
Analisis Perubahan Berat Badan Balita dengan Estimator *Penalized Spline* Kuadrat

Muhammad Jayzul Usrah^{1*}, Anna Islamiyati¹, Anisa¹

¹Departemen Statistika, Fakultas MIPA,

Universitas Hasanuddin, Makassar, 90245, Indonesia.

* Corresponding author, email: mjayzulusrah@gmail.com

Abstract

Nonparametric regression is a regression approach that is used when one of the parametric assumptions are not fulfilled. One of the estimators in nonparametric regression is penalized spline. The growth pattern of toddler that varied each month of observation make the suitable regression approach is nonparametric penalized spline regression because of its high flexibility. This study aims to obtain an estimate of the growth model for toddler in South Sulawesi. The optimal model obtained with a minimum GCV value of $4.87E-05$ using two point knots that is 14 and 56 with lamda 100. The estimation results show that there are 3 intervals of change patterns in the growth of toddler in South Sulawesi.

Keywords: GCV, Growth of toddler, Knots Point, Nonparametric Regression, Spline.

Abstrak

Regresi nonparametrik merupakan salah satu pendekatan regresi yang digunakan ketika salah satu asumsi parametrik tidak terpenuhi. Salah satu estimator pada regresi nonparametrik yakni *penalized spline*. Tumbuh kembang anak yang bervariasi pada setiap bulan pengamatan menyebabkan pendekatan regresi yang cocok adalah regresi nonparametrik *penalized spline* karena sifat fleksibilitasnya yang tinggi. Penelitian ini bertujuan untuk memperoleh estimasi model pertumbuhan balita di Sulawesi Selatan. Diperoleh model yang optimal dengan nilai GCV minimum $4,87E-05$ menggunakan dua titik knot yakni 14 dan 56 serta lamda 100. Hasil estimasi menunjukkan bahwa terdapat 3 interval pola perubahan pada pertumbuhan balita di Sulawesi Selatan.

Kata Kunci: GCV, Pertumbuhan Balita, Regresi Nonparametrik, Spline, Titik Knot.

1. Pendahuluan

Regresi merupakan suatu metode yang digunakan untuk menyelidiki pola hubungan dan pengaruh beberapa variabel prediktor terhadap variabel [1]. Berdasarkan model, regresi terbagi atas tiga pendekatan yakni regresi parametrik, regresi nonparametrik dan regresi semiparametrik [2]. Regresi parametrik digunakan apabila bentuk kurva fungsi regresi diketahui atau terdapat informasi masa lalu yang lengkap mengenai pola datanya. Namun terdapat kasus dimana pola data tidak mengikuti model regresi parametrik, maka digunakan pendekatan regresi nonparametrik untuk mengestimasi fungsi regresi. Sifat fleksibel terhadap suatu data membuat regresi nonparametrik cocok digunakan pada data yang tidak memiliki model tertentu [3].

Pada regresi nonparametrik terdapat beberapa estimator diantaranya spline *truncated* [4,5,6], spline *smoothing* [7] dan spline *penalized* [8, 9, 10]. Spline merupakan potongan-potongan polinomial yang memiliki sifat tersegmen dan kontinu serta berorde tertentu

yang saling bersambung pada titik-titik knot. Titik knot merupakan titik perpaduan yang terjadi karena terdapat perubahan perilaku pola pada interval yang berlainan [11, 12]. Dalam regresi spline penting untuk memperhatikan titik-titik knot data. Dari titik-titik knot itu dapat digunakan untuk memperoleh model matematis yang optimal. Keunggulan *penalized spline* ada pada kemampuan estimasi yang lebih akurat karena sudah melibatkan titik knot dan parameter penghalus secara simultan dalam mengontrol kemulusan kurva.

Penelitian mengenai regresi nonparametrik *penalized spline* telah banyak dilakukan, diantaranya, Zia, Suparti, dan Safitri (2017) menggunakan *penalized spline* untuk memodelkan hubungan harga penutupan saham dan kurs USD [13]. Islamiyati, Raupong dan Anisa (2019) menggunakan *penalized spline* untuk mengidentifikasi pola perubahan gula darah pasien diabetes [14]. Oleh sebab itu, penelitian ini menggunakan regresi nonparametrik dengan estimator *penalized spline* kuadratik untuk memodelkan pertumbuhan balita di Sulawesi Selatan.

2. Material dan Metode

Data yang digunakan dalam penelitian ini adalah data sekunder dari hasil pengukuran balita. Data berasal dari posyandu di 12 kabupaten/kota di provinsi Sulawesi Selatan pada tahun 2019. Data terdiri dari angka berat badan balita sebagai variabel respon (y) dan usia balita sebagai variabel prediktor (x). Model yang digunakan adalah model regresi nonparametrik dengan estimator *penalized spline* kuadratik :

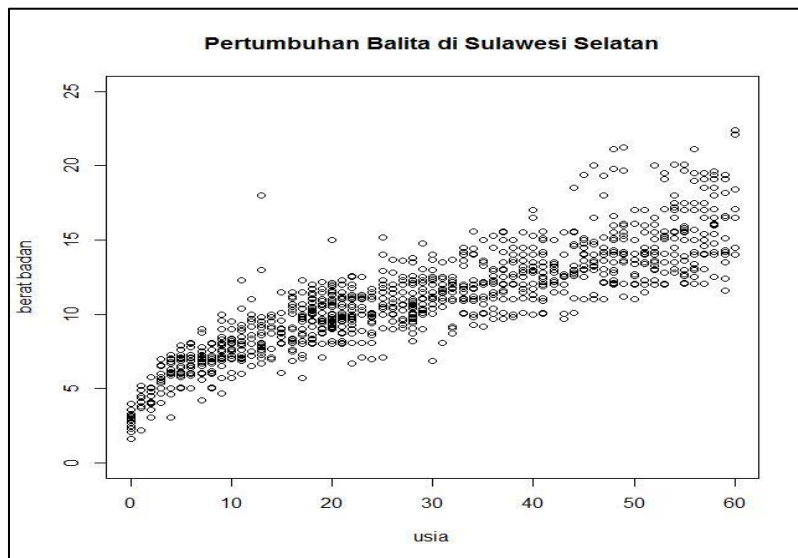
$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 (x - K_1)_+^2 + \dots + \beta_{2+d} (x - K_d)_+^2 \quad (1)$$

Nilai d menunjukkan banyaknya titik knot. Pemilihan titik knot optimal dalam model spline dapat menggunakan metode *Generalized Cross Validation* (GCV). Knot yang optimal berkaitan dengan nilai GCV yang terkecil.

Tahap pertama untuk memperoleh model regresi nonparametrik *penalized spline* kuadratik adalah membuat *scatter plot* untuk melihat pola perubahan berat badan balita berdasarkan usia balita. Tahap kedua yakni mencari titik knot dan parameter penghalus optimum yang mempunyai nilai GCV minimum. Selanjutnya tahap ketiga adalah membandingkan nilai GCV model regresi nonparametrik *penalized spline* kuadratik untuk mendapatkan model terbaik. Tahap terakhir yakni menginterpretasikan pola hubungan berat badan balita dengan usia balita di Sulawesi Selatan berdasarkan model terbaik.

3. Hasil dan Diskusi

Tahapan pemodelan berat badan balita berdasarkan usia dimulai dengan membuat *scatter plot* untuk melihat pola hubungan antara berat badan balita dengan usia yang ditunjukkan pada Gambar 1.



Gambar 1. *Scatter Plot* variabel respon dan prediktor pada data pertumbuhan balita di Sulawesi Selatan.

Berdasarkan Gambar 1 terlihat bahwa semakin usia balita bertambah maka keragaman dari berat badan balita semakin besar. Terlihat pula bahwa terdapat beberapa data outlier pada data, sehingga model estimasi parametrik tidak dapat digunakan. Oleh karena itu, untuk memodelkan data pertumbuhan balita di Sulawesi Selatan digunakan pendekatan regresi nonparametrik dengan estimator *penalized spline* kuadratik.

Pemodelan dilakukan dengan memilih titik knot dan parameter penghalus optimum untuk satu hingga tiga titik knot. Pemilihan titik knot dilakukan melalui *trial and error* dengan mengambil titik yang berada pada interval nilai variabel respon hingga didapatkan nilai GCV paling minimum.

Tabel 1. Nilai GCV *Penalized Spline Kuadratik* satu dan dua titik knot

No.	Titik Knot		Lamda	GCV	Titik Knot		Lamda	GCV
	K1				K1	K2		
1	5		0	6.76E-05	5	10	500	5.47E-05
2	10		200	5.42E-05	5	15	1500	4.98E-05
3	13		400	5.06E-05	5	20	100	4.90E-05
4	14		500	4.99E-05	10	15	3000	5.02E-05
5	15		600	4.98E-05	10	20	2000	4.97E-05
6	16		700	4.99E-05	14	55	400	4.97E-05
7	17		900	5.01E-05	14	56	100	4.87E-05*
8	20		1400	5.11E-05	15	20	6100	5.02E-05
9	25		2600	5.59E-05	20	30	2000	5.28E-05
10	30		3800	6.26E+05	25	40	3500	5.69E-05

Sumber: Data diolah 2020

Tabel 2. Nilai GCV *Penalized Spline Kuadrat* 3 titik knot

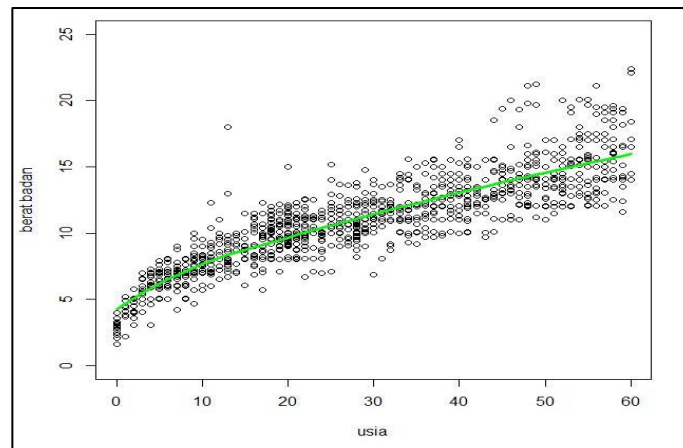
No.	Titik Knot			Lamda	GCV
	K1	K2	K3		
1	5	10	15	3100	5.02E-05
2	5	10	20	2100	4.95E-05
3	5	10	23	1000	4.94E-05
4	10	15	20	7300	4.98E-05
5	10	15	25	3600	5.00E-05
6	15	20	25	4200	5.10E-05
7	15	20	30	4000	5.13E-05
8	20	25	30	1600	5.31E-05
9	25	30	35	500	5.65E-05
10	30	35	40	400	6.14E-05

Sumber: Data diolah 2020

Berdasarkan Tabel 1 dan 2 nilai GCV minimum terbaik berada pada 2 titik knot optimal yakni K1 = 14 dan K2 = 56 dengan lamda 100 dimana nilai GCV yang diperoleh adalah sebesar 4.87E-05. Selanjutnya dilakukan estimasi model regresi nonparametrik *penalized spline* kuadrat menggunakan titik knot optimal K1 = 14 dan K2 = 56 dan lamda 100 dan diperoleh model persamaan sebagai berikut:

$$\hat{y} = 4,2735 + 0,4216x - 0,0083x^2 + 0,0077(x - 14)_+^2 + 0,0004(x - 56)_+^2 \quad (2)$$

Model regresi pada Persamaan (2) menunjukkan bahwa pola kuadrat yang terjadi pada pertumbuhan balita di Sulawesi Selatan optimal pada dua titik knot yaitu 14 dan 56. Hal ini bersesuaian dengan Gambar 2, berat badan balita di Provinsi Sulawesi Selatan memiliki 3 pola perubahan, yaitu berat badan balita naik hingga pada usia 14 bulan, selanjutnya naik perlahan hingga usia 56 bulan, dan kembali naik setelah usia 56 bulan.



Gambar 2. Plot Estimasi kurva perubahan berat badan balita di Sulawesi Selatan

Berikut estimasi model terbaik yaitu model *penalized spline* kuadratik dengan dua titik knot ($K1 = 14$ & $K2 = 56$) dan lamda 100 sebagai berikut:

$$\hat{y} = \begin{cases} 4,2735 + 0,4216x - 0,0083x^2 & 0 \leq x < 14 \\ 5,7827 + 0,206x - 0,0006x^2 & 14 \leq x \leq 56 \\ 7,0371 + 0,1612x - 0,0002x^2 & 56 \leq x \end{cases} \quad (3)$$

4. Kesimpulan

Berdasarkan hasil dan pembahasan, diperoleh kesimpulan sebagai berikut :

1. Data pertumbuhan balita di Sulawesi Selatan dapat dimodelkan dengan menggunakan regresi nonparametrik *penalized spline kuadratik* dengan model matematis sebagai berikut :

$$\hat{y} = 4,2735 + 0,4216x - 0,0083x^2 + 0,0077(x - 14)_+^2 + 0,0004(x - 56)_+^2$$

2. Model terbaik pertumbuhan balita di Sulawesi Selatan berada pada dua titik knot yakni 14 dan 56 dan parameter penghalus 100 dengan nilai GCV yang diperoleh sebesar 4,87E-05

Penelitian ini masih mempertimbangkan satu faktor yaitu berat badan balita dalam mengevaluasi pertumbuhan balita di Sulawesi Selatan. Oleh sebab itu, kajian ini dapat dikembangkan ke penggunaan prediktor lain dan estimator lainnya.

Daftar Pustaka

- [1] Islamiyati, A., Fatmawati & Chamidah, N. Estimation of Covariance Matrix on Bi-Response Longitudinal Data Analysis with Penalized Spline Regression. *J. Phys.: Conf. Ser.* 979 012093, 2018
- [2] Lestari, B., & Budiantara, I. N. Spline Estimator of Triple Response Nonparametric Regression Model. *Jurnal Ilmu Dasar*, Vol. 11, 17-22, 2010.
- [3] Rumlawang, F.Y., Aulele, S.N. & Kasim, N. Penentuan Model Regresi Nonparametrik Spline Pada Data Pertumbuhan Balita Di Desa Nania Provinsi Maluku Tahun 2013-2014. *Barekeng : Jurnal Ilmu Matematika dan Terapan*, 12(1): 27-32, 2018.
- [4] Islamiyati, A. Taksiran Kurva Regresi Spline pada Data Longitudinal dengan Kuadrat Terkecil. *Jurnal Matematika, Statistika & Komputasi*, 11(1) : 97-102, 2014.
- [5] Ramdhani, Z.A., Islamiyati, A. & Raupong. Hubungan Faktor Kolesterol Terhadap Gula Darah Diabetes dengan Spline Kubik Terbobot. *ESTIMASI: Journal of Statistics and Its Application*, 1 (1) : 32-39, 2020.

- [6] Arifin, S., Islamiyati, A. & Raupong. Kemampuan Estimator Spline Linear dalam Analisis Komponen Utama. *ESTIMASI: Journal of Statistics and Its Application*, 1 (1) : 40-47, 2020.
- [7] Lestari, B., Fatmawati & Budiantara, I.N. Smoothing Spline Estimator in Multiresponse Nonparametric Regression for Predicting Blood Pressure and Heart Rate. *International Journal of Academic and Applied Research*, 3 (9) : 1-8, 2019.
- [8] Islamiyati, A., Fatmawati & Chamidah, N. Fungsi Goodness of Fit dalam Kriteria Penalized Spline pada Estimasi Regresi Nonparametrik Birespon untuk Data Longitudinal. *Proseding Seminar Nasional Matematika dan Aplikasinya*. UNAIR Surabaya, 2017.
- [9] Islamiyati, A., Fatmawati & Chamidah, N. Penalized Spline Estimator With Multi Smoothing Parameters in Biresponse Multipredictor Regression Model for Longitudinal Data. *Songklanakarin Journal of Science and Technology*, 42 (4) : 897-909, 2020.
- [10] Islamiyati, A., Sunusi, N., Kalondeng, A., Wati, F. & Chamidah, N. Use of Two Smoothing Parameters in Penalized Spline Estimator for Bi-Variate Predictor Non-Parametric Regression Model. *Journal of Sciences, Islamic Republic of Iran*, 31 (2) : 175-183, 2020.
- [11] Ahmad, R.R., Ghazali, N., Rambeli, A.S., Din, U. K. S. & Hassan, N. Application of Cubic Spline in the Implementation of Braces for the Case of a Child. *Journal of Mathematics and Statistics*, 8 (1) : 144-149, 2012.
- [12] Islamiyati, A., Fatmawati & Chamidah, N. Changes in Blood Glucose 2 Hours After Meals in Type 2 Diabetes Patients Based On Length of Treatment at Hasanuddin University Hospital, Indonesia. *Rawal Medical Journal*, 45 (1) : 31-34, 2020.
- [13] Zia, G.N., Suparti, & Safitri, D. Pemodelan Regresi Spline Menggunakan Metode Penalized Spline Pada Data Longitudinal (Studi Kasus: Harga Penutupan Saham Lq45 Sektor Keuangan Dengan Kurs Usd Terhadap Rupiah Periode Januari 2011- Januari 2016). *Gaussian : Jurnal Statistika UNDIP*, 6(2) : 221-230,2017.
- [14] Islamiyati, A., Raupong & Anisa. Use of Penalized Spline Linear to Identify Change in Pattern of Blood Sugar Based On the Weight of Diabetes Patients. *International Journal of Academic and Applied Research*, 3(12) : 75-78, 2019.