

Pemodelan Faktor-Faktor yang Mempengaruhi Angka Morbiditas di Jawa Timur Menggunakan Regresi Nonparametrik Spline

Krisna Wulandari, I Nyoman Budiantara, dan Madu Ratna

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: i_nyoman_b@statistika.its.ac.id

Abstrak—Morbiditas adalah kondisi seseorang dikatakan sakit apabila keluhan kesehatan yang dirasakan menyebabkan terganggunya aktivitas sehari-hari yaitu tidak dapat melakukan kegiatan bekerja, mengurus rumah tangga, dan kegiatan normal sebagaimana biasanya. Semakin tinggi morbiditas, menunjukkan derajat kesehatan penduduk yang semakin buruk. Berdasarkan hasil Survei Sosial Ekonomi Nasional (SUSENAS), pada tahun 2010 angka morbiditas Jawa Timur sebesar 28,4%, sedangkan pada tahun 2014 angka morbiditas Jawa Timur mencapai 30,21%. Adanya peningkatan angka morbiditas tersebut tentunya disebabkan oleh faktor-faktor baik gaya hidup maupun kondisi lingkungan di wilayah tersebut. Oleh karena itu, dilakukan sebuah penelitian berupa pemodelan untuk mengetahui faktor-faktor yang mempengaruhi morbiditas di Jawa Timur. Berdasarkan data yang diperoleh dari SUSENAS Jawa Timur tahun 2014, menunjukkan bahwa pola hubungan antara morbiditas penduduk dengan faktor-faktor yang mempengaruhinya tidak berpola tertentu, sehingga pada penelitian ini digunakan metode regresi nonparametrik spline. Hasil penelitian ini menunjukkan bahwa model regresi terbaik menggunakan 3 titik knot dengan 6 variabel yang signifikan yaitu kepadatan penduduk, rata-rata lama sekolah, persentase penduduk miskin, Upah Minimum Kabupaten, persentase rumah tangga *Open Defecation* (OD), dan persentase rumah tangga dengan jarak sumber air minum ke tempat penampungan kotoran lebih dari 10 meter. Nilai kebaikan model atau R^2 yang diperoleh yaitu sebesar 89,72%.

Kata Kunci—Angka Morbiditas, GCV, Jawa Timur, Regresi Nonparametrik Spline, Titik Knot.

I. PENDAHULUAN

TUJUAN pembangunan berkelanjutan atau *Sustainable Development Goals* (SDGs) yang telah disepakati oleh Perserikatan Bangsa-Bangsa (PBB) mempunyai fokus utama, yaitu meningkatkan kesejahteraan manusia dalam berbagai aspek kehidupan. Salah satu aspek yang menjadi alat pengukur tingkat kesejahteraan manusia di Indonesia disebut Indeks Pembangunan Manusia (IPM). Sumber Daya Manusia (SDM) yang sehat secara fisik diharapkan menjadi manusia berkualitas yang dapat ikut berperan dalam pembangunan untuk mewujudkan kesejahteraan rakyat [1]. Sehat adalah suatu keadaan yang sempurna baik secara fisik, mental, dan sosial, serta bebas dari penyakit atau kelemahan. Oleh karena itu, indikator kesehatan suatu daerah dapat ditinjau dari jumlah penduduk yang mengalami kesakitan atau terjangkit suatu penyakit [2].

Angka morbiditas merupakan ukuran tentang kesakitan atau

keluhan kesehatan di suatu wilayah [3]. Angka morbiditas mempunyai peranan yang lebih penting dibandingkan dengan angka kematian [4]. Hal ini disebabkan apabila angka kesakitan tinggi, maka akan memicu kematian sehingga menyebabkan angka kematian juga tinggi, sehingga angka harapan hidup di suatu wilayah akan rendah.

Jawa Timur merupakan provinsi dengan luas wilayah terbesar di Pulau Jawa. Pada tahun 2010 angka morbiditas penduduk Jawa Timur sebesar 28,4%, namun terjadi peningkatan pada tahun 2014 menjadi 30,21%. Peningkatan tersebut disebabkan oleh faktor-faktor yang mempengaruhi. Menurut Lembaga Demografi UI [5] serta penelitian tentang morbiditas yang dilakukan Hanum [6], Ardhiyanti [7], Fuhrer[8], dan Arola [9] faktor determinan morbiditas merupakan faktor sosial, ekonomi, dan budaya.

Berdasarkan pola data variabel prediktor terhadap variabel respon yang tidak mengikuti pola tertentu, maka akan dilakukan pemodelan menggunakan regresi nonparametrik spline [10]. Penelitian dengan metode regresi nonparametrik spline telah dilakukan oleh Ramadhani [11] tentang jumlah pengangguran terbuka. Metode ini mempunyai sifat tersegmen, sehingga mempunyai fleksibilitas tinggi dan dapat menyesuaikan diri terhadap karakteristik lokal suatu data [12].

II. TINJAUAN PUSTAKA

A. Regresi Nonparametrik Spline

Regresi nonparametrik merupakan suatu metode Statistika yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk fungsinya, hanya diasumsikan *smooth* (mulus) dalam arti termuat dalam suatu ruang fungsi tertentu. Regresi nonparametrik merupakan regresi yang sangat fleksibel dalam memodelkan pola data [12]. Model regresi nonparametrik secara umum dapat disajikan sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i, \quad i=1,2,3,\dots,n \quad (1)$$

dimana y_i merupakan variabel respon, x_i merupakan variabel prediktor, $f(x_i)$ merupakan fungsi regresi yang tidak mengikuti pola tertentu, dan $\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$. Apabila kurva regresi f merupakan model aditif dan dihipotesiskan dengan fungsi spline maka diperoleh model regresi sebagai berikut,

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

dimana :

$$f(x_{ji}) = \sum_{h=0}^q \beta_{hj} x_{ji}^h + \sum_{l=1}^m \beta_{(q+l)j} (x_{ji} - k_{lj})_+^q + \varepsilon_i, i=1,2,\dots,n$$

dengan,

$$(x_{ji} - k_{lj})_+^q = \begin{cases} (x_{ji} - k_{lj})^q, & x_{ji} \geq k_{lj} \\ 0, & x_{ji} \leq k_{lj} \end{cases} \quad (3)$$

dan $k_{1j}, k_{2j}, \dots, k_{mj}$ merupakan titik knot yang menunjukkan bentuk perubahan perilaku dari fungsi pada sub-sub interval tertentu. Nilai q merupakan derajat polinomial. Kurva polinomial derajat satu disebut kurva linier, kurva polinomial derajat dua disebut kurva kuadratik serta derajat tiga disebut kurva kubik. Persamaan (2) dapat diuraikan sebagai berikut:

$$y_i = \beta_{01} + \beta_{11}x_{i1} + K + \beta_{q1}x_{i1}^q + a_{11}(x_{i1} - k_{11})_+^q + K + a_{m1}(x_{i1} - k_{m1})_+^q + \beta_{02} + \beta_{12}x_{i2} + K + \beta_{q2}x_{i2}^q + a_{12}(x_{i2} - k_{12})_+^q + K + a_{m2}(x_{i2} - k_{m2})_+^q + K + \beta_{0p} + \beta_{1p}x_{ip} + K + \beta_{qp}x_{ip}^q + a_{1p}(x_{ip} - k_{1p})_+^q + K + a_{mp}(x_{ip} - k_{mp})_+^q + \varepsilon_i$$

Estimasi parameter model regresi nonparametrik spline dilakukan dengan menggunakan metode *Ordinary Least Square* (OLS) sebagai berikut.

$$\hat{\beta}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}_{OLS} \quad (4)$$

B. Pemilihan Titik Knot Optimal

Salah satu metode yang digunakan untuk pemilihan titik knot optimal adalah *Generalized Cross Validation* (GCV). Metode GCV mempunyai sifat optimal asimtotik jika dibandingkan dengan metode lain, misalnya *Cross Validation* (CV) [13]. Model spline dengan nilai GCV terkecil dari titik knot optimal merupakan model spline yang terbaik. Metode GCV dapat dituliskan sebagai berikut [12],

$$GCV(k) = \frac{MSE(k)}{[n^{-1} \text{trace}(\mathbf{I} - \mathbf{A})]^2} \quad (5)$$

dimana \mathbf{I} adalah matriks identitas, n merupakan jumlah pengamatan, $k = (k_1, k_2, \dots, k_m)$ merupakan titik-titik knot, dan

$$MSE(k) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (6)$$

serta \mathbf{A} didapat dari matriks $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'$.

C. Pengujian Parameter Model

Diberikan model regresi nonparametrik spline derajat q dengan knot-knot k_1, k_2, \dots, k_m :

$$y_i = \sum_{h=0}^q \beta_h x_i^h + \sum_{l=1}^m \beta_{q+l} (x_i - k_l)_+^q + \varepsilon_i \quad (7)$$

1. Pengujian Parameter Model Secara Serentak

Uji serentak dilakukan untuk mengetahui signifikansi parameter model regresi secara bersama-sama.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{q+m} = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, q + m$$

Nilai $q+m$ merupakan banyak parameter dalam model regresi nonparametrik spline kecuali β_0

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{MSR}{MSE} \quad (8)$$

H_0 ditolak apabila $F_{hitung} > F_{\alpha; (s, n-(q+m)-1)}$ atau $p\text{-value} < \alpha$.

2. Pengujian Parameter Model Secara Individu

Pengujian secara individu dilakukan untuk mengetahui apakah parameter secara individual mempunyai pengaruh yang signifikan terhadap variabel respon, dengan hipotesis sebagai berikut:

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, k = 1, 2, \dots, q + m$$

Pengujian secara individu dilakukan dengan menggunakan uji t [15]. Statistik uji yang digunakan adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_k}{\sqrt{\text{var}(\hat{\beta}_k)}} \quad (9)$$

dengan

$$\text{var}(\hat{\beta}_k) = \text{diag}[(\mathbf{X}'\mathbf{X})^{-1} \hat{\sigma}^2]_k$$

dimana $\hat{\sigma}^2$ merupakan *MSE*. H_0 ditolak apabila $|t_{hitung}| > t_{\alpha/2; (n-(q+m)-1)}$ atau $p\text{-value} < \alpha$. R^2 digunakan sebagai indikator kebaikan model, yang diberikan sebagai berikut.

$$R^2 = \frac{\hat{\beta}^T \mathbf{X}^T \mathbf{y} - n\bar{y}}{\mathbf{y}^T \mathbf{y} - n\bar{y}^2} \times 100\% \quad (10)$$

D. Pengujian Asumsi Residual

1. Uji Normalitas Kolmogorov-Smirnov

Uji *Kolmogorov-Smirnov* bertujuan untuk mengetahui apakah suatu data telah mengikuti suatu distribusi tertentu.

$$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$$

$$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon)$$

Statistik uji yang digunakan adalah sebagai berikut.

$$D = \text{Sup}_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (11)$$

Tolak H_0 apabila $D > D_{\alpha}$.

2. Asumsi Identik

Asumsi identik terpenuhi apabila varians antar residual sama, atau tidak terjadi heteroskedastisitas [15].

$$\text{var}(y_i) = \text{var}(\varepsilon_i) = \sigma^2; i = 1, 2, \dots, n \quad (12)$$

Uji identik dapat menggunakan uji Glejser. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, n$$

Langkah pertama dalam uji Glejser yaitu apabila telah didapatkan ε dari model regresi, selanjutnya dilakukan regresi $|\varepsilon_i|$ terhadap X sehingga diperoleh nilai F_{hitung} . Statistik uji yang digunakan sebagaimana Persamaan (13).

$$F_{hitung} = \frac{\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|)}{v-1} \quad (13)$$

$$= \frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)}{n-v}$$

H_0 ditolak apabila $F_{hitung} > F_{\alpha;(v-1,n-v)}$ atau $p\text{-value} < \alpha$, dimana nilai v adalah banyaknya parameter model Glejser.

E. Morbiditas

Morbiditas adalah kondisi seseorang dikatakan sakit apabila keluhan kesehatan yang dirasakan menyebabkan terganggunya aktivitas sehari-hari yaitu tidak dapat melakukan kegiatan bekerja, mengurus rumah tangga, dan kegiatan normal sebagaimana biasanya [3]. Berikut merupakan rumus angka morbiditas.

$$AM = \frac{JPKK}{JP} \times 100 \tag{14}$$

Keterangan :

- AM : angka morbiditas
- JPKK : jumlah penduduk yang mengalami keluhan kesehatan dan terganggunya aktivitas
- JP : jumlah penduduk

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari Badan Pusat Statistika Provinsi Jawa Timur, Laporan Survei Sosial Ekonomi Nasional (SUSENAS) dan Profil Kesehatan Provinsi Jawa Timur tahun 2014. Unit penelitian yang digunakan merupakan 38 kabupaten dan kota di provinsi Jawa Timur.

B. Variabel Penelitian

Variabel respon yang digunakan dalam penelitian ini adalah angka morbiditas penduduk provinsi Jawa Timur, sedangkan variabel prediktor yang diduga berpengaruh adalah sebagai berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
Y	Angka morbiditas
X ₁	Kepadatan penduduk
X ₂	Angka buta huruf
X ₃	Rata-rata lama sekolah penduduk 15 tahun keatas
X ₄	Persentase penduduk miskin
X ₅	Upah Minimum Kabupaten/ Kota (UMK)
X ₆	Persentase rumah tangga <i>Open Defecation</i> (OD)
X ₇	Persentase rumah tangga dengan jarak sumber air minum ke tempat penampungan kotoran lebih dari 10 meter

C. Langkah-Langkah Penelitian

Langkah-langkah analisis dalam penelitian adalah sebagai berikut.

1. Mendeskripsikan angka morbiditas di Jawa Timur dan faktor-faktor yang mempengaruhinya.
2. Membuat *scatter plot* antara variabel respon dengan masing-masing variabel prediktor untuk mengetahui pola hubungan yang terjadi.
3. Memodelkan variabel respon (angka morbiditas) menggunakan model regresi nonparametrik spline dengan berbagai titik knot.
4. Memilih titik knot optimal berdasarkan nilai GCV yang paling minimum.

5. Mendapatkan model regresi spline terbaik dengan titik knot optimal.
6. Melakukan uji signifikansi parameter secara serentak dan parsial.
7. Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN) dari model regresi spline.
8. Membuat interpretasi model dan menarik kesimpulan.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Angka Morbiditas di Jawa Timur

Pada tahun 2014 angka morbiditas di Jawa Timur mencapai 30,21%. Hal ini berarti terdapat 30 penduduk dari 100 penduduk di Jawa Timur yang mengalami keluhan sakit. Angka morbiditas tertinggi terletak pada Kota Kediri yaitu sebesar 41,24, sedangkan angka morbiditas terendah terletak pada Kabupaten Sidoarjo yaitu sebesar 22,31. Kota Kediri mempunyai angka morbiditas tertinggi di Jawa Timur karena sejak tahun 2006 kondisi lingkungan Kota Kediri tercemar oleh polusi PT. Gudang Garam, sehingga mengakibatkan warga Kota Kediri mengalami gangguan ISPA. Kabupaten Sidoarjo memiliki angka morbiditas terendah disebabkan karena fasilitas kesehatan Kabupaten Sidoarjo yang baik. Pada tahun 2013 Kabupaten Sidoarjo membentuk Program Pembangunan Terpadu Berbasis Kawasan Kemiskinan (PTBK2). Berikut merupakan karakteristik data dari tujuh faktor yang diduga berpengaruh terhadap angka morbiditas di Jawa Timur.

Tabel 2. Statistika Deskriptif Faktor-faktor yang diduga Mempengaruhi Pengangguran di Indonesia

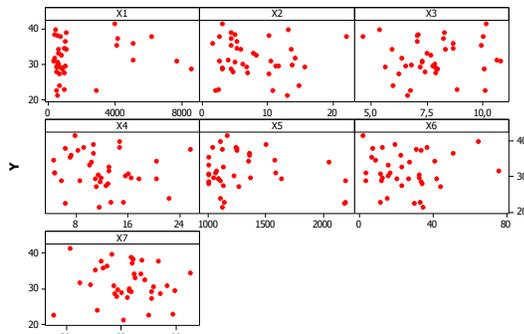
Variabel	Mean	Varians	Min	Maks
X ₁	1817	4708298	387	8562
X ₂	7,699	23,462	1,480	22,07
X ₃	7,750	2,345	4,620	10,81
X ₄	12,096	24,926	4,590	70,72
X ₅	1293,1	120395,4	1000	2200
X ₆	25,20	278,26	1,60	76,17
X ₇	63,93	258,36	23,19	98,31

Berdasarkan Tabel 2, kepadatan penduduk (X₁) terendah terletak pada Kabupaten Pacitan, sedangkan kepadatan penduduk tertinggi terletak pada Kota Surabaya, karena Kota Surabaya merupakan ibukota Provinsi Jawa Timur. Angka buta huruf (X₂) terendah terletak pada Kota Pasuruan, sedangkan angka buta huruf tertinggi terletak pada Kabupaten Sampang. Rata-rata lama sekolah (X₃) terendah terletak pada Kabupaten Sampang, sedangkan rata-rata lama sekolah tertinggi terletak pada Kota Malang, karena fasilitas pendidikan Kota Malang yang baik yaitu memiliki 80 perguruan tinggi. Persentase penduduk miskin tertinggi (X₄) adalah Kabupaten Sampang, karena infrastruktur yang kurang memadai, sedangkan Kota Batu menjadi kota dengan persentase penduduk miskin terendah. Kota Surabaya merupakan kota dengan UMK tertinggi. Untuk persentase rumah tangga OD tertinggi adalah Kabupaten Situbondo, sedangkan persentase OD (X₆) terendah adalah Kota Kediri. Persentase OD tertinggi mayoritas terdapat pada daerah tapal kuda, hal ini dikarenakan sumber air bersih yang kurang. Wilayah dengan persentase rumah tangga yang memiliki jarak sumber air minum ke tempat penampungan kotoran lebih dari 10 meter (X₇) terendah adalah Kabupaten

Sidoarjo, sedangkan nilai tertinggi adalah Kota Batu. Hal ini dikarenakan Kabupaten Sidoarjo merupakan kabupaten yang tergolong pada kepadatan penduduk tinggi di Jawa Timur.

B. Pola Data Antara Angka Morbiditas dengan Variabel yang Diduga Mempengaruhi

Langkah pertama dalam melakukan analisis regresi adalah membuat *scatter plot* untuk mengetahui pola hubungan variabel prediktor terhadap variabel respon yaitu angka morbiditas. Gambar 1 menunjukkan bahwa pola hubungan antara angka morbiditas dengan seluruh variabel prediktor tidak membentuk pola tertentu, sehingga metode yang digunakan adalah regresi nonparametrik spline.



Gambar 1 Scatterplot variabel prediktor terhadap respon

C. Pemilihan Titik Knot Optimal

Nilai GCV minimum pada pemilihan titik knot optimal dengan satu titik knot, dua titik knot, tiga titik knot, dan kombinasi titik knot ditampilkan sebagai berikut.

Tabel 3.
Nilai GCV Minimum Titik Knot Optimal

Model	GCV
1 Titik Knot	33,46
2 Titik Knot	31,88
3 Titik Knot	18,28
Kombinasi Titik Knot (1,2,2,3,3,2)	24,34

Berdasarkan kriteria pemilihan model terbaik diketahui bahwa nilai GCV paling minimum dihasilkan oleh model regresi nonparametrik spline dengan tiga titik knot.

D. Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline

1. Uji Serentak

Pengujian secara serentak ini dilakukan untuk menguji estimasi parameter model secara bersamaan (simultan). Nilai α yang digunakan sebesar 0,05. Berikut ini adalah hasil analisis ragam model regresi nonparametrik spline.

Tabel 4.
Analisis Ragam Uji Serentak

Sumber Variasi	db	JK	RJK	F	p-value
Regresi	28	960,131	34,29	5,302	0,0063
Residual	9	58,208	4,467		
Total	37	1018,34			

Berdasarkan Tabel 4 dapat diketahui bahwa nilai *p-value* sebesar 0,0063. Nilai ini lebih kecil daripada nilai α yaitu 0,05 sehingga tolak H_0 yang artinya minimal terdapat satu parameter

yang signifikan dalam model. Untuk mengetahui parameter mana yang signifikan dalam model, dilakukan pengujian secara individu.

2. Uji Individu

Pengujian secara individu dilakukan untuk mengetahui variabel yang berpengaruh signifikan terhadap angka morbiditas di Jawa Timur. Berikut ini adalah hasil pengujian signifikansi parameter model secara individu. Berdasarkan Tabel 5 diketahui terdapat lima variabel yang berpengaruh secara signifikan terhadap angka morbiditas yaitu $X_1, X_3, X_5, X_6,$ dan X_7 .

Tabel 5
Hasil Uji Individu

Variabel	Parameter	Estimasi	p-value	Ket.
X_1	β_0	2,15	0,518	Tidak Signifikan
	β_1	0,01	0,000	Signifikan
	β_2	-0,14	0,001	Signifikan
	β_3	0,53	0,002	Signifikan
	β_4	-0,40	0,003	Signifikan
X_2	β_5	-0,08	0,860	Tidak Signifikan
	β_6	-4,70	0,127	Tidak Signifikan
	β_7	23,09	0,092	Tidak Signifikan
	β_8	-15,74	0,173	Tidak Signifikan
X_3	β_9	-2,26	0,113	Tidak Signifikan
	β_{10}	-7,91	0,055	Tidak Signifikan
	β_{11}	7,20	0,028	Signifikan
	β_{12}	0,46	0,876	Tidak Signifikan
X_4	β_{13}	-0,60	0,153	Tidak Signifikan
	β_{14}	-0,84	0,539	Tidak Signifikan
	β_{15}	10,56	0,252	Tidak Signifikan
	β_{16}	-11,58	0,184	Tidak Signifikan
X_5	β_{17}	0,01	0,239	Tidak Signifikan
	β_{18}	-0,09	0,017	Signifikan
	β_{19}	1,33	0,050	Signifikan
	β_{20}	-1,30	0,060	Tidak Signifikan
X_6	β_{21}	-0,26	0,004	Signifikan
	β_{22}	1,39	0,047	Signifikan
	β_{23}	-13,83	0,010	Signifikan
	β_{24}	13,45	0,007	Signifikan
X_7	β_{25}	0,90	0,001	Signifikan
	β_{26}	-3,08	0,002	Signifikan
	β_{27}	12,04	0,002	Signifikan
	β_{28}	-9,98	0,002	Signifikan

E. Pemilihan Titik Knot Optimal Dengan Enam Variabel Prediktor

Selanjutnya dilakukan pemodelan kembali dengan menggunakan enam variabel prediktor yaitu dengan menghapus variabel X_2 terlebih dahulu. Variabel tersebut dihapus karena pada tahun 2014, indikator pendidikan telah diganti menjadi rata-rata lama sekolah atau variabel X_3 .

Tabel 6.
Nilai GCV Minimum Titik Knot Optimal Enam Prediktor

Model	GCV
1 Titik Knot	29,04
2 Titik Knot	28,66
3 Titik Knot	20,56
Kombinasi Titik Knot (3,2,2,3,3,3)	22,95

Berdasarkan Tabel 6 diketahui bahwa model dengan tiga titik knot menghasilkan GCV paling minimum yaitu 20,56.

F. Penaksiran Parameter Model Regresi Nonparametrik Spline Enam Variabel Prediktor

Model tersebut merupakan model dengan nilai GCV terkecil, didapatkan model terbaik yaitu model regresi nonparametrik spline dengan tiga titik knot dengan enam variabel prediktor.

$$\hat{y} = 8,556 + 0,001x_1 + 0,229(x_1 - 3223,22)_+^1 - 0,354(x_1 - 3556,90)_+^1 + 0,124(x_1 - 4224,24)_+^1 - 4,476x_3 + 68,324(x_3 - 6,77)_+^1 - 77,828(x_3 - 7,02)_+^1 + 15,036(x_3 - 7,53)_+^1 - 1,687x_4 - 1,603(x_4 - 11,95)_+^1 + 6,429(x_4 - 12,81)_+^1 - 2,88(x_4 - 14,55)_+^1 + 0,0138x_5 - 0,162(x_5 - 1416,33)_+^1 - 0,382(x_5 - 1465,31)_+^1 + 0,211(x_5 - 1563,27)_+^1 + 0,42x_6 - 1,338(x_6 - 27,47)_+^1 + 0,291(x_6 - 30,51)_+^1 + 1,255(x_6 - 36,60)_+^1 + 0,77x_7 - 4,469(x_7 - 49,25)_+^1 + 5,641(x_7 - 52,32)_+^1 - 1,924(x_7 - 58,45)_+^1$$

G. Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline

Seperti prosedur sebelumnya, akan dilakukan pengujian signifikansi parameter baik secara serentak maupun parsial.

Tabel 7. Analisis Ragam Uji Serentak Enam Prediktor

Sumber Variasi	db	JK	RJK	F	p-value
Regresi	24	925,694	38,58	4,731	0,00274
Residual	13	106,024	8,156		
Total	37	1031,99			

Pada tingkat signifikan 5%, p-value bernilai kurang dari α, sehingga tolak H₀. Hal ini menunjukkan bahwa minimal terdapat satu parameter yang tidak sama dengan nol atau secara bersama enam variabel prediktor berpengaruh secara signifikan terhadap angka morbiditas penduduk di Jawa Timur.

Tabel 8. Hasil Uji Individu Enam Prediktor

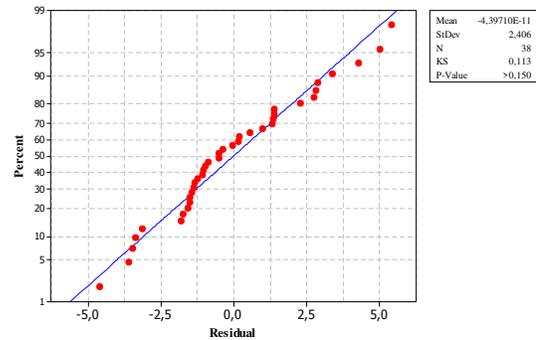
Variabel	Parameter	Estimasi	p-value	Ket.
X ₁	β ₀	8,56	0,000	Signifikan
	β ₁	0,00	0,739	Tidak Signifikan
	β ₂	0,23	0,001	Signifikan
	β ₃	-0,35	0,001	Signifikan
	β ₄	0,12	0,002	Signifikan
X ₃	β ₅	-4,48	0,040	Signifikan
	β ₆	68,32	0,000	Signifikan
	β ₇	-77,83	0,000	Signifikan
	β ₈	15,04	0,035	Signifikan
X ₄	β ₉	-1,69	0,019	Signifikan
	β ₁₀	-1,60	0,594	Tidak Signifikan
	β ₁₁	6,43	0,128	Tidak Signifikan
	β ₁₂	-2,88	0,140	Tidak Signifikan
X ₅	β ₁₃	0,01	0,044	Signifikan
	β ₁₄	0,16	0,184	Tidak Signifikan
	β ₁₅	-0,38	0,044	Signifikan
	β ₁₆	0,21	0,010	Signifikan
X ₆	β ₁₇	0,42	0,023	Signifikan
	β ₁₈	-1,34	0,186	Tidak Signifikan
	β ₁₉	0,30	0,814	Tidak Signifikan
	β ₂₀	1,25	0,054	Tidak Signifikan
X ₇	β ₂₁	0,77	0,008	Signifikan
	β ₂₂	-4,47	0,040	Signifikan
	β ₂₃	5,64	0,033	Signifikan
	β ₂₄	-1,92	0,019	Signifikan

Berdasarkan Tabel 8 diketahui bahwa keenam prediktor berpengaruh secara signifikan terhadap angka morbiditas di Jawa Timur. Variabel tersebut adalah kepadatan penduduk, rata-rata lama sekolah, persentase penduduk miskin, Upah Minimum Kabupaten (UMK), persentase rumah tangga *Open Defecation* (OD) dan persentase rumah tangga dengan jarak sumber air minum ke tempat penampungan kotoran lebih dari 10 meter.

H. Pengujian Asumsi Residual

1. Uji Normalitas Residual

Uji normalitas digunakan untuk mengetahui apakah residual telah mengikuti pola distribusi normal.



Gambar 3 Uji Normalitas Residual Kolmogorov-Smirnov

Berdasarkan Gambar 3 nilai KS yang diperoleh sebesar 0,113 dengan p-value lebih besar dari 0,150. Pada tingkat signifikan 5% didapatkan nilai KS tabel atau D_(1-α) sebesar 0,194, sehingga nilai KS_{hitung} kurang dari KS_{tabel} dan p-value lebih dari 0,05. Berdasarkan perumusan hipotesis pada BAB II, maka gagal tolak hipotesis awal, sehingga residual dari model regresi nonparametrik spline yang didapatkan telah berdistribusi normal.

2. Asumsi Identik

Berikut ini adalah hasil analisis ragam uji Glejser

Tabel 9. Analisis Ragam Uji Glejser

Sumber Variasi	db	JK	RJK	F Hitung	p-value
Regresi	24	40,57802	1,691	0,643	0,831
Residual	13	34,15664	2,627		
Total	37	74,73466			

Berdasarkan Tabel 9 didapatkan nilai F_{hitung} uji *Glejser* sebesar 0,643 dengan p-value sebesar 0,831. Pada tingkat signifikan 5%, maka p-value kurang dari α, sehingga gagal tolak H₀. Hal ini berarti bahwa residual memiliki varians yang homogen atau telah memenuhi asumsi identik.

I. Nilai Koefisien Determinasi (R²)

Berdasarkan perhitungan didapatkan nilai R² sebesar 89,72%. Hal ini berarti model regresi nonparametrik spline yang didapatkan mampu menjelaskan variabilitas angka morbiditas di Jawa Timur sebesar 89,72%. Nilai tersebut mendekati 100%, sehingga model sudah cukup baik.

J. Interpretasi Model Regresi Nonparametrik Spline

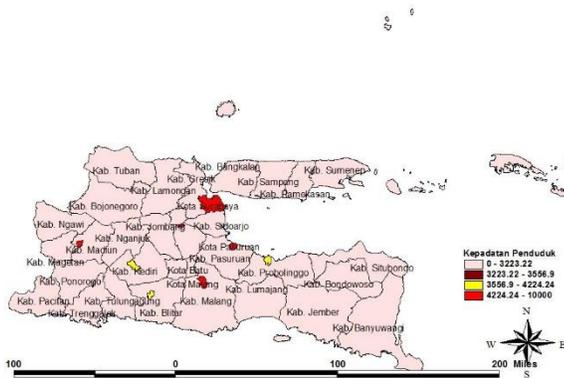
Berikut merupakan interpretasi dari model regresi nonparametrik spline yang didapatkan. Apabila X₃, X₄, X₅, X₆,

dan X_7 dianggap konstan, maka pengaruh kepadatan penduduk (X_1) terhadap angka morbiditas adalah

$$\hat{y} = 8,556 + 0,001x_1 + 0,229(x_1 - 3223,22)_+^1 - 0,354(x_1 - 3556,90)_+^1 + 0,124(x_1 - 4224,24)_+^1 + C$$

$$= \begin{cases} 8,556 + 0,001x_1 & , x_1 < 3223,22 \\ -729,561 + 0,23x_1 & , 3223,22 \leq x_1 < 3556,90 \\ 529,581 - 0,124x_1 & , 3556,90 \leq x_1 < 4224,24 \\ 5,775 + 0,0005x_1 & , x_1 \geq 4224,24 \end{cases}$$

Berdasarkan model yang diperoleh dapat diinterpretasikan bahwa apabila wilayah dengan kepadatan penduduk kurang dari 3223,23 penduduk per km² naik sebesar 10 penduduk per km², maka angka morbiditas akan naik sebesar 0,01 persen. Wilayah yang berada pada interval adalah seluruh wilayah kabupaten di Jawa Timur beserta Kota Batu. Apabila wilayah dengan kepadatan penduduk berada antara 3223,22 dan 3556,90 penduduk per km² naik sebesar 10 penduduk per km², maka angka morbiditas akan naik sebesar 2,3 persen. Apabila wilayah dengan kepadatan penduduk berada antara 3556,90 dan 4224,24 penduduk per km² naik sebesar 10 penduduk per km², maka angka morbiditas akan turun sebesar 1,24 persen. Wilayah yang termasuk pada kategori tersebut yaitu Kota Kediri, Kota Blitar, dan Kota Probolinggo. Pada daerah ini pengaruh kepadatan penduduk dengan angka morbiditas cenderung turun karena wilayah pada interval tersebut merupakan wilayah perkotaan dengan kepadatan penduduk yang tidak terlalu tinggi serta jumlah fasilitas kesehatan yang lebih banyak dibanding wilayah kabupaten. Apabila wilayah dengan kepadatan penduduk lebih dari 4224,24 penduduk per km² naik sebesar 10 penduduk per km², maka angka morbiditas akan naik sebesar 0,005 persen, wilayah yang termasuk yaitu Kota Malang, Kota Pasuruan, Kota Mojokerto, Kota Madiun, dan Kota Surabaya.



Gambar 5. Kepadatan Penduduk Berdasarkan Interval Titik Knot

Selanjutnya, interpretasi variabel rata-rata lama sekolah, persentase penduduk miskin, UMK, perserntase rumah tangga OD, dan persentase rumah tangga dengan jarak sumber air minum ke tempat pembuangan kotoran lebih dari 10 meter dilakukan dengan cara yang serupa seperti yang telah dijelaskan.

K. Rekomendasi Untuk Pemerintah Jawa Timur

World Health Organization (WHO) (2013) menyatakan bahwa target angka morbiditas kasar suatu wilayah sebesar

kurang dari 25% Berdasarkan Tabel 4.9 diketahui bahwa model optimis didapatkan dari nilai minimum seluruh variabel prediktor. Provinsi Jawa Timur akan mencapai target WHO apabila berada pada model optimis yaitu dengan angka morbiditas sebesar 15,16% dan 23,06%. Angka morbiditas yang sesuai target akan terpenuhi apabila seluruh variabel prediktor berada pada nilai minimum.

Tabel 9.

Skenario Model Angka Morbiditas							
X ₁	X ₃	X ₄	X ₅	X ₆	X ₇	Prediksi	Model
387	10,16	4,59	1000	1,6	23,19	15,159	Optimis
403	10,66	4,8	1000	2,9	32,48	23,059	
442	10,81	4,86	1000	3,13	37,5	27,129	
726	7,28	11,07	1120	20,04	63,87	35,460	Middle
731	7,41	11,28	1131	22,08	64,56	34,317	
750	7,49	11,53	1135	22,83	64,79	33,175	
6236	4,62	20,49	2190	51,11	88,56	35,551	Pesimis
7691	5,36	22,38	2195	65,13	89,7	42,374	
8562	5,62	25,8	2200	76,17	98,31	49,698	

V. KESIMPULAN DAN SARAN

Angka morbiditas tertinggi terletak pada Kota Kediri yaitu sebesar 41,24%, sedangkan angka morbiditas terendah terletak pada Kabupaten Sidoarjo yaitu sebesar 22,31%. Model regresi nonparametrik spline terbaik diperoleh dengan menggunakan tiga titik knot dengan nilai koefisien determinasi (R²) sebesar 89,69% dengan enam variabel yang berpengaruh secara signifikan yaitu kepadatan penduduk, rata-rata lama sekolah, persentase penduduk miskin, UMK, perserntase rumah tangga OD, dan persentase rumah tangga dengan jarak sumber air minum ke tempat pembuangan kotoran lebih dari 10 meter.

Saran untuk penelitian selanjutnya sebaiknya menambah jumlah variabel yang diduga berpengaruh terhadap angka morbiditas di Provinsi Jawa Timur, sehingga diharapkan akan mendapatkan model yang lebih sesuai. Bagi pemerintah, sebaiknya memperhatikan variabel yang mempunyai pengaruh cukup besar pada angka morbiditas di Provinsi Jawa Timur seperti persentase penduduk miskin dan persentase rumah tangga OD.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik Republik Indonesia. 2009. *Statistik Kesehatan Republik Indonesia Tahun 2009*. Badan Pusat Statistik Republik Indonesia.
- [2] Badan Pusat Statistik Jawa Timur. 2014. *Hasil Survei Sosial Ekonomi Nasional Tahun 2014 Provinsi Jawa Timur*. Badan Pusat Statistik Jawa Timur.
- [3] Sirusa BPS. 2016. *Angka Kesakitan (Morbiditas)*, diakses dari <https://sirusa.bps.go.id/> tanggal 10 Agustus 2016.
- [4] Kartasasmita, C.B. 2009. *Epidemiologi Tuberkulosis. Sari Pediatri*. Vol 11, No 2, Agustus 2009, 124-9.
- [5] Lembaga Demografi UI. 2010. *Dasar-Dasar Demografi*. Jakarta : Salemba Empat.
- [6] Hanum, D. 2013. *Faktor-faktor yang Mempengaruhi Morbiditas Penduduk Jawa Timur dengan Multivariate Geographically Weighted Regression (MGWR)*. Penelitian, Jurusan Statistika, ITS.
- [7] Ardhianti, N.L.P.D. 2013. *Peningkatan Angka Morbiditas di Provinsi Bali*. Thesis, Jurusan Ilmu Ekonomi, Universitas Udayana..
- [8] Fuhrer, R., Shipley, M.J., Chastang, J.F., Schmaus, A., Niedhammer, I., Stansfeld, S.A., Goldberg, M. dan Marmot, M.G. 2002. *Socioeconomic Position, Health, and Possible Explanations: A Tale of Two Cohorts. American Journal of Public Health*, Vol 92 No 8 Agustus 2002, 1290-1294.

- [9] Arola, H., Pitkanen, M., Nygrad, C.H, Huhtala, H., dan Manka, M.L. 2003. *The Connection Between Age, Job Controll, and Sicknes Absences Among Finishing Food Workers. Occupational Medicine*, Vol 53 No 3: 229-230.
- [10] Budiantara, I.N. 2009. *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.
- [11] Ramadhani, R.K. 2016. *Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Pengangguran Terbuka Menggunakan Regresi Nonparametrik Spline di Indonesia*. Penelitian, Jurusan Statistika, ITS.
- [12] Eubank, R. 1988. *Spline Smoothing and Nonparametric Regression*. Marcel Dekker Inc. New York.
- [13] Budiantara, I.N. 2000. Metode U, GLM, CV, dan GCV dalam Regresi Nonparametrik Spline. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6, 41-45.