, model campuran spline *truncated* dan kernel dapat digunakan untuk memodelkan dan menjelaskan hubungan antara Indikator Kemiskinan dalam hal ini Persentase Penduduk Miskin, Indeks Kedalaman Kemiskinan, dan Indeks Keparahan Kemiskinan dengan variabel-variabel prediktor Indeks Pembangunan Manusia (IPM), Tingkat Pengangguran Terbuka (TPT), dan laju PDRB/kapita sebesar 89,86%.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik. 2023. *Profil Kemiskinan di Indonesia Maret 2023*. Badan Pusat Statistik, Jakarta.
- [2] Budiantara, I. N., 2019. Regresi Linear Berganda dan Asumsinya. *Jurnal Statistika dan Aplikasinya*, Vol. 18, No. 2, 150–165.
- [3] Budiantara, I. N., 2019. Regresi Nonparametrik dengan Pendekatan Spline. *Jurnal Matematika dan Statistika*, Vol. 19, No. 2, 123–145.

- [4] Chamidah, N. & Saifudin, T., 2013. Estimation of Children Growth Curve Based on Kernel Smoothing in Multi-Response Nonparametric Regression. *Applied Mathematical Sciences*, Vol. 7, No. 37, 1839-1847.
- [5] Cox, D. D. & O'sullivan, F., 1996. Penalized Likelihood-type Estimators for Generalized Nonparametric Regression. *Journal of Multivariate Analysis*, Vol. 56, No.2, 185-206.
- [6] Craven, P. & Wahba, G., 1978. Smoothing Noisy Data with Spline Functions. *Numerische Mathematik*, Vol. 31, No. 4, 377-403.
- [7] Dani, A. T. R. & Adrianingsih, N. Y., 2021. Pemodelan Regresi Nonparametrik dengan Estimator Spline *Truncated* dan Deret Fourier. *Jambura Journal of Mathematics*, Vol. 3, No. 1, 26-36.
- [8] Dewanti, P., 2020. Estimator Campuran Spline *Truncated* dan Kernel dalam Regresi Nonparametrik Birespon untuk Pemodelan Faktor- Faktor Yang Mempengaruhi Keberhasilan Pembangunan di Provinsi Jawa Timur. *Institut Teknologi Sepuluh Nopember*, Surabaya.
- [9] Eubank, R. L. 1999. *Nonparametric Regression and Spline Smoothing 2nd Edition*. Marcel Dekker, New York.
- [10] Fan, S. & Rao, N., 2003. Public Spending in Developing Countries: Trends, Determination, And Impact. *International Food Policy Research Institute (IFPRI)*, EPTD Discussion Paper, No. 99.
- [11] Härdle, W. 1994. *Applied Nonparametric Regression*. Cambrige University Press, New York.
- [12] Pratiwi, L. P. S. & Wijaya, I M. P. P., 2022. Perbandingan Estimator Spline dan Kernel dalam Model Regresi Nonparametrik Menggunakan Data PDRB di Indonesia. *Jurnal Ilmiah Matematika dan Terapan*, Vol. 19, No. 2, 144-152.
- [13] Ramadhani, E., Salwa, N. & Mazaya, M. S., 2020. Identifikasi Faktor-Faktor yang Memengaruhi Angka Harapan Hidup di Sumatera Tahun 2018 Menggunakan Analisis Regresi Spasial Pendekatan Area. *Journal of Data Analysis*, Vol.3, No. 2, 62-75.
- [14] Tripena, A. & Prabowo, A., 2021. Estimated Spline in Nonparametric Regression with a Generalized Cross Validation and Unbiased Risk Approach. *Proceeding of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore*, 3788-3798.
- [15] Wahba, G. 1990. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia.
- [16] Wang, Y., Li, Z. & Liu, X., 2000. Multivariate Nonparametric Regression Analysis Using Smoothing Splines. *Journal of Statistical Planning and Inference*, Vol. 92, No. 1, 27-42.