

PEMODELAN JUMLAH PENDERITA DEMAM BERDARAH DENGUE (DBD) DI JAWA TIMUR MENGGUNAKAN ANALISIS SPLINE TRUNCATED TAHUN 2023**Nala Kamila Azizy**Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu
email: nalaziee@gmail.com**Idhia Sriliana**Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu
email: idhiasriliana@unib.ac.id***Ranisyah Anggraini**Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu
email: ranisyahanggraini892@gmail.com**Aditama Rouliber Simanullang**Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu
email: aditamasimanullang16@gmail.com**Nur Azmi Setiawati**Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu
email: nurazmisetiawati@gmail.com**Pepi Novianti**Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu
email: pie_novianti@unib.ac.id**Abstrak**

Demam Berdarah Dengue (DBD) merupakan salah satu penyakit tropis yang masih menjadi permasalahan kesehatan di Indonesia, terutama di wilayah dengan kepadatan penduduk tinggi seperti di Jawa Timur. Untuk memahami pola penyebaran kasus DBD, penelitian ini menggunakan metode *Spline Truncated* orde pertama dalam pemodelan jumlah penderita DBD. Metode ini dipilih karena kemampuannya dalam menangkap pola hubungan non-linier yang mungkin terjadi pada data jumlah kasus DBD berdasarkan variabel lainnya. Penelitian ini bertujuan untuk menganalisis jumlah penderita DBD menggunakan pendekatan regresi *spline*, dengan menentukan titik knot secara tepat guna memperoleh model terbaik. Hasil penelitian menunjukkan bahwa model *Spline Truncated* dengan tiga titik knot mampu menggambarkan pola perubahan jumlah penderita DBD dengan akurasi yang tinggi, ditunjukkan oleh nilai *Generalized Cross Validation* (GCV) sebesar 293,4351 dan koefisien determinasi (R^2) sebesar 69,55%. Hasil penelitian menunjukkan variabel persentase kepadatan penduduk, sumber air minum ledeng, dan akses sanitasi layak berpengaruh signifikan terhadap jumlah kasus penderita DBD di Jawa Timur.

Kata Kunci: Demam Berdarah Dengue, *Spline Truncated*, Regresi Nonparametrik.

Abstract

Dengue Hemorrhagic Fever (DHF) is one of the tropical diseases that is still a health problem in Indonesia, especially in areas with high population density such as in East Java. To understand the distribution pattern of DHF cases, this study used the first-order Truncated Spline method in modeling the number of DHF patients. This method was chosen because of its ability to capture non-linear relationship patterns that may occur in the data of the number of DHF cases based on other variables. This study aims to analyze the number of DHF patients using a spline regression approach, by determining the knot points appropriately to obtain the best model. The results showed that the Truncated Spline model with three knot points was able to describe the pattern of changes in the number of DHF patients with high accuracy, indicated by the Generalized Cross Validation (GCV) value of 293.4351 and the coefficient of determination (R^2) of 69.55%. The results showed that the variable percentage of population density, piped drinking water sources, and access to proper sanitation had a significant effect on the number of DHF cases in East Java.

Keywords: *Dengue Hemorrhagic Fever, Spline Truncated, Nonparametric Regression.*

PENDAHULUAN

Demam Berdarah Dengue (DBD) adalah penyakit infeksi yang disebabkan oleh *Dengue Virus* (DENV) serotype 1-4 yang ditransmisikan oleh nyamuk *Aedes aegypti* (Nugraheni dkk., 2023). Nyamuk *Aedes aegypti* memiliki suatu karakteristik yang dapat menentukan persebaran kejadian infeksi yakni habitatnya yang pada umumnya berada di wilayah beriklim tropis, dengan curah hujan yang tinggi, serta bersuhu panas maupun lembab. Selain itu, nyamuk ini juga suka pada tempat penampungan dan genangan air sebagai tempat berkembang biaknya (Kemenkes RI 2022, 2023). *Dengue Virus* melemahkan sistem imun manusia, menyebabkan penurunan ketahanan tubuh secara cepat. DBD telah lama menjadi penyakit menular yang menjadi salah satu penyebab utama kematian di seluruh dunia (Mustafidah & Purnama, 2024).

Berdasarkan data Dinas Kesehatan Prov. Jawa Timur kasus DBD per Desember 2024 telah terjadi peningkatan baik di awal maupun di akhir tahun. Tercatat, pada Semester I tahun 2024 jumlah kasus DBD mencapai 21.959 kasus dan pada Semester II sebesar 7.537 kasus. Ada beberapa faktor yang mempengaruhi kejadian penyakit Demam Berdarah Dengue (DBD), salah satunya adalah faktor lingkungan sosial seperti kepadatan penduduk, dan dukungan tenaga medis (Andayani dkk., 2024). Sanitasi lingkungan berhubungan dengan proses pertumbuhan dan perkembangbiakan nyamuk *Aedes aegypti*, yang cenderung berkembang biak di wadah-wadah yang dapat menampung air hujan dan tidak dilapisi tanah (Inayah dkk., 2021).

Sehubungan dengan tingginya kasus DBD di Provinsi Jawa Timur tahun 2024, penelitian ini bermaksud untuk memahami lebih lanjut bagaimana faktor lingkungan sosial berkontribusi terhadap kejadian DBD dengan menggunakan data tahun sebelumnya, yaitu tahun 2023. Analisis akan dilakukan menggunakan regresi nonparametrik *spline truncated* untuk memodelkan jumlah kejadian DBD di Jawa Timur. Fungsi dasar yang digunakan dalam pendekatan regresi *spline* adalah fungsi *truncated*, dan orde dari model regresi *spline truncated* dapat dikenali melalui pola hubungan antar variabel. Salah satu aspek penting yang tidak boleh diabaikan adalah penentuan jumlah dan posisi titik knot, karena keduanya sangat memengaruhi bentuk

keseluruhan kurva regresi (Ramadhan dkk., 2023). Metode regresi nonparametrik dengan estimator *spline* ini mampu menangani data yang memiliki pola perilaku yang berubah-ubah pada interval-interval tertentu (Husain dkk., 2024).

Penelitian sebelumnya oleh Mubarak & Budiantara (2012) memodelkan kematian akibat DBD di Jawa Timur menggunakan regresi *spline multivariable*. Dari 13 variabel prediktor yang dianalisis, ditemukan bahwa angka bebas jentik, curah hujan, dan persentase penduduk miskin berpengaruh signifikan terhadap kematian DBD, dengan model terbaik menggunakan *spline* linier dengan satu titik knot. Penelitian ini akan menggunakan variabel prediktor yang berbeda untuk menggali faktor-faktor baru yang berpotensi memengaruhi jumlah kejadian DBD di Jawa Timur tahun 2023. Dengan menerapkan regresi *spline truncated*, diharapkan dapat diperoleh model yang lebih adaptif terhadap pola perubahan data dan memberikan pemahaman yang lebih mendalam mengenai hubungan antara faktor lingkungan sosial dengan kejadian DBD. Hasil penelitian ini diharapkan dapat menjadi dasar rekomendasi kebijakan pencegahan DBD yang lebih tepat sasaran di Provinsi Jawa Timur.

KAJIAN TEORI

REGRESI NONPARAMETRIK SPLINE TRUNCATED

Regresi nonparametrik *spline truncated* adalah metode regresi nonparametrik yang paling banyak digunakan (Wongkar, Ruliana, & Fahmuddin, 2023). Pendekatan *spline* merupakan bagian dari pendekatan dalam regresi nonparametrik yang dicirikan dengan potongan (truncated) polinomial tersemen dan kontinu (Arman dkk., 2024). Secara umum, bentuk dari fungsi *spline truncated* derajat m dengan titik-titik knot K_1, K_2, \dots, K_r dapat dinyatakan sebagai berikut (Dani dkk., 2024):

$$m(x_i) = \sum_{j=0}^m \gamma_j x_i^j + \sum_{k=1}^r \gamma_{m+k} (x_i - K_k)_+^m \quad (1)$$

Di mana fungsi *truncated* diberikan oleh:

$$(x_i - K_k)_+^m = \begin{cases} (x_i - K_k)_+^m & x \geq K_k \\ 0 & x < K_k \end{cases} \quad (2)$$

Jika $m = 1$, maka disebut sebagai *spline linear*, $m = 2$ disebut sebagai *spline kuadratik*, $m = 3$ disebut sebagai *spline kubik*.

PEMILIHAN TITIK KNOT

Salah satu metode dalam pemilihan titik knot optimal adalah *Generalized Cross Validation* (GCV). Model regresi *spline* terbaik dengan titik knot optimal berdasarkan nilai GCV minimum. Berikut rumus yang digunakan untuk menghitung nilai GCV (Tetrapoik dkk., 2023):

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[I - A(K_1, K_2, \dots, K_r)])^2} \quad (3)$$

Di mana:

n : jumlah observasi

I : matriks identitas berukuran $n \times n$

$A(K_1, K_2, \dots, K_r)$: matriks $X(X^T X)^{-1} X^T$

$$MSE(K_1, K_2, \dots, K_r) : n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

PENGUJIAN ASUMSI RESIDUAL

Metode uji Shapiro-Wilk merupakan salah satu alternatif prosedur dalam pengujian normalitas. Misalkan X_1, X_2, \dots, X_n merupakan statistik terurut dari suatu pengamatan, maka statistik uji Shapiro-Wilk (W) adalah (Ahadi & Zain, 2023):

$$W = \frac{(\sum_1^n a_i x_i)^2}{(\sum_1^n x_i - \bar{x})^2} \quad (4)$$

Selanjutnya W dibandingkan dengan daerah kritis untuk dilakukan pengambilan keputusan. Tolak H_0 jika nilai $W <$ nilai kritis atau $p\text{-value} < \alpha$.

Uji identik atau uji heteroskedastisitas merupakan keadaan di mana asumsi homoskedastisitas tidak terpenuhi. Uji Breusch-Pagan merupakan salah satu metode statistik yang dapat mendekripsi heteroskedastisitas. Uji ini digunakan untuk menguji adanya efek waktu, individu atau keduanya dengan hipotesis sebagai berikut (Indrasetianingsih & Wasik, 2020).

$$H_0: \sigma_e^2 = 0 \text{ (Common Effect Model)}$$

$$H_1: \text{minimal ada } \sigma_e^2 \neq 0 \text{ (Random Effect Model)}$$

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N [T\bar{e}]^2}{\sum_{i=1}^N \sum_{t=1}^T \sigma_{it}^2} - 1 \right]^2 \quad (5)$$

Di mana T adalah jumlah unit time series, N adalah jumlah unit cross section, dan e_{it} adalah residual pada unit ke- i dan waktu ke- t .

Uji Durbin-Watson adalah uji formal untuk mendekripsi ada tidaknya autokorelasi dengan nilai DW yang mendekati 2 dianggap bahwa model tidak ada autokorelasi. Untuk mendekripsi adanya autokorelasi pada model regresi dapat digunakan uji statistik Durbin-Watson dengan hipotesis (Madany dkk., 2022):

$$H_0: \rho = 0 \text{ (tidak autokorelasi)}$$

$$H_1: \rho = 1 \text{ (ada autokorelasi)}$$

Statistik uji yang digunakan adalah:

$$DW = \frac{\sum_{t=2}^N (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=2}^N \varepsilon_t^2} \quad (6)$$

Di mana d_l dan d_u adalah batas bawah dan batas atas nilai kritis yang dapat dicari melalui tabel Durbin-Watson berdasarkan k jumlah variabel bebas dan n adalah jumlah smapel yang relevan.

METODE

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari data Badan Pusat Statistik Provinsi Jawa Timur tahun 2023. Unit penelitian yang diamati adalah masing-masing kabupaten/kota di Provinsi Jawa Timur yang terdiri dari 38 kabupaten/kota. Variabel yang digunakan meliputi variabel dependen (Y) dan variabel independen (X). Jumlah kasus penyakit DBD per 100.000 penduduk (Y), Persentase kepadatan penduduk (penduduk / km²) (X_1), Persentase rumah tangga dengan sumber air minum ledeng (X_2), dan Persentase rumah tangga yang memiliki akses terhadap sanitasi layak (X_3). Berikut adalah langkah-langkah yang dilakukan dalam penelitian ini:

1. Mengumpulkan data dari laman BPS Provinsi Jawa Timur
2. Analisis deskriptif pada setiap variabel.
3. Membuat *scatterplot* untuk melihat pola hubungan antara variabel dependen dan masing-masing variabel independen.
4. Memodelkan variabel respon menggunakan *spline truncated* orde 1 (*spline linear*) dengan berbagai titik knot.
5. Menentukan titik knot optimal yang didasarkan pada nilai GCV dan R^2 .
6. Menetapkan model *spline truncated* optimal.
7. Menguji signifikansi parameter secara serentak dan individu.
8. Membuat *plot* prediksi pada model *spline truncated*.
9. Melakukan uji asumsi residual IIDN (Identik, Independen, dan Distribusi Normal).

HASIL DAN PEMBAHASAN

ANALISIS DESKRIPТИF

Provinsi Jawa Timur terdiri dari 38 Kabupaten/Kota dengan rincian 29 Kabupaten dan 9 Kota. Berikut ini merupakan statistika deskriptif

data Jumlah kasus penyakit Demam Berdarah Dengue (DBD) dan faktor lain yang mempengaruhinya.

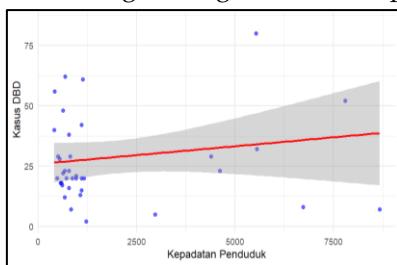
Tabel 1. Hasil Analisis Deskriptif

	Y	X_1	X_2	X_3
Minimum	2	0,27	50,3	410
kuartil 1	17,25	3,803	81,58	644,5
Median	21,5	6,965	87,84	855,5
Mean	28,71	9,924	84,59	1937,9
Kuartil 3	36,5	14,043	93,77	1204,2
Maksimum	95	36,61	98,18	8667

Berdasarkan Tabel 1, Distribusi kasus DBD (Y) cukup bervariasi dengan rata-rata 28,71, tetapi ada daerah dengan jumlah kasus yang jauh lebih tinggi yaitu di Kota Probolinggo. Rata-rata Kepadatan Penduduk (X_1) sangat tinggi yaitu sebesar 1937,9 penduduk/km² dengan beberapa daerah sangat padat hingga 8667 penduduk/km². Distribusinya tidak merata. Rata-rata 9,92% menunjukkan penggunaan Sumber Air Ledeng (X_2) yang masih rendah dengan daerah sangat minim akses terdapat di Kabupaten Gresik. Mayoritas daerah memiliki Akses Sanitasi Layak (X_3) yang baik dengan rata-rata 84,59%, tetapi ada daerah dengan akses yang cukup rendah terdapat di Kabupaten Bangkalan.

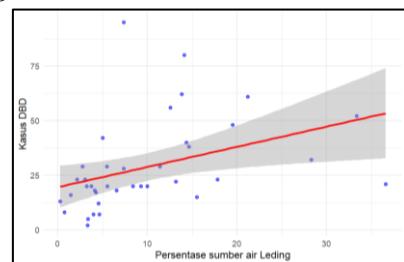
POLA HUBUNGAN ANTAR VARIABEL

Berikut ini merupakan *scatterplot* pada variabel dependen dan masing-masing variabel independen:

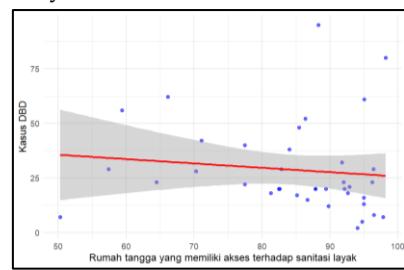
Gambar 1. Scatterplot variabel Y dan X_1

Pada Gambar 1, *scatter plot* menunjukkan hubungan antara kepadatan penduduk dengan jumlah penderita DBD. Jika terlihat pola positif, di mana semakin tinggi kepadatan penduduk maka jumlah kasus DBD juga meningkat, hal ini dapat dijelaskan oleh tingginya interaksi manusia yang mempermudah penyebaran virus dengue serta banyaknya tempat potensial bagi nyamuk berkembang biak. Namun, jika pola tidak terlalu jelas atau menyebar, berarti ada faktor lain yang lebih dominan dalam mempengaruhi jumlah kasus DBD,

seperti pola hidup masyarakat atau kebersihan lingkungan.

Gambar 2. Scatterplot variabel Y dan X_2

Pada Gambar 2, hubungan antara persentase rumah tangga dengan sumber air minum ledeng dan jumlah kasus DBD bisa bersifat kompleks. Jika *scatterplot* menunjukkan tren positif, artinya wilayah dengan akses air ledeng lebih tinggi justru mengalami peningkatan kasus DBD. Sebaliknya, jika pola menunjukkan tren negatif, maka akses air bersih mungkin membantu mengurangi risiko penyebaran penyakit dengan menghilangkan sumber air yang digunakan nyamuk untuk bertelur.

Gambar 3. Scatterplot variabel Y dan X_3

Pada Gambar 3, *scatterplot* menggambarkan bagaimana akses sanitasi layak berkaitan dengan jumlah penderita DBD. Jika *scatterplot* menunjukkan tren menurun, maka wilayah dengan sanitasi yang lebih baik memiliki kasus DBD yang lebih sedikit. Namun, jika tidak ada pola hubungan yang jelas atau terdapat fluktuasi, ini bisa menunjukkan bahwa sanitasi layak bukan satu-satunya faktor yang berpengaruh, dan ada variabel lain yang lebih dominan dalam menentukan jumlah kasus DBD.

MODEL SPLINE TRUNCATED DENGAN SATU TITIK KNOT

Berikut ini merupakan Nilai GCV dengan menggunakan 1 titik knot pada regresi nonparametrik *spline truncated*.

Tabel 2. Nilai GCV dengan Satu Titik Knot

No	Lokasi Titik Knot			GCV
	X_1	X_2	X_3	
1.	4538,500	18,440000	74,240	303,9592
2.	4882,542	19,954167	76,235	308,1323

3.	4194,458	16,925833	72,245	308,8709
----	----------	-----------	--------	----------

Berdasarkan Tabel 2, diperoleh nilai GCV minimum yaitu sebesar 303,9592. Model regresi nonparametrik *spline truncated* dapat dituliskan sebagai berikut.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 (x_{i1} - 4538,5)_+ + \hat{\beta}_3 x_{i2} + \hat{\beta}_4 (x_{i2} - 18,44)_+ + \hat{\beta}_5 x_{i3} + \hat{\beta}_6 (x_{i3} - 74,24)_+$$

MODEL SPLINE TRUNCATED DENGAN DUA TITIK KNOT

Berikut ini merupakan Nilai GCV dengan menggunakan 2 titik knot pada regresi nonparametrik *spline truncated*.

Tabel 3. Nilai GCV dengan Dua Titik Knot

No	Lokasi Titik Knot			GCV
	X ₁	X ₂	X	
1.	5570,625	22,98250	80,225	311,9088
	6602,750	27,52500	86,210	
2.	5914,667	22,98250	82,220	312,7654
	6602,750	27,52500	86,210	
3.	5570,625	22,98250	80,225	312,9032
	6258,708	26,01083	84,215	

Berdasarkan Tabel 3, diperoleh nilai GCV minimum yaitu sebesar 311,9088. Model regresi nonparametrik *spline truncated* dapat dituliskan sebagai berikut.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 (x_{i1} - 5570,625)_+ + \hat{\beta}_3 (x_{i1} - 6602,750)_+ + \hat{\beta}_4 x_{i2} + \hat{\beta}_5 (x_{i2} - 22,9825)_+ + \hat{\beta}_6 (x_{i2} - 27,525)_+ + \hat{\beta}_7 x_{i3} + \hat{\beta}_8 (x_{i3} - 80,225)_+ + \hat{\beta}_9 (x_{i3} - 86,21)_+$$

MODEL SPLINE TRUNCATED DENGAN TIGA TITIK KNOT

Berikut ini merupakan Nilai GCV dengan menggunakan 3 titik knot pada regresi nonparametrik *spline truncated*.

Tabel 4. Nilai GCV dengan Tiga Titik Knot

No	Lokasi Titik Knot			GCV
	X ₁	X ₂	X ₃	
1.	4194,458	16,92583	72,245	293,4351
	5226,583	21,46833	78,23	
	6602,75	27,525	86,21	
2.	4538,5	18,44	74,24	293,6163
	4882,542	19,95417	76,235	
	8322,958	35,09583	96,185	
3.	3850,417	15,41167	70,25	293,8536
	5226,583	21,46833	78,23	
	6602,75	27,525	86,21	

Berdasarkan Tabel 4, diperoleh nilai GCV minimum yaitu sebesar 293,4351. Model regresi

nonparametrik *spline truncated* dapat dituliskan sebagai berikut.

$$\begin{aligned}\hat{y}_i = & \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 (x_{i1} - 4194,458)_+ \\ & + \hat{\beta}_3 (x_{i1} - 5226,583)_+ \\ & + \hat{\beta}_4 (x_{i1} - 6602,75)_+ + \hat{\beta}_5 x_{i2} \\ & + \hat{\beta}_6 (x_{i2} - 16,92583)_+ \\ & + \hat{\beta}_7 (x_{i2} - 21,46833)_+ \\ & + \hat{\beta}_8 (x_{i2} - 27,525)_+ + \hat{\beta}_9 x_{i3} \\ & + \hat{\beta}_{10} (x_{i3} - 72,245)_+ \\ & + \hat{\beta}_{11} (x_{i3} - 78,23)_+ \\ & + \hat{\beta}_{12} (x_{i3} - 86,21)_+\end{aligned}$$

PEMILIHAN KNOT OPTIMAL

Berdasarkan Tabel 2, 3, dan 4 maka dapat dibuat tabel ringkasan hasil nilai GCV yang terbaik.

Tabel 5. Perbandingan 3 Model

No	Lokasi Titik Knot			GCV	R ²
	X ₁	X ₂	X ₃		
1.	4538,500	18,44000	74,240	303,9592	51,29%
	5570,625	22,98250	80,225		
2.	6602,750	27,52500	86,210	311,908	59,22%
	4194,458	16,925	72,245		
3.	5226,583	21,468	78,23	293,4351	69,55%
	6602,75	27,525	86,21		

Tabel 5 menunjukkan nilai GCV terkecil dan koefisien determinasi (R²) terbesar diperoleh pada model regresi nonparametrik *spline truncated* dengan tiga titik knot. Jadi, dapat disimpulkan bahwa pemodelan regresi nonparametrik *spline truncated* dengan tiga titik knot merupakan model dengan titik knot yang paling optimal.

PENGUJIAN SIGNIFIKANSI PARAMETER

Pengujian signifikansi parameter dilakukan untuk melihat apakah model *Spline Truncated* memberikan kontribusi yang signifikan dalam menjelaskan variabel respon (y). Terdapat dua pengujian yang dilakukan, yaitu uji simultan (serentak) dan uji parsial. Hipotesis pada uji simultan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{12} = 0$$

$$H_1: \text{Minimal terdapat satu } \beta_i \neq 0, i = 1, 2, \dots, 12$$

Dengan menggunakan taraf signifikansi 5%, kriteria penolakannya adalah jika $F_{\text{hitung}} > F_{\text{tabel}}$ atau nilai $P_{\text{value}} < \alpha$ maka tolak H_0 .

Tabel 6. Hasil Uji Serentak

Sumber Keragaman	Jumlah Kuadrat	Rataan Kuadrat	F _{hitung}	P _{value}
Regresi	10978,7	914,8914	7,235 175	0,000 069

Error	4805,119	126,4505		
Total	15783,82			

Berdasarkan Tabel 6, diperoleh nilai F_{hitung} sebesar 7,235175 dan P_{value} sebesar 0,000069. Hal ini menunjukkan bahwa dengan taraf signifikansi 5% atau $\alpha = 0,05$ maka tolak H_0 , karena nilai $P_{value} < \alpha$. Artinya, secara simultan variabel independen berpengaruh signifikan terhadap variabel dependen atau setidaknya terdapat satu variabel yang signifikan.

Pengujian secara parsial digunakan untuk melihat parameter mana yang signifikan terhadap model. Hipotesis pada uji parsial adalah sebagai berikut.

$$H_0: \beta_i = 0 \text{ vs } H_1: \beta_i \neq 0, i = 1, 2, \dots, 12$$

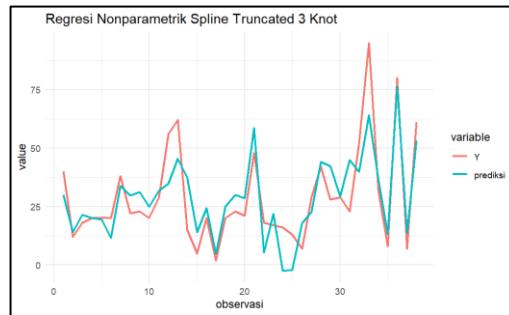
Dengan menggunakan taraf signifikansi 5%, kriteria penolakannya adalah jika $t_{hitung} > t_{tabel}$ atau nilai $P_{value} < \alpha$ maka tolak H_0 .

Tabel 7. Hasil Uji Parsial

Variabel	Parameter	Estimasi	P_{value}	Keputusan
Intersep	β_0	-50,5733	0,206	Tidak signifikan
X_1	β_1	0,006281	0,050	Signifikan
	β_2	1,12985	0,042	Signifikan
	β_3	1,153155	0,071	Tidak signifikan
	β_4	0,058148	0,020	Signifikan
	β_5	-0,11336	0,002	Signifikan
X_2	β_6	0,049059	0,028	Signifikan
	β_7	7,328245	0,039	Signifikan
	β_8	-25,5346	0,000	Signifikan
	β_9	24,98684	0,000	Signifikan
X_3	β_{10}	-5,31661	0,025	Signifikan
	β_{11}	4,955883	0,092	Tidak signifikan
	β_{12}	-3,43959	0,053	Tidak signifikan

Berdasarkan Tabel 7, hasil estimasi parameter secara parsial menggunakan metode regresi nonparametrik *spline truncated* dengan tiga titik knot didapatkan 9 parameter yang signifikan. Parameter yang signifikan terdiri dari $\beta_1, \beta_2, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9$, dan β_{10} , hal ini dikarenakan nilai $p\text{-value} < \alpha$.

PLOT PREDIKSI SPLINE TRUNCATED



Gambar 2. Plot prediksi

Berdasarkan *plot* hasil regresi nonparametrik *spline truncated* dengan 3 knot, terlihat bahwa garis prediksi cukup mengikuti pola data aktual, yang menunjukkan bahwa model berhasil menangkap tren utama dalam data. Namun, terdapat beberapa penyimpangan pada beberapa titik, terutama di bagian puncak dan lembah yang tajam, yang mengindikasikan bahwa model mungkin masih mengalami sedikit *error* dalam menangkap fluktuasi ekstrem. Meskipun demikian, secara keseluruhan, model ini cukup baik dalam merepresentasikan pola data, dengan kesalahan prediksi yang relatif kecil. Evaluasi lebih lanjut seperti analisis residual, dapat dilakukan untuk memastikan bahwa model ini tidak mengalami masalah asumsi seperti independen atau identik.

PENGUJIAN ASUMSI RESIDUAL

Suatu model regresi dengan parameter signifikan dan memenuhi kriteria terbaik tetapi melanggar asumsi residual tidak disarankan untuk dipakai untuk menggambarkan pola hubungan antara variabel prediktor dan variabel respon. Asumsi residual dalam analisis regresi meliputi residual berdistribusi Normal, identik dan independen. Asumsi- asumsi ini sangat penting untuk dilakukan karena bertujuan untuk melihat akurasi model yang terbentuk. Uji Normalitas dapat membantu memastikan koefisien regresi *spline truncated*-nya tidak bias. Berikut adalah hasil uji asumsi residual normal.

Tabel 8. Uji Shapiro Wilk

W	P_{value}
0,95875	0,1728

Berdasarkan Tabel 8, dengan menggunakan taraf signifikansi 5% maka dapat disimpulkan bahwa $p\text{-value} > \alpha$. Artinya pada pengujian asumsi ini residual berdistribusi normal.

Uji identik atau uji homoskedastisitas bertujuan untuk melihat varians residual bersifat konstan atau berubah-ubah. Berikut adalah hasil uji asumsi identik.

Tabel 9. Uji Breush Pagan

BP	P_value
3,2907	0,06967

Berdasarkan Tabel 9, dengan menggunakan taraf signifikansi 5% maka dapat disimpulkan bahwa $p-value > \alpha$. Artinya pada pengujian asumsi ini residual bersifat konstan atau homoskedastisitas.

Uji independen atau uji autokorelasi bertujuan untuk melihat apakah variabel independen saling berkorelasi atau tidak. Berikut adalah hasil uji independen.

Tabel 10. Uji Durbin Watson

DW	P_value
1,9074	0,3791

Berdasarkan Tabel 10, dengan menggunakan taraf signifikansi 5% maka dapat disimpulkan bahwa $p-value > \alpha$. Artinya pada pengujian asumsi ini residual tidak berkorelasi.

INTERPRETASI MODEL REGRESI NONPARAMETRIK SPLINE TRUNCATED

Model yang diperoleh pada regresi nonparametrik *spline truncated* orde satu dengan tiga titik knot adalah sebagai berikut

$$\hat{y}_i = 0,006x_{i1} + 1,129(x_{i1} - 4194,458)_+ + 0,058(x_{i1} - 6602,75)_+ - 0,113x_{i2} + 0,049(x_{i2} - 16,925)_+ + 7,328(x_{i2} - 21,468)_+ - 25,534(x_{i2} - 27,525)_+ + 24,986x_{i3} - 5,316(x_{i3} - 72,245)_+$$

Interpretasi dari setiap variabel yang berpengaruh secara signifikan terhadap jumlah penderita DBD adalah sebagai berikut.

1. Pengaruh dari persentase kepadatan penduduk (penduduk/km²) (X_1) terhadap Jumlah kasus penyakit DBD per 100.000 penduduk (Y) dengan mengasumsikan variabel prediktor lain konstan dapat dituliskan sebagai berikut.

$$\hat{y}_i = 0,006x_{i1} + 1,129(x_{i1} - 4194,458)_+ + 0,058(x_{i1} - 6602,75)_+ \\ 0,006x_{i1} ; x_{i1} < 4194,458 \\ 1,135x_{i1} - 4735,54 ; 4194,458 \leq x_{i1} < 6602,75 \\ 1,193x_{i1} - 5118,49 ; x_{i1} \geq 6602,75$$

2. Pengaruh dari persentase rumah tangga dengan sumber air minum ledeng (X_2) terhadap jumlah kasus penyakit DBD per 100.000 penduduk (Y)

dengan mengasumsikan variabel prediktor lain konstan dapat dituliskan sebagai berikut.

$$\hat{y}_i = -0,113x_{i2} + 0,049(x_{i2} - 16,925)_+ + 7,328(x_{i2} - 21,468)_+ - 25,534(x_{i2} - 27,525)_+$$

$$\hat{y}_i = \begin{cases} -0,113x_{i2} ; x_{i2} < 16,925 \\ -0,064x_{i2} - 0,8293 ; 16,925 \leq x_{i2} < 21,468 \\ 7,264x_{i2} - 158,146 ; 21,468 \leq x_{i2} < 27,525 \\ -18,27x_{i2} + 544,677 ; x_{i2} \geq 27,525 \end{cases}$$

3. Pengaruh dari persentase rumah tangga yang memiliki akses terhadap sanitasi layak (X_3) terhadap jumlah kasus penyakit DBD per 100.000 penduduk (Y) dengan mengasumsikan variabel prediktor lain konstan dapat dituliskan sebagai berikut.

$$\hat{y}_i = 24,986x_{i3} - 5,316(x_{i3} - 72,245)_+$$

$$\hat{y}_i = \begin{cases} 24,986x_{i3} ; x_{i3} < 72,245 \\ 19,67x_{i3} + 384,054 ; x_{i3} \geq 72,245 \end{cases}$$

Diperoleh hasil berdasarkan hubungan antara persentase kepadatan penduduk (penduduk/km²) terhadap jumlah kasus penyakit DBD per 100.000 penduduk. Jika persentase kepadatan penduduk terhadap jumlah kasus penyakit DBD kurang dari 4194,458 dan bertambah 1 satuan maka jumlah kasus penyakit DBD cenderung meningkat sebesar 0,006 per 100.000 penduduk. Apabila persentase kepadatan penduduk terhadap jumlah kasus penyakit DBD berada dalam rentang 4194,458 hingga 6602,75 dan bertambah 1 satuan maka jumlah kasus penyakit DBD mengalami peningkatan sebesar 1,135 per 100.000 penduduk. Apabila persentase kepadatan penduduk terhadap jumlah kasus penyakit DBD melebihi 6602,75 dan bertambah 1 satuan maka jumlah kasus penyakit DBD mengalami peningkatan sebesar 1,193 per 100.000 penduduk.

PENUTUP

SIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, diperoleh kesimpulan bahwa jumlah penderita Demam Berdarah Dengue (DBD) cukup bervariasi dengan rata-rata 28,71, tetapi ada daerah dengan jumlah kasus yang jauh lebih tinggi yaitu di Kota Probolinggo. Pemilihan knot optimal yang dilihat dari nilai GCV minimum dan diperoleh nilai GCV minimum yaitu sebesar 293.4351 terdapat pada pemodelan dengan menggunakan tiga titik knot. Model yang dihasilkan dengan menggunakan tiga

titik knot memiliki nilai koefisien determinasi (R^2) sebesar 69,55%. Artinya variabel independen dapat menjelaskan 69,55% variabel dependen dan sisanya dijelaskan oleh variabel independen lain.

Berdasarkan hasil analisis, pemerintah perlu memprioritaskan upaya pengendalian DBD pada area dengan kepadatan penduduk tinggi dan melakukan investigasi mendalam terkait sistem sumber air minum ledeng, termasuk kemungkinan kebocoran atau kebiasaan penampungan air yang tidak tepat, mengingat keduanya signifikan meningkatkan kasus DBD. Sementara itu, akses sanitasi layak perlu dievaluasi kembali efektivitasnya dalam konteks pencegahan DBD, karena hasil menunjukkan hal yang bertentangan dengan prinsip kesehatan masyarakat.

SARAN

Pada penelitian ini penulis masih membatasi metode *spline truncated* pada orde linier dengan tiga titik knot. Permasalahan yang lebih detail belum dikaji secara mendalam. Oleh karena itu, saran untuk penelitian selanjutnya adalah mencoba analisis dengan metode regresi nonparametrik *spline* orde kuadratik dan kubik atau menggunakan metode regresi nonparametrik lainnya.

DAFTAR PUSTAKA

- Ahadi, G. D., & Zain, N. N. L. E. (2023). Pemeriksaan Uji Kenormalan dengan Kolmogorov-Smirnov, Anderson-Darling dan Shapiro-Wilk. *Eigen Mathematics Journal*, 6(1), 11–19. <https://doi.org/10.29303/emj.v6i1.131>
- Andayani, S. A., Avita, H., Agustina, D., Fitri, E. W., & Handayani, E. (2024). Hubungan Antara Peran Kader Demam Berdarah dan Indeks Bebas Larva (LFI) Terhadap Kasus Demam Berdarah. 5(2), 1027–1033.
- Arman, A. N., Lemido, R., Siswanto, S., & Kalondeng, A. (2024). Pemodelan Regresi Nonparametrik Dengan Estimator Spline Polynomial Truncated Pada Data Jumlah Wisatawan Nusantara. *MATHunesa: Jurnal Ilmiah Matematika*, 12(1), 127–133. <https://doi.org/10.26740/mathunesa.v12n1.p127-133>
- Dani, A. T. R., Putra, F. B., Ni'matuzzahroh, L., Ratnasari, V., & Budiantara, I. N. (2024). *Regresi Parametrik dan Nonparametrik, Teori dan Aplikasi dengan Software R*.
- Husain, H., Dewi, A. F., & Wardani, A. E. (2024). Modeling of Indonesian Stunting Prevalence Using Spline Truncated Non-parametric Regression. *Journal of Analytical Research*, 3(1), 1–13.
- Inayah, Z., Aldawiyah, K. K., Maraqonitatillah, N., Meidyna, S. N., & Chasanah, A. G. (2021). Pengaruh Kebersihan Lingkungan terhadap Penyebaran Demam Berdarah Dengue (DBD) di Wilayah Puskesmas Gending Kabupaten Gresik. 07(1), 139–151.
- Indrasetianingsih, A., & Wasik, T. K. (2020). Model Regresi Data Panel Untuk Mengetahui Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Pulau Madura. *Jurnal Gaussian*, 9(3), 355–363. <https://doi.org/10.14710/j.gauss.v9i3.28925>
- Kemenkes RI 2022. (2023). Profil Kesehatan Indonesia. Dalam *Pusdatin.Kemenkes.Go.Id*. <https://www.kemkes.go.id/downloads/resources/download/pusdatin/profil-kesehatan-indonesia/Profil-Kesehatan-2021.pdf>
- Madany, N., Ruliana, & Rais, Z. (2022). Regresi Data Panel dan Aplikasinya dalam Kinerja Keuangan terhadap Pertumbuhan Laba Perusahaan Idx Lq45 Bursa Efek Indonesia. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(2), 79–94. <https://doi.org/10.35580/variansium28>
- Mubarak, R., & Budiantara, I. N. (2012). Analisis Regresi Spline Multivariabel Untuk Pemodelan Kematian Penderita Demam Berdarah Dengue (DBD) Di Jawa Timur. *Jurnal Sains dan Seni ITS*, 1(1), D224–D229. http://ejurnal.its.ac.id/index.php/sains_seni/article/view/2025/335
- Mustafidah, M. E., & Purnama, M. D. (2024). Pengelompokan Kabupaten /Kota Di Jawa Timur Berdasarkan Indikator Kasus Dbd Menggunakan Complete Linkage Dan Average Linkage. *Jurnal Ilmiah Matematika*, 12(2), 337–343.
- Nugraheni, E., Rizqoh, D., & Sundari, M. (2023). Manifestasi Klinis Demam Berdarah Dengue (Dbd). *Jurnal Kedokteran dan Kesehatan : Publikasi Ilmiah Fakultas Kedokteran Universitas Sriwijaya*, 10(3), 267–274. <https://doi.org/10.32539/jkk.v10i3.21425>
- Ramadhan, M. R., Darnah, D., & Wahyuningsih, S. (2023). Model Regresi Nonparametrik Spline Truncated Pada Indeks Pembangunan Manusia di Indonesia. *Eksponensial*, 14(2), 91. <https://doi.org/10.30872/eksponensial.v14i2.1090>
- Tetrapoik, A. E., Lembang, F. K., Ilwaru, V. Y. I., & Lewaherilla, N. (2023). Pemodelan Regresi Nonparametrik Spline Dan Aplikasinya Pada Indeks Kebahagiaan Provinsi Di Indonesia. *Jurnal Riset dan Aplikasi Matematika (JRAM)*,

- 7(1), 56–65.
<https://doi.org/10.26740/jram.v7n1.p56-65>
- Wongkar, D. C., & Fahmuddin, M. S. (2023). Analisis Regresi Nonparametrik Spline Truncated untuk Menganalisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 5(2), 55–63. <https://doi.org/10.35580/variansiunm101>