

PEMODELAN PREVALENSI STUNTING INDONESIA MENGUNAKAN REGRESI NONPARAMETRIK *SPLINE* *TRUNCATED*

Modeling of Indonesian Stunting Prevalence Using Spline Truncated Non-parametric Regression

Hartina Husain*, Ardiana Fatma Dewi**, Ayunytyah Eka Wardani***

*Prodi Sains Data Institut Teknologi Bacharuddin Jusuf Habibie, E-mail: hartinahusain@ith.ac.id

**Prodi Tadris Matematika Institut Agama Islam Negeri Kediri

***Prodi Pengembangan Masyarakat Islam Institut Agama Islam Negeri Parepare

ABSTRAK

Stunting merupakan keadaan gagal tumbuh pada anak yang terjadi disebabkan terjadi kekurangan gizi kronik sehingga tinggi badan anak lebih pendek jika dibandingkan dengan seusianya. Tujuan penelitian ini untuk membuat pemodelan faktor-faktor yang mempengaruhi prevalensi stunting di Indonesia berdasarkan studi literatur yang dilakukan dengan menggunakan metode regresi nonparametrik *spline truncated*. Data yang digunakan berupa data sekunder mengenai prevalensi stunting dan beberapa faktor-faktor yang diduga mempengaruhinya yaitu Persentase penduduk dengan jaminan kesehatan, Persentase penduduk merokok sebulan terakhir (Usia ≥ 15 Tahun), Rata-rata usia perkawinan pertama ibu, Persentase ibu yang melahirkan Anak Lahir Hidup (ALH) dua tahun terakhir dan ALH terakhir yang dilahirkan dengan Berat Badan Lahir Rendah (BBLR), dan Persentase anak umur 0-23 bulan yang pernah diberi ASI. Data bersumber dari publikasi Kementerian Kesehatan dan Badan Pusat Statistik (BPS) tahun 2022. Hasil penelitian didapatkan model terbaik yaitu regresi nonparametrik *spline truncated* 3 knot dengan kriteria kebaikan model yaitu $GCV = 22.16$, $R^2 = 88,49\%$, dan $MSE = 3.24$.

Kata Kunci : *GCV, Spline truncated, Stunting*

ABSTRACT

Stunting is a condition of failure to thrive in children that occurs due to chronic malnutrition so that children's height is shorter when compared to their age. The study aimed to model the factors that influence the prevalence of stunting in Indonesia based on a literature study conducted using the spline truncated nonparametric regression method. The data used is in the form of secondary data regarding the prevalence of stunting and several factors that are thought to influence it, namely the percentage of the population with health insurance, the percentage of the population smoking in the last month (Age ≥ 15 years), the average age of first marriage to mothers, the percentage of mothers who gave birth to live births in the last two years and the last birth to live who were born with low birth weight (LBW), and the percentage of children aged 0-23 months who have ever been breastfed. The data is sourced from the publication of the Ministry of Health and the Central Bureau of Statistics (BPS) in 2022. The results of the study obtained the best model, namely the nonparametric spline truncated regression of 3 knots with the goodness of the model are $GCV = 22.16$, $R^2 = 88,49\%$ and $MSE = 3.24$.

Keywords : GCV, Spline truncated, Stunting

I. PENDAHULUAN

Stunting merupakan suatu permasalahan serius yang sedang dihadapi dunia (Rifada, 2021), khususnya negara miskin dan negara berkembang seperti di negara Indonesia (Ashari 2023). Stunting menjadi salah satu fokus pemerintah Indonesia dalam menyusun strategi nasional untuk

pengecahan stunting. Stunting merupakan kondisi kurangnya tinggi badan pada anak yang disebabkan kekurangan gizi dan berlangsung dalam jangka waktu yang lama. Hal ini menyebabkan tinggi badan anak yang lebih rendah jika dibandingkan dengan usia seharusnya. Di Indonesia, penurunan prevalensi stunting diupayakan untuk sejalan dengan target skala global, yaitu target dari *World Health Assembly* (WHA) yang bertujuan untuk mengurangi angka prevalensi stunting hingga 40% di tahun 2025. Sejalan dengan itu, target Tujuan Pembangunan Berkelanjutan atau disebut juga *Sustainable Development Goals* (SDGs) adalah mencapai eliminasi semua bentuk kekurangan gizi pada tahun 2030. Untuk mencapai tujuan tersebut, perlu dilakukan upaya dalam mempercepat penurunan stunting kondisi saat ini sehingga prevalensi stunting pada balita dapat menurun menjadi 19,4% di tahun 2024.

Berdasarkan data Kementerian Kesehatan angka stunting mengalami penurunan dari tahun 2021 sebesar 24,4% turun di tahun 2022 menjadi 21,6% . Pemerintah menyampaikan melalui Menteri Koordinator Bidang Pembangunan Manusia dan Kebudayaan (Menko PMK) menargetkan setiap tahunnya angka stunting turun sebesar 3%. Namun kenyataannya, pada saat ini masalah stunting masih menjadi fokus pemerintah Indonesia. Banyak faktor yang menyebabkan masalah stunting (Brilliantita, 2022). Dampak stunting pada anak dapat berpengaruh dalam waktu dekat maupun di masa yang akan datang. Dampak jangka pendek termasuk munculnya gangguan ataupun kerusakan pada perkembangan otak anak, penurunan tingkat kecerdasan (IQ), dan sistem kekebalan tubuh yang lemah sehingga berakibat rentan terserang infeksi atau penyakit (Purnamasari, 2022). Untuk itu diperlukan suatu penanganan stunting khususnya di Indonesia melalui pemahaman faktor-faktor apa saja yang dapat mempengaruhi prevalensi stunting.

Penelitian mengenai stunting sudah sering dilakukan oleh peneliti lain sebelumnya dengan bermacam-macam metode statistik yang digunakan. Hamsah (2023) menyatakan bahwa banyak faktor yang dimungkinkan akan berpengaruh terhadap stunting di Indonesia. Salah satunya penelitian mengenai stunting dengan pemodelan *Geographically Weighted Regression* (GWR) pada kasus stunting di Provinsi NTT (Bele, 2022). Berdasarkan penelitian tersebut terdapat beberapa faktor berpengaruh signifikan pada stunting di Provinsi NTT yaitu persentase ibu hamil yang berisiko Kekurangan Energi Kronik (KEK), persentase bayi mendapat imunisasi lengkap, persentase perempuan yang tamat SMA/SMK/MA/Paket C, persentase penduduk miskin, persentase bayi mendapat ASI eksklusif, dan persentase perempuan bawah umur yang pernah kawin.

Penelitian lain mengenai stunting juga pernah dilakukan dengan metode regresi nonparametrik *spline truncated* pada tahun 2016 (Anindita, 2016). Hasil yang diperoleh dari penelitian tersebut model terbaik yang diperoleh dengan metode regresi non-parametrik *spline truncated* dengan kombinasi jumlah knot 3,1,2,2,3 pada masing-masing variabel yang

mempengaruhinya. Kelima variabel independen tersebut antara lain persentase balita mendapat imunisasi lengkap, persentase ibu hamil risiko KEK, persentase balita mendapat asi selama 6 bulan, persentase rumah tangga memiliki sanitasi layak dan persentase bayi lahir mendapat IMD yang secara keseluruhan memberikan berpengaruh secara signifikan terhadap model. Sementara itu, koefisien determinasi yang diperoleh dari model tersebut sebesar 80,77%.

Indonesia sebagai salah satu negara dengan kasus stunting yang masih tinggi menunjukkan urgensinya dalam menangani hal tersebut. Berbagai upaya dilakukan guna menurunkan angka prevalensi stunting termasuk dengan semakin banyaknya riset mengenai stunting. Metode regresi nonparametrik sesuai jika digunakan pada data *cross section* yaitu data yang disusun berdasarkan individu dalam penelitian ini berupa provinsi yang ada di Indonesia. Adapun estimator *spline truncated* digunakan dalam menganalisis data dengan pola yang tidak dapat diketahui bentuk hubungannya serta berubah-ubah pada sub-sub interval tertentu. Berdasarkan observasi awal berupa grafik pencar pada data prevalensi stunting didapatkan pola yang berubah-ubah sehingga pada penelitian ini akan dilakukan pemodelan faktor-faktor yang mempengaruhi prevalensi stunting di Indonesia berdasarkan studi literatur yang dilakukan dengan menggunakan metode regresi nonparametrik *spline truncated*.

II. TINJAUAN PUSTAKA

II.1 Analisis Regresi Nonparametrik

Analisis regresi merupakan metode statistik untuk mengetahui keeratan hubungan antara variabel respon dan satu atau lebih variabel prediktor. Ada beberapa pendekatan dalam regresi yang digunakan antara lain parametrik, nonparametrik, dan semiparametrik. Jika pola data diketahui linier, kuadrat, atau kubik, maka regresi metode yang digunakan adalah metode regresi parametrik (Hardle, 1990). Namun pada kenyataannya, ditemukan bahwa tidak seluruh data membentuk suatu pola tertentu (Husain, 2021). Apabila hubungan antara variabel respon dan prediktor tidak diketahui polanya, maka dalam memodelkan hubungan antara keduanya menggunakan metode regresi nonparametrik. Adapun model regresi nonparametrik secara umum dapat dituliskan sebagai berikut ini:

$$y_j = f(x_j) + \varepsilon_j, \quad j = 1, 2, 3 \dots n \quad (1)$$

dengan,

y_j : Variabel terikat/respon

$f(x_j)$: Fungsi pada regresi nonparametrik variabel ke- p dan observasi ke- j

ε_j : Error yang memiliki asumsi *IID* $(0, \sigma^2)$

Pendekatan regresi nonparametrik membebaskan peneliti untuk memilih estimator yang sesuai dengan pola data yang tersedia (Dewi, 2023). Pada regresi nonparametrik ini terdapat beberapa pendekatan estimator yang sering digunakan yaitu Kernel, Polinomial Lokal, Spline,

Histogram, Deret Fourier, Deret Ortogonal, k-NN, MARS, Neural Network (NN), Wavelets, dan pendekatan lain sebagainya.

II.2 Analisis Regresi Nonparametrik *Spline truncated*

Regresi nonparametrik dengan estimator spline merupakan teknik estimasi yang sesuai digunakan dalam menangani data yang memiliki pola perilaku berubah-ubah pada sub-sub interval tertentu. Model regresi nonparametrik spline dapat dituliskan seperti pada persamaan (1) dimana $f(x_j)$ merupakan kurva regresi yang didekati dengan fungsi spline yang memiliki orde/tingkatan p . Adapun spline sangat tergantung pada titik knot yaitu letak titik perubahan pola pada data (Eubank, 1999). Maka pada persamaan regresi *spline truncated* diberikan titik knot K_1, K_2, \dots, K_m . Adapun $f(x_j)$ dapat diberikan oleh persamaan berikut:

$$f(x_j) = \sum_{t=0}^p \delta_0 x_j^t + \sum_{t=0}^m \delta_{p+t} (x_j - K_t)_+^p \quad (2)$$

Substitusikan persamaan (1) ke persamaan (2) sehingga diperoleh persamaan regresi nonparametrik spline seperti pada persamaan (3) sebagai berikut:

$$y_j = \sum_{t=0}^p \delta_0 x_j^t + \sum_{t=0}^m \delta_{p+t} (x_j - K_t)_+^p + \varepsilon_j \quad (3)$$

Dengan fungsi $(x_j - K_t)_+^p$ merupakan fungsi truncated (potongan) yang diberikan oleh:

$$(x_j - K_t)_+^p = \{(x_j - K_t)^p, x_j \geq K_t \quad 0, x_i < K_t \quad (4)$$

dimana p yaitu orde *spline truncated* dan K yaitu knot yang terpilih.

II.3 Pemilihan Titik Knot Optimal

Titik knot yaitu titik perpaduan bersama yang terjadi pola perubahan perilaku d dari suatu fungsi pada interval yang berbeda. Pemilihan titik knot pada regresi nonparametrik penting digunakan karena untuk menentukan model terbaik berdasarkan nilai tertentu. Pada beberapa penelitian tentang regresi nonparametrik metode yang seringkali digunakan dalam pemilihan titik knot yang optimal yaitu dengan *Generalized Cross Validation* (GCV). Berikut merupakan fungsi GCV pada regresi nonparametric secara umum dapat dituliskan sebagai berikut (Eubank, 1999):

$$GCV = \frac{MSE(K_1, K_2, \dots, K_m)}{(n^{-1} \text{trace}[I - A(K_1, K_2, K_3, \dots, K_m)])^2} \quad (5)$$

dengan

$$MSE(K_1, K_2, K_3 \dots, K_m) = n^{-1} \sum_{j=1}^n (y_j - \hat{f}(x_j))^2$$

Dimana $K_1, K_2, K_3 \dots, K_m$ yaitu titik knot awal hingga knot ke- m , $GCV(K_1, K_2, K_3 \dots, K_m)$ yaitu vektor berisikan nilai GCV dari titik-titik knot, dan n yaitu banyaknya observasi.

II.4 Prevalensi Stunting

Stunting adalah keadaan ketidakmampuan balita untuk tumbuh secara optimal disebabkan karena kekurangan gizi yang berlangsung dalam jangka waktu panjang. Dampaknya mengakibatkan anak menjadi terlalu pendek dibandingkan dengan standar yang telah ditetapkan *World Health Organization* (WHO) (Kemenkes RI, 2013). Adapun masalah pertumbuhan pendek pada balita mengindikasikan adanya gizi kronis yang dialami oleh ibu atau calon ibu selama masa kehamilan, masa janin, ataupun masa bayi atau balita serta masalah lainnya yang turut berperan secara tidak langsung terhadap kesehatan (Kemenkes RI, 2016).

Balita perlu mendapatkan perhatian khusus dalam berbagai aspek untuk mencegah stunting yang dapat menyebabkan terhambatnya perkembangan mental, pertumbuhan fisik, dan status kesehatan pada anak (Setiawan et al., 2018). Pertumbuhan dan perkembangan balita dapat dilihat dengan beberapa indikator pada status gizi. Suatu standar antropometri yang menjadi penilaian status gizi anak yaitu pendek atau sangat pendek merupakan keadaan yang berdasarkan indeks panjang badan menurut tinggi badan atau umur atau dikenal dengan istilah stunted/pendek atau severely stunted/sangat pendek. Skor Z pada kategori pendek yaitu -3 Standar deviasi hingga <-2 Standar deviasi serta sangat pendek yaitu <-3 standar deviasi (Kemenkes RI, 2018).

III. METODE PENELITIAN

Penelitian ini menggunakan pendekatan secara kuantitatif yaitu analisis menggunakan metode statistika untuk mengidentifikasi pola hubungan maupun menarik kesimpulan dari hasil analisis data. Data yang digunakan yaitu data sekunder periode tahun 2022 yang diperoleh dari publikasi Kementerian Kesehatan dan Badan Pusat Statistik (BPS). Tipe data yaitu *cross section* yang terdiri atas 34 provinsi di Indonesia. Adapun pada penelitian menggunakan enam variabel yang terdiri atas satu variabel respon (Y) dan lima variabel prediktor (X) yang tertera pada Tabel 1 berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
Y	Prevalensi stunting
X_1	Persentase penduduk dengan jaminan kesehatan
X_2	Persentase penduduk merokok sebulan terakhir (Usia ≥ 15 Tahun)
X_3	Rata-rata usia perkawinan pertama ibu
X_4	Persentase ibu yang melahirkan Anak Lahir Hidup (ALH) dua tahun terakhir dan ALH terakhir yang dilahirkan dengan Berat Badan Lahir Rendah (BBLR).
X_5	Persentase anak umur 0-23 bulan yang pernah diberi ASI

Selanjutnya tahapan-tahapan dalam analisis data pada penelitian ini yaitu sebagai berikut:

1. Melakukan analisis statistika secara deskriptif pada data penelitian

2. Membuat grafik pencar dari masing-masing variabel prediktor terhadap variabel respon untuk melihat pola hubungannya
3. Membuat pemodelan data dengan regresi nonparametrik *spline truncated* linear dengan satu, dua, dan tiga titik knot.
4. Memilih titik knot optimal didasarkan pada nilai GCV minimum.
5. Memperoleh model regresi *spline truncated* dengan titik knot optimum
6. Menghitung nilai koefisien determinasi model
7. Melakukan interpretasi model

IV. ANALISIS DAN PEMBAHASAN

IV.1 Analisis Statistika Deskriptif

Sebagai tahap awal, dilakukan analisis secara deskriptif terhadap masing-masing variabel yang digunakan dalam penelitian. Hal ini bertujuan untuk menggambarkan karakteristik data secara ringkas agar didapatkan pemahaman yang lebih baik mengenai data tersebut.

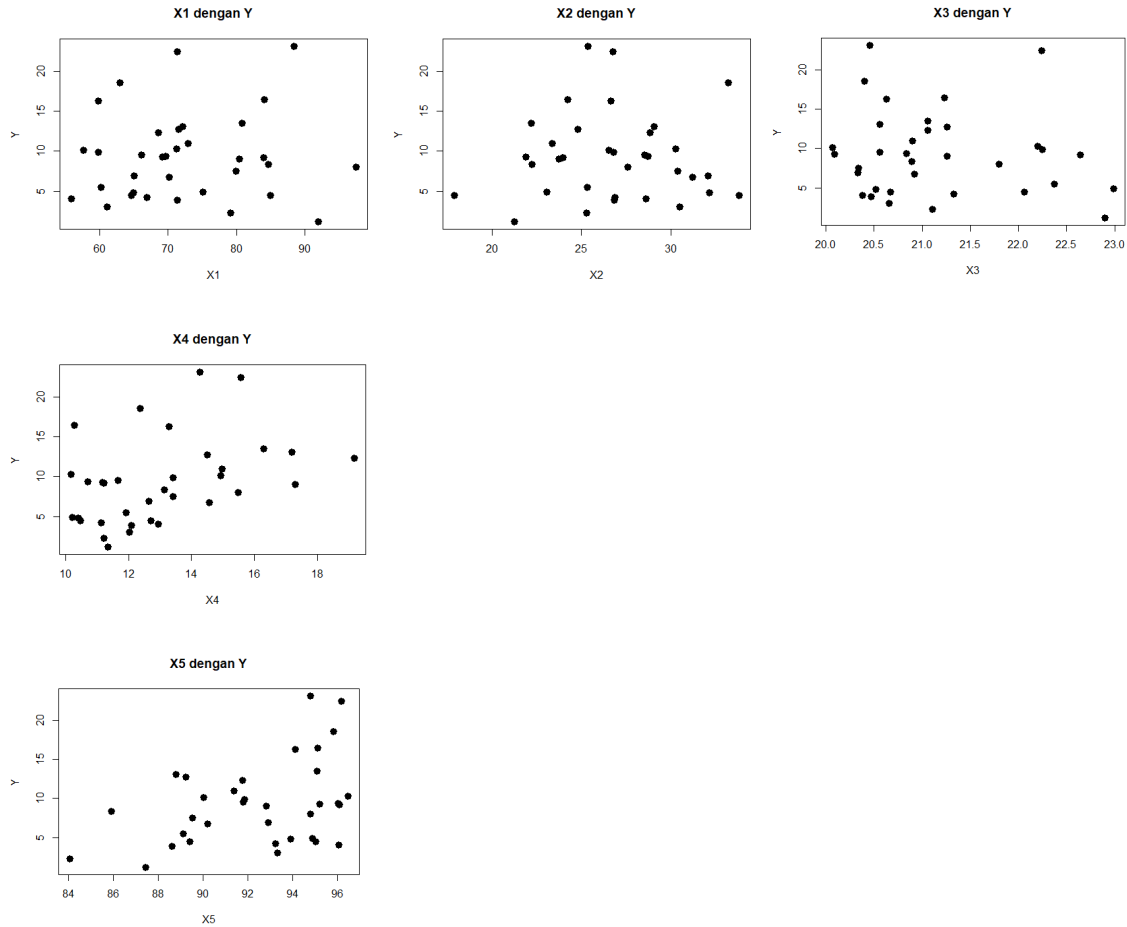
Tabel 2. Analisis Statistika Deskriptif

Variabel	Minimum	Maximum	Mean	Std. Deviation
y	1.20	23.10	9.31	5.38
x_1	55.91	97.50	72.49	10.35
x_2	17.91	33.81	26.76	3.73
x_3	20.07	22.99	21.16	0.83
x_4	10.17	19.19	13.06	2.31
x_5	84.06	96.48	92.38	3.26

Tabel 2 menunjukkan nilai minimum, maksimum, mean/rata-rata dan standar deviasi dari masing-masing variabel. Pada variabel y diperoleh nilai prevalensi stunting terkecil yaitu 1,20 berada di Provinsi DKI Jakarta dan terbesar yaitu 23,10 berada di Provinsi Sulawesi Barat. Sementara itu rata-rata dan standar deviasinya masing-masing diperoleh 9,31 dan 5,38. Hal ini menunjukkan bahwa rata-rata prevalensi stunting pada anak di Indonesia sebesar 9,31% . Sementara itu, nilai standar deviasi yang tinggi menunjukkan data prevalensi stunting pada populasi anak di Indonesia cukup beragam pada setiap provinsi.

IV.2 Grafik Pencar

Grafik pencar (*scatter plot*) merupakan diagram yang berisikan titik-titik data yang tersebar dalam bidang kartesius yang bertujuan untuk melihat pola hubungan antara dua variabel. Pada penelitian ini dilakukan pada setiap variabel prediktor terhadap variabel respon.



Gambar 1. Grafik Pencar antara Prevalensi stunting dengan 5 Variabel yang diduga mempengaruhinya

Gambar 1 menunjukkan pola hubungan kelima variabel prediktor terhadap prevalensi stunting tidak memiliki pola tertentu. Secara visual, tidak terlihat pola data yang membentuk secara linear ataupun pola tertentu sehingga kurang tepat menggunakan pendekatan secara parametrik. Adapun pada beberapa pola hubungan data terlihat berubah-ubah pada sub-sub interval tertentu seperti pada variabel Rata-rata usia perkawinan pertama ibu (X_3) berubah di sekitar angka 21 sampai 21,5, Persentase ibu yang melahirkan Anak Lahir Hidup (ALH) dalam dua tahun terakhir dan ALH terakhir yang dilahirkan dengan Berat Badan Lahir Rendah (BBLR) (X_4) berubah di sekitar angka 14 sampai 16, dan Persentase anak umur 0-23 bulan yang pernah diberi ASI (X_5) berubah di sekitar angka 92 dan 94. Dengan demikian dapat digunakan pendekatan nonparametrik *spline truncated* linear dalam memodelkan prevalensi stunting di Indonesia.

IV.3 Pemodelan dengan Regresi Nonparametrik *Spline truncated*

Selanjutnya dilakukan pemodelan prevalensi stunting di Indonesia tahun 2022 menggunakan regresi nonparametrik *spline truncated* linear dengan 1, 2 dan 3 titik knot. Kemudian dari ketiga pemodelan tersebut digunakan model terbaik dengan kriteria nilai GCV terkecil

Tabel 3. Perbandingan nilai GCV model satu titik knot

Titik Knot					GCV
K_1	K_2	K_3	K_4	K_5	
57,89	18,67	20,21	10,60	84,65	29,21*
87,60	30,02	22,29	17,04	93,52	29,45
89,58	30,78	22,43	17,47	94,11	29,82
85,62	29,27	22,15	16,61	92,93	30,48
95,52	33,05	22,85	18,76	95,88	30,64

Tabel 3 menunjukkan titik knot pada masing-masing variabel prediktor dengan satu titik knot berdasarkan lima urutan nilai GCV terkecil. Terlihat nilai GCV paling kecil yaitu 29,21 dengan letak titik knot pada masing-masing prediktor yaitu 57,89; 18,67; 20,21; 10,60; dan 84,65. Selanjutnya, dengan melakukan substitusi estimasi parameter-parameter dan letak titik knot didapatkan model regresi nonparametrik *spline truncated* satu knot sebagai berikut :

$$\hat{y} = \hat{\delta}_0 + \hat{\delta}_1 x_1 + \hat{\delta}_2(x_1 - K_1) + \hat{\delta}_3 x_2 + \hat{\delta}_4(x_2 - K_2) + \hat{\delta}_5 x_3 + \hat{\delta}_6(x_3 - K_3) + \hat{\delta}_7 x_4 + \hat{\delta}_8(x_4 - K_4) + \hat{\delta}_9 x_5 + \hat{\delta}_{10}(x_5 - K_5)$$

$$\hat{y} = 0,46 + 4,74 x_1 + 1,60(x_1 - 57,89) + 12,85 x_2 - 4,49(x_2 - 18,67) - 6,05 x_3 - 4,77(x_3 - 20,21) - 2,06 x_4 - 14,69(x_4 - 10,60) + 5,63 x_5 + 7,02(x_5 - 84,65)$$

Tabel 4. Perbandingan nilai GCV model dua titik knot

Titik Knot										GCV
K_{11}	K_{12}	K_{21}	K_{22}	K_{31}	K_{32}	K_{41}	K_{42}	K_{51}	K_{52}	
59,87	75,71	19,42	25,48	20,35	21,46	11,02	14,47	85,24	89,97	28,31*
57,89	75,71	18,67	25,48	20,21	21,46	10,60	14,47	84,65	89,97	28,82
61,85	73,73	20,18	24,72	20,49	21,32	11,46	14,03	85,83	89,38	28,89
59,87	73,73	19,42	24,72	20,35	21,32	11,03	14,03	85,24	89,38	28,92
61,85	75,71	20,18	25,48	20,49	21,46	11,46	14,47	85,83	89,97	29,36

Tabel 4 menunjukkan titik knot pada masing-masing variabel prediktor dengan dua titik knot berdasarkan lima urutan nilai GCV terkecil. Terlihat nilai GCV paling kecil yaitu 29,21 dengan letak titik knot pada prediktor X_1 yaitu 59,87 dan 75,71. Kemudian letak titik knot pada prediktor X_2 , X_3 , X_4 dan X_5 yaitu masing-masing 19,42; 25,48, 20,35; 21,46, 11,02; 14,47; dan 85,24; 89,97. Selanjutnya, dengan melakukan substitusi estimasi parameter-parameter dan letak titik knot didapatkan model regresi nonparametrik *spline truncated* dua knot sebagai berikut :

$$\hat{y} = \hat{\delta}_0 + \hat{\delta}_1x_1 + \hat{\delta}_2(x_1 - K_{11}) + \hat{\delta}_3(x_1 - K_{12}) + \hat{\delta}_4x_2 + \hat{\delta}_5(x_2 - K_{21}) + \hat{\delta}_6(x_2 - K_{22}) + \hat{\delta}_7x_3 + \hat{\delta}_8(x_3 - K_{31}) + \hat{\delta}_9(x_3 - K_{32}) + \hat{\delta}_{10}x_4 + \hat{\delta}_{11}(x_4 - K_{41}) + \hat{\delta}_{12}(x_4 - K_{42}) + \hat{\delta}_{13}x_5 + \hat{\delta}_{14}(x_5 - K_{51}) + \hat{\delta}_{15}(x_5 - K_{52})$$

$$\hat{y} = -428,87 + 3,53x_1 - 8,12(x_1 - 59,87) - 5,05(x_1 - 75,71) - 4,67x_2 + 6,26(x_2 - 19,42) - 3,60(x_2 - 25,48) - 0,02x_3 + 9,84(x_3 - 20,35) - 2,94(x_3 - 21,46) - 0,57x_4 + 6,97(x_4 - 11,02) + 7,64(x_4 - 14,47) - 3,06x_5 - 8,13(x_5 - 85,24) + 3,61(x_5 - 89,97)$$

Tabel 5. Perbandingan nilai GCV model tiga titik knot

Titik Knot															GCV
K_{11}	K_{12}	K_{13}	K_{21}	K_{22}	K_{23}	K_{31}	K_{32}	K_{33}	K_{41}	K_{42}	K_{43}	K_{51}	K_{52}	K_{53}	
61,85	73,73	91,56	20,18	24,72	31,54	20,49	21,32	22,57	11,46	14,03	17,90	85,83	89,38	94,71	22,16*
59,87	77,69	89,58	19,42	26,24	30,78	20,34	21,60	22,43	11,03	14,90	17,47	85,24	90,57	94,11	23,50
59,87	75,71	91,56	19,42	25,48	31,54	20,34	21,46	22,57	11,03	14,47	17,90	85,24	89,97	94,71	23,76
61,85	73,73	89,58	20,18	24,72	30,78	20,48	21,32	22,43	11,46	14,03	17,47	85,83	89,38	94,11	23,83
57,89	81,66	87,60	18,67	27,75	30,02	20,21	21,88	22,29	10,60	15,75	17,04	84,65	91,75	93,52	24,13

Tabel 5 menunjukkan titik knot pada masing-masing variabel prediktor dengan tiga titik knot berdasarkan lima urutan nilai GCV terkecil. Terlihat nilai GCV paling kecil yaitu 22,16 dengan letak titik knot pada prediktor X_1 yaitu 61,85; 73,73; dan 91,56. Kemudian letak titik knot pada prediktor X_2 , X_3 , X_4 dan X_5 yaitu masing-masing 20,18; 24,72; 31,54, 20,49; 21,32; 22,57, 11,46; 14,03; 17,90; dan 85,83; 89,38; 94,71. Selanjutnya, dengan melakukan substitusi estimasi parameter-parameter dan letak titik knot didapatkan model regresi nonparametrik *spline truncated* tiga knot sebagai berikut :

$$\hat{y} = \hat{\delta}_0 + \hat{\delta}_1x_1 + \hat{\delta}_2(x_1 - K_{11}) + \hat{\delta}_3(x_1 - K_{12}) + \hat{\delta}_4(x_1 - K_{13}) + \hat{\delta}_5x_2 + \hat{\delta}_6(x_2 - K_{21}) + \hat{\delta}_7(x_2 - K_{22}) + \hat{\delta}_8(x_2 - K_{23}) + \hat{\delta}_9x_3 + \hat{\delta}_{10}(x_3 - K_{31}) + \hat{\delta}_{11}(x_3 - K_{32}) + \hat{\delta}_{12}(x_3 - K_{33}) + \hat{\delta}_{13}x_4 + \hat{\delta}_{14}(x_4 - K_{41}) + \hat{\delta}_{15}(x_4 - K_{42}) + \hat{\delta}_{16}(x_4 - K_{43}) + \hat{\delta}_{17}x_5 + \hat{\delta}_{18}(x_5 - K_{51}) + \hat{\delta}_{19}(x_5 - K_{52}) + \hat{\delta}_{20}(x_5 - K_{53})$$

$$\hat{y} = -9,28 + 4,86x_1 - 2,37(x_1 - 61,85) - 0,66(x_1 - 73,73) - 7,18(x_1 - 91,56) - 1,61x_2 - 6,39(x_2 - 20,18) + 2,13(x_2 - 24,72) - 1,61(x_2 - 31,54) + 4,38x_3 - 3,95(x_3 - 20,49) + 0,03(x_3 - 21,32) - 7,47(x_3 - 22,57) + 6,65x_4 + 13,17(x_4 - 11,46) + 13,07(x_4 - 14,03) - 5,95(x_4 - 17,90) + 0,73x_5 + 2,02(x_5 - 85,83) - 0,74(x_5 - 89,38) + 9,36(x_5 - 94,71)$$

IV.4 Pemilihan Model Terbaik

Dari ketiga pemodelan yang telah dilakukan untuk 1,2, dan 3 titik knot kemudian dibandingkan kembali untuk didapatkan model terbaik. Kriteria pemilihannya didasarkan pada nilai GCV terkecil.

Tabel 6. Perbandingan model dengan 1,2, dan 3 titik knot

Titik Knot	GCV	R^2	MSE
1 titik knot	29,21	48,26%	14,56
2 titik knot	28,31	71,80%	7,94
3 titik knot	22,16	88,49%	3,24

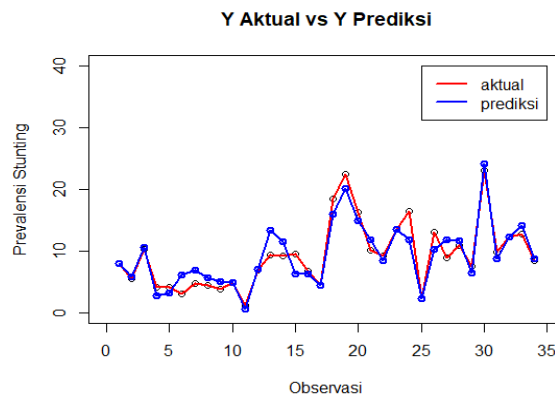
Tabel 6 menunjukkan perbandingan nilai GCV dari model regresi nonparametrik menggunakan 1, 2, dan 3 titik knot dimana model dengan 3 titik knot diperoleh GCV 22,16 lebih rendah dibandingkan kedua model lainnya. Dengan demikian, model tersebut menjadi model terbaik dalam memodelkan prevalensi stunting di Indonesia. Adapun nilai koefisien determinasi (R^2) menunjukkan bahwa sebanyak 88,49% faktor-faktor yang mempengaruhi prevalensi stunting yaitu persentase penduduk dengan jaminan kesehatan, persentase penduduk merokok sebulan terakhir (Usia \geq 15 Tahun), rata-rata usia perkawinan pertama ibu, persentase ibu yang melahirkan Anak Lahir Hidup (ALH) dua tahun terakhir dan ALH terakhir yang dilahirkan dengan Berat Badan Lahir Rendah (BBLR), dan persentase anak umur 0-23 bulan yang pernah diberi ASI. Sementara sisanya yaitu 11,51% dipengaruhi oleh faktor-faktor lainnya. Sementara itu, nilai *Mean Square Error* (MSE) model 3 titik knot diperoleh 3,24 yang nilainya terkecil dibandingkan model dengan 1 titik knot ataupun 2 knot. Hal ini menunjukkan kinerja model lebih baik dengan 3 titik knot dalam memprediksi nilai prevalensi stunting di Indonesia.

Tabel 7. Perbandingan nilai aktual dan nilai prediksi model terbaik

Provinsi	y	\hat{y}	Provinsi	y	\hat{y}
Aceh	8,0	8,03	Nusa Tenggara Barat	18,5	15,99
Sumatera Utara	5,5	5,85	Nusa Tenggara Timur	22,4	20,19
Sumatera Barat	10,3	10,57	Kalimantan Barat	16,3	14,87
Riau	4,2	2,8	Kalimantan Tengah	10,1	11,83
Jambi	4,1	3,27	Kalimantan Selatan	9,3	8,45
Sumatera Selatan	3,1	6,14	Kalimantan Timur	13,5	13,47
Bengkulu	4,8	6,9	Kalimantan Utara	16,4	11,88
Lampung	4,5	5,71	Sulawesi Utara	2,3	2,3
Kep. Bangka Belitung	3,9	5,08	Sulawesi Tengah	13,1	10,36
Kep. Riau	4,9	4,91	Sulawesi Selatan	9	11,82
DKI Jakarta	1,2	0,69	Sulawesi Tenggara	11	11,69
Jawa Barat	6,9	7,11	Gorontalo	7,5	6,45
Jawa Tengah	9,4	13,33	Sulawesi Barat	23,1	24,05
DI Yogyakarta	9,2	11,6	Maluku	9,9	8,85
Jawa Timur	9,5	6,27	Maluku Utara	12,3	12,3

Banten	6,8	6,39	Papua Barat	12,7	14,13
Bali	4,5	4,5	Papua	8,4	8,79

Tabel 7 menunjukkan hasil perbandingan antara nilai prevalensi stunting aktual (y) dan hasil prediksi (\hat{y}) menggunakan model regresi nonparametrik *spline truncated* dengan tiga titik knot. Dari hasil perbandingan, terlihat bahwa pada beberapa provinsi menunjukkan selisih yang relatif kecil antara nilai aktual dan prediksi seperti pada provinsi Aceh, Sumatera Utara, Sumatera Utara, Kep.Riau, Kalimantan Timur, Sulawesi Tenggara, dan Papua. Sementara itu, pada Provinsi Bali, Sulawesi Utara, dan Maluku Utara didapatkan hasil yang sama antara nilai prevalensi stunting aktual dan prediksi. Pemahaman lebih lanjut dapat dilihat melalui Gambar 2.



Gambar 2. Plot data y dan \hat{y}

Gambar 2 menunjukkan plot perbandingan data aktual (y) dan data hasil prediksi (\hat{y}) dari prevalensi stunting di Indonesia. Terlihat bahwa antara plot \hat{y} (berwarna biru) cenderung mengikuti plot y (berwarna merah) sehingga pemodelan yang didapatkan mampu memprediksi data aktual dari prevalensi stunting di Indonesia.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian sebelumnya, diperoleh bahwa model terbaik dalam memodelkan prevalensi stunting di Indonesia yaitu model regresi nonparametrik *spline truncated* dengan tiga titik knot. Adapun persamaannya yaitu sebagai berikut:

$$\begin{aligned} \hat{y} = & -9,28 + 4,86x_1 - 2,37(x_1 - 61,85) - 0,66(x_1 - 73,73) - 7,18(x_1 - 91,56) - 1,61x_2 \\ & - 6,39(x_2 - 20,18) + 2,13(x_2 - 24,72) - 1,61(x_2 - 31,54) + 4,38x_3 \\ & - 3,95(x_3 - 20,49) + 0,03(x_3 - 21,32) - 7,47(x_3 - 22,57) + 6,65x_4 \\ & + 13,17(x_4 - 11,46) + 13,07(x_4 - 14,03) - 5,95(x_4 - 17,90) + 0,73x_5 \\ & + 2,02(x_5 - 85,83) - 0,74(x_5 - 89,38) + 9,36(x_5 - 94,71) \end{aligned}$$

Dimana diperoleh nilai GCV sebesar 22,16 , R^2 sebesar 88,49% , dan MSE sebesar 3,24.

Sementara untuk penelitian lainnya dapat digunakan estimator selain *spline truncated* seperti kernel, deret fourier, polinomial, dan lain sebagainya. Dapat pula menggunakan campuran dua atau tiga estimator regresi nonparametrik lainnya. Sementara itu, untuk data penelitian dapat digunakan dalam menganalisis isu-isu strategis saat ini ataupun kasus lainnya.

DAFTAR PUSTAKA

- Anindita, S. D. (2018). *Pemodelan Persentase Balita Stunting di Indonesia Menggunakan Regresi Nonparametrik Spline truncated*. Undergraduate thesis, Institut Teknologi Sepuluh Nopember (ITS) Surabaya.
- Ashari, R., Basyir, V., Afriwardi, Mayetti, Yusrawati, & Desmawati. (2023). *Factors Related to Stunting Incidence in Toddlers Aged 24-59 Months in the Working Area of Kambang Community Health Center, Pesisir Selatan District*. *Scientific Periodical of Public Health and Coastal* 5(2),2023 , pp. 530-549.
- Bele, M. G. L., Hermanto, E. M. P., & Fitriani, F. (2022). *Pemodelan Geographically Weighted Regression pada Kasus Stunting di Provinsi Nusa Tenggara Timur Tahun 2020*. *Jurnal Statistika dan Aplikasinya* Volume 6 Issue 2, December 2022.
- Brianiannita, A., Ismail, L., & Supu, L. (2022). *Faktor Risiko Kejadian Stunting pada Anak Usia 6-7 Tahun. 2022*. *Jurnal Ilmiah Kesehatan*, Vol. 4, No.1, pp 90-97.
- Dewi, A. F., Budiantara, I. N., & Ratnasari, V. (2023). *Parameter estimation of spline truncated, kernel, and Fourier series mixed estimators in semiparametric regression*. *AIP Conference Proceedings* 080005. Volume 2540, Issue 1.
- Eubank, R. L. (1999). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker.
- Hamzah, S. R. (2023). *Analisis Faktor Risiko Kejadian Stunting Pada Balita (24-59 Bulan) Di Kota Kotamobagu*. *Journal Health & Science: Gorontalo Journal Health and Science Community*. Vol. 7 No. 2.
- Hardle, W. (1990). *Applied Non-parametrik Regression*. Cambridge : Cambridge University Press
- Husain, H., Budiantara, I. N., & Zain, I. (2021). *Mixed estimator of spline truncated, Fourier series, and kernel in biresponse semiparametric regression model*. *IOP Conference Series: Earth and Environmental Science*. 880 012046.
- Kemenkes RI. (2013). *Riset Kesehatan Dasar*. Jakarta
- Kemenkes RI. (2016). *Situasi Balita Pendek*. Pusat Data dan Informasi Kesehatan RI. Jakarta
- Kemenkes RI. (2018). *Buletin Stunting. Kementerian Kesehatan RI*, 301(5), 1163–1178.
- Purnamasari, I., Widiyati, F., & Sahli, M. (2022). *Analisis Faktor Risiko yang mempengaruhi kejadian stunting pada Balita*. *Jurnal Penelitian dan Pengabdian Kepada Masyarakat UNSIQ*, Vol.9 No.1, pp 48-56.
- Rachmawati, J. (2018). *Pemetaan Persentase Dan Jumlah Balita Gizi Buruk Di Provinsi Nusa Tenggara Timur Tahun 2017 Dengan Basis Titik Geographically Weighted Regression*

Dan Geographically Weighted Poisson Regression. Undergraduate thesis, Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

- Rifada, M., Chamidah, N., Ningrum, & R. A., Muniroh, L. (2023). *Stunting Determinants Among Toddlers In Probolinggo District Of Indonesia Using Parametric And Nonparametric Ordinal Logistic Regression Models*. *Communications in Mathematical Biology and Neuroscience*, Vol. 2023.
- Sekretariat Percepatan Pencegahan Stunting. (2019). *Strategi Nasional Percepatan Pencegahan Anak Kerdil (Stunting)*. Sekretariat Wakil Presiden Republik Indonesia Edisi Kedua, Cetakan Pertama, Juli 2019
- Setiawan, E., Machmud, R., & Masrul. (2018). Faktor-Faktor yang Berhubungan dengan Kejadian Stunting pada Anak Usia 24-59 Bulan di Wilayah Kerja Puskesmas Andalas Kecamatan Padang Timur Kota Padang Tahun 2018. *Jurnal Kesehatan Andalas*, 7(2), 275–284.