

# Model Regresi Kuantil Spline Orde Dua Dalam Menganalisis Perubahan Trombosit Pasien Demam Berdarah

Anisa<sup>1</sup>, Anna Islamiyati<sup>1,\*</sup>, Sitti Sahriman<sup>1</sup>, Jusmawati Massalesse<sup>2</sup>, Bunga Aprilia<sup>1</sup>

<sup>1</sup>Departemen Statistika, Fakultas MIPA, Universitas Hasanuddin, Makassar 90245, Indonesia

<sup>2</sup>Departemen Matematika, Fakultas MIPA, Universitas Hasanuddin, Makassar 90245, Indonesia

\*Corresponding author. Email: [annaislamiyati701@gmail.com](mailto:annaislamiyati701@gmail.com)

## ABSTRAK

Regresi kuantil dapat digunakan untuk menganalisis data yang mengandung pencilan termasuk data DBD. Spline mampu mengidentifikasi beberapa pola perubahan pada model regresi sehingga penelitian ini menggunakan model regresi kuantil spline dengan orde dua dalam menganalisis data DBD yang terjadi di Kota Makassar. Pada artikel ini, penulis menganalisis bentuk pola perubahan yang terjadi pada trombosit berdasarkan perubahan kandungan hematokrit dari pasien DBD. Kuantil yang dipilih adalah kuartil 0,25; 0,50; dan 0,75 dengan 3 titik knot. Berdasarkan hasil analisis, nilai GCV minimum diperoleh pada penggunaan titik knot yaitu 30,30; 44,80; 47,10 untuk kuartil 0,25; 0,50; dan 0,75. Hal tersebut menunjukkan bahwa pada setiap kuartil, ada empat pola perubahan secara kuadrat yang terjadi pada jumlah trombosit pasien DBD. Kurva parabola yang terbentuk pada setiap segmentasi pola menunjukkan ada saat trombosit mengalami peningkatan dan ada saat trombosit mengalami penurunan. Akan tetapi, rata-rata trombosit mengalami penurunan drastis utamanya pada saat hematokrit mencapai 47,10%.

## Kata Kunci:

Hematokrit; Regresi Kuantil; Spline Orde Dua; Trombosit

## ABSTRACT

Quantile regression can be used to analyze data containing outliers including DHF data. The spline is able to identify several patterns of change in the regression model, so this study uses a second-order quantile spline regression model in analyzing DHF data that occurred in Makassar City. In this article, the authors analyze the pattern of changes that occur in platelets based on changes in the hematocrit content of DHF patients. The selected quantiles are quartiles 0.25; 0.50; and 0.75 with 3-knot points. Based on the results of the analysis, the minimum GCV value obtained at the use of knot points is 30.30; 44.80; 47.10 for the 0.25 quartile; 0.50; and 0.75. This shows that in each quartile, there are four patterns of quadratic changes that occur in the platelet count of DHF patients. The parabolic curve formed in each pattern segmentation shows that there are times when platelets are increasing and there are times when platelets are decreasing. However, the average platelets decreased drastically, especially when the hematocrit reached 47.10%.

---

**Keywords:**

Hematocrit; Quantile Regression; Second Order Spline; Platelets

---

**Format Sitasi:**

A. Anisa, A. Islamiyati, S. Sahriman, J. Massalesse, and B. Aprilia, "Model Regresi Kuantil Spline Orde Dua Dalam Menganalisis Perubahan Trombosit Pasien Demam Berdarah", *Jambura J. Math.*, vol. 5, No. 1, pp. 38–45, 2023, doi: <https://doi.org/10.34312/jjom.v5i1.16086>

---

## 1. Pendahuluan

Regresi kuantil merupakan salah satu regresi yang dapat digunakan untuk data yang mengandung outlier. Kandungan outlier pada data akan menyebabkan sebaran data tidak berbentuk simetris sehingga ada pelanggaran asumsi kenormalan sehingga dapat menggunakan regresi kuantil [1]. Regresi kuantil membagi data menjadi beberapa kuantil tertentu sehingga kondisi data akan terlihat secara lengkap baik pada bagian median maupun pada ekor sebaran data [2]. Beberapa penelitian DBD yang menggunakan regresi kuantil diaplikasikan pada kota yang berbeda, diantaranya DBD di Kota Surabaya [3], DBD di Kota Palu [4], dan DBD di Kota Makassar [5].

Pada perkembangan selanjutnya, kondisi data semakin bervariasi dan terdapat beberapa kasus yang menunjukkan pola hubungan antara prediktor dengan respon tidak berpola parametrik. Oleh sebab itu, berkembang regresi nonparametrik yang dapat digunakan dalam berbagai jenis data aplikasi. Terdapat beberapa estimator dalam regresi nonparametrik, diantaranya spline [6], kernel [7], polinomial lokal [8], dan deret fourier [9]. Untuk estimator spline, perkembangannya juga sudah sangat pesat karena memiliki fleksibilitas tinggi yaitu data yang akan mencari pola data yang sesuai [10]. Jenis estimator spline yang telah dikembangkan peneliti, diantaranya spline truncated [11], penalized spline [12], spline smoothing [13] dan spline poisson [14].

Estimator spline juga telah dikembangkan penggunaannya dalam regresi kuantil diantaranya spline kubik restricted kuantil [15] dan regresi kuantil spline linier [16]. Akan tetapi, selain pola linier dan kubik, terdapat beberapa data yang dapat saja mengikuti pola kuadratik karena adanya pola perubahan naik turun pada data. Oleh sebab itu, penelitian ini menggunakan regresi kuantil spline kuadratik dalam menganalisis data DBD di Kota Makassar dengan melibatkan tiga titik knot. Estimator yang digunakan adalah spline truncated yang melibatkan titik knot dalam fungsi *goodness of fit*. Hal tersebut terkait dengan kecenderungan pola yang terjadi pada hubungan antara hematokrit dengan trombosit pasien DBD. Perubahan naik dan turun dari trombosit pasien telah ditunjukkan pada penelitian sebelumnya dengan menggunakan spline biprediktor [17] dan analisis survival MARS [18].

## 2. Metode

Data DBD di Kota Makassar dianalisis melalui pendekatan regresi kuantil dengan estimator spline truncated kuadratik. Data memuat variabel prediktor yaitu hematokrit pasien DBD dan variabel respon yaitu trombosit pasien DBD. Variabel tersebut dimodelkan ke dalam regresi kuantil spline kuadratik berikut:

$$y_i(\theta) = \beta_0(\theta) + \beta_1(\theta)x_i + \beta_2(\theta)x_i^2 + \beta_3(\theta)(x_i - K_1)_+^2 + \dots + \beta_p(\theta)(x_i - K_r)_+^2 + \varepsilon_i(\theta) \tag{1}$$

dengan  $\theta$  adalah kuantil 0,25; 0,50; dan 0,75,  $\beta$  adalah koefisien regresi kuantil spline kuadrat,  $K$  adalah titik knot,  $p$  adalah banyaknya koefisien regresi, dan  $r$  adalah banyaknya titik knot.

Selanjutnya Persamaan (1) dapat dinyatakan dalam bentuk matriks sebagai berikut:

$$Y = X\beta(\theta) + \varepsilon(\theta) \tag{2}$$

dengan,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}; X = \begin{bmatrix} 1 & x_1 & x_1^2 & (x_1 - K_k)_+^2 & \dots & (x_1 - K_r)_+^2 \\ 1 & x_2 & x_2^2 & (x_2 - K_k)_+^2 & \dots & (x_2 - K_r)_+^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & (x_n - K_k)_+^2 & \dots & (x_n - K_r)_+^2 \end{bmatrix};$$

$$\beta(\theta) = \begin{bmatrix} \beta_0(\theta) \\ \beta_1(\theta) \\ \beta_2(\theta) \\ \vdots \\ \beta_p(\theta) \end{bmatrix}; \varepsilon(\theta) = \begin{bmatrix} \varepsilon_1(\theta) \\ \varepsilon_2(\theta) \\ \vdots \\ \varepsilon_n(\theta) \end{bmatrix}.$$

Dalam kasus regresi kuantil, besarnya deviasi atas dan deviasi bawah tidak sama. Sehingga harus diboboti dengan  $\theta$  untuk bagian positif (deviasi atas) dan  $1 - \theta$  untuk bagian negatif (deviasi bawah). Oleh sebab itu, estimasi parameter  $\beta$  diperoleh dengan meminimumkan kriteria estimasi berikut:

$$\hat{\beta}(\theta) = \min_{\beta \in R^p} \theta \sum_{i=1; y > q}^n |\varepsilon_{1i}(\theta)| + (1 - \theta) \sum_{i=1; y < q}^n |\varepsilon_{2i}(\theta)| \tag{3}$$

Dengan fungsi kendala

$$X\beta(\theta) + \varepsilon_{1i}(\theta) - \varepsilon_{2i}(\theta) = y; \varepsilon_1, \varepsilon_2 \geq 0$$

Selanjutnya melalui pendekatan metode simpleks, didapatkan estimasi parameter  $\hat{\beta}(\theta) = [\hat{\beta}_0(\theta) \hat{\beta}_1(\theta) \hat{\beta}_2(\theta) \dots \hat{\beta}_p(\theta)]'$ , sehingga diperoleh estimasi model regresi kuantil spline kuadrat yaitu:

$$\hat{y}(\theta) = X\hat{\beta}(\theta) \tag{4}$$

Kuantil spline optimal dengan titik-titik knot optimal yaitu dengan menggunakan kriteria GCV minimum. Fungsi GCV didefinisikan sebagai berikut:

$$\begin{aligned}
 GCV(K_K) &= \frac{MSE(K_K)}{[n^{-1}trace(I - A(K_K))]^2} \\
 &= \frac{n^{-1}y^T(I - A(K_K))^T(I - A(K_K))y}{[n^{-1}trace(I - A(K_K))]^2}
 \end{aligned}$$

dengan  $K_K = (K_1, K_2, \dots, K_r)$  adalah titik knot dan matriks  $A(K_K)$  adalah

$$A(K_K) = X[K] \left( X[K]' X[K] \right)^{-1} X[K]'$$

### 3. Hasil dan Pembahasan

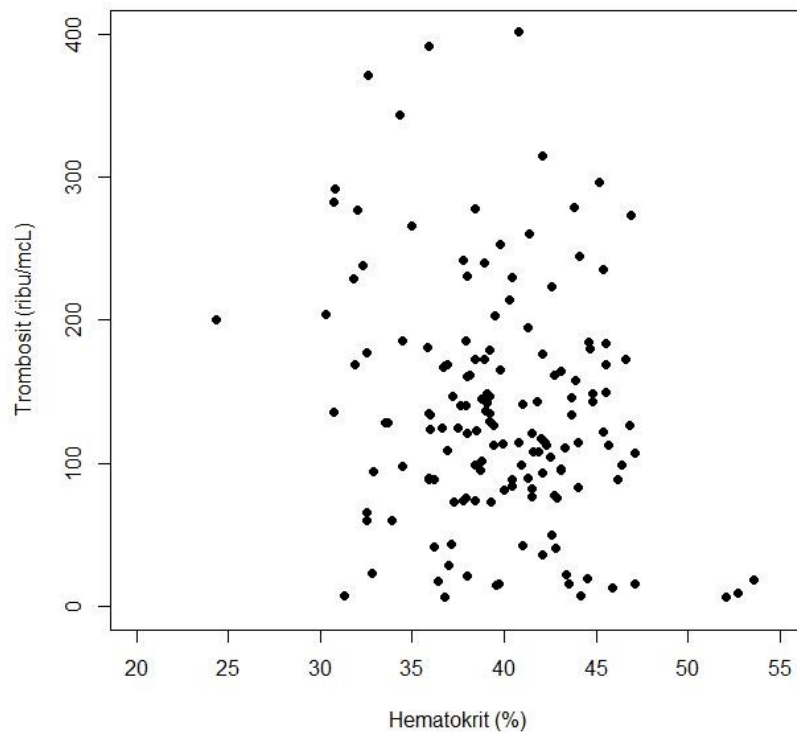
Penggunaan regresi nonparametrik digunakan pada saat plot data tidak mengikuti pola parametrik. Hal tersebut diidentifikasi melalui *scatter plot* seperti yang ditunjukkan pada Gambar 1. Plot data pada Gambar 1 menunjukkan bahwa sebaran data trombosit dengan hematokrit cenderung bervariasi. Variasi tersebut ditunjukkan oleh ada pasien yang didiagnosis menderita DBD dengan memiliki nilai hematokrit normal tetapi jumlah trombosit yang abnormal. Sebaliknya, ada pula pasien yang didiagnosis DBD memiliki trombosit normal tetapi jumlah hematokrit di bawah normal yaitu <38%. Hal tersebut menyebabkan tidak ada kecenderungan plot data berbentuk parametrik, akan tetapi sebaran data menjadi tidak beraturan sehingga model estimasi regresi parametrik tidak dapat digunakan. Selain itu, distribusi data yang heterogen menyebabkan perlunya dimodelkan dalam berbagai kuartil melalui regresi kuantil. Oleh karena itu, data jumlah trombosit pasien DBD dengan nilai hematokrit dianalisis dengan menggunakan pendekatan regresi kuantil nonparametrik melalui estimator spline orde dua.

Penelitian ini menggunakan 3 titik knot dalam spline orde dua karena plot data yang sangat tidak beraturan. Dalam hal ini, diperlukan analisis dalam berbagai segmentasi pola perubahan yang terjadi berdasarkan titik-titik knot yang terpilih pada setiap kuartil 0,25, 0,50, dan 0,75. Titik knot yang terpilih adalah sama untuk setiap kuantil yang diujikan.

**Tabel 1.** Nilai GCV dari estimasi model spline orde dua dengan 3 titik knot

No	Titik Knot			GCV
	K <sub>1</sub>	K <sub>2</sub>	K <sub>3</sub>	
1	30,70	39,80	46,90	6.318,084
2	30,50	44,20	47,10	6.303,352
3	32,50	44,60	47,10	6.301,656
4	40,40	45,20	47,10	6.321,220
5	43,10	45,20	47,10	6.306,688
6	32,50	44,80	47,10	6.298,575
7	30,70	44,80	47,10	6.291,574
8	30,70	44,80	46,90	6.292,949
<b>9</b>	<b>30,30</b>	<b>44,80</b>	<b>47,10</b>	<b>6.291,477</b>
10	31,80	44,80	46,90	6.296,674

Berdasarkan Tabel 1 diperoleh nilai GCV dari estimasi model regresi spline orde dua dengan beberapa variasi tiga titik knot. Nilai GCV minimum diperoleh pada titik knot  $K_1 = 30,30$ ,  $K_2 = 44,80$ ; dan  $K_3 = 47,10$  yaitu sebesar 6.291,477. Berdasarkan 3 titik knot



**Gambar 1.** Plot pengaruh nilai hematokrit (x) terhadap jumlah trombosit (y)

optimal tersebut, selanjutnya digunakan dalam analisis model regresi kuantil spline orde dua dan diperoleh taksiran persamaan regresi seperti pada Tabel 2.

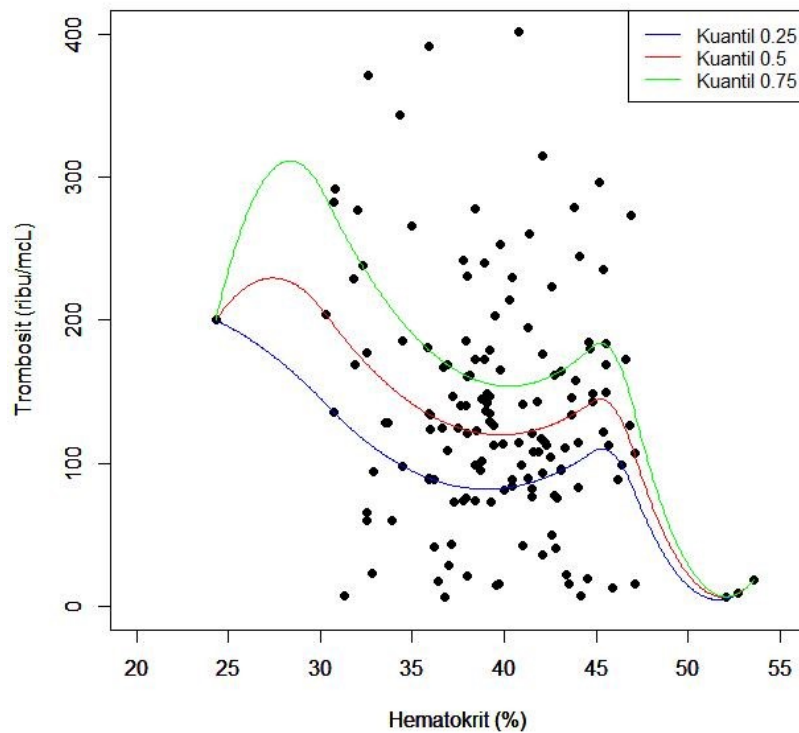
**Tabel 2.** Estimasi model kuantil spline kuadratik tiga titik knot

Parameter	Kuantil		
	0,25	0,5	0,75
$\hat{\beta}_0$	-28,392	-2063,300	-5.135,52
$\hat{\beta}_1$	24,781	167,302	384,229
$\hat{\beta}_2$	-0,633	-3,052	-6,776
$\hat{\beta}_3$	1,411	3,979	8,102
$\hat{\beta}_4$	-9,956	-11,901	-15,194
$\hat{\beta}_5$	12,842	15,256	19,061

Berdasarkan Tabel 2 diperoleh estimasi model kuantil spline kuadratik dengan tiga titik knot ( $K_1 = 30,30$ ,  $K_2 = 44,80$  dan  $K_3 = 47,10$ ) pada kuantil 0,25 ; 0,50 ; 0,75 sebagai berikut:

$$\begin{aligned} \hat{y}(0,25) &= -28,392 + 24,781x - 0,633x^2 + 1,411(x - 30,3)_+^2 - 9,956(x - 44,8)_+^2 \\ &\quad + 2,842(x - 47,1)_+^2 \\ \hat{y}(0,50) &= -2.063,3 + 167,302x - 3,052x^2 + 3,979(x - 30,3)_+^2 - 11,901(x - 44,8)_+^2 \\ &\quad + 15,256(x - 47,1)_+^2 \\ \hat{y}(0,75) &= -5.135,52 + 384,229x - 6,776x^2 + 8,102(x - 30,3)_+^2 - 15,194(x - 44,8)_+^2 \\ &\quad + 19,061(x - 47,1)_+^2 \end{aligned}$$

Model regresi untuk setiap kuartil yang diperoleh berbeda, tetapi kecenderungan perubahan trombosit tampak tidak berbeda secara signifikan. Persamaan regresi kuantil spline orde dua yang telah diperoleh pada setiap kuartil juga ditunjukkan pada Gambar 2.



**Gambar 2.** Estimasi kurva regresi kuantil spline orde dua dengan 3 titik knot antara jumlah trombosit dengan nilai hematokrit pasien DBD

Gambar 2 menunjukkan kurva regresi kuantil spline orde dua dengan 3 titik knot pada kuartil 0,25; 0,50; 0,75. Pada kuartil 0,75, pola perubahan trombosit terlihat masih mengalami peningkatan secara kuadrat sebelum hematokrit pasien DBD mencapai 30,30%. Setelah itu, untuk kandungan hematokrit pada interval 30,30-44,80% menunjukkan pola parabola terbuka ke atas. Ada kecenderungan trombosit menurun drastis setelah trombosit di atas 30,30% dan perlahan meningkat sangat kecil saat mencapai 44,80% hematokrit pasien. Pola terakhir setelah mencapai 44,80% hematokrit yaitu di atas ukuran normal hematokrit seseorang, terlihat jumlah trombosit semakin mengalami penurunan sangat drastis. Pola yang sama terlihat pada kuartil 0,50, hal yang berbeda hanya pada ketajaman kenaikan jumlah trombosit sebelum mencapai 30,30% dan 44,80% hematokrit. Kenaikan trombosit pada kuartil 0,50 lebih rendah dibandingkan dengan kondisi pada kuartil 0,75. Begitu pun pada kuartil 0,25, kenaikan trombosit semakin kecil dan hanya terjadi pada hematokrit sebelum 44,80%.

Pola perubahan trombosit pada pasien DBD berdasarkan model regresi kuantil spline orde dua menunjukkan kecenderungan trombosit menurun kuadrat dan hal tersebut juga terjadi dalam batas interval hematokrit normal. Hasil ini tentu dapat menjadi informasi bagi tim medis bahwa diperlukan juga langkah-langkah awal perawatan

sebelum nilai hematokrit pasien meningkat di atas 45%. Peningkatan hematokrit akan menyebabkan trombosit turun karena terjadi kebocoran plasma dan trombosit yang berperan untuk menutupi kebocoran tersebut sehingga memicu menurunnya jumlah trombosit.

### 4. Kesimpulan

Perubahan jumlah trombosit pasien DBD berdasarkan jumlah hematokrit melalui model regresi kuantil spline orde dua menunjukkan ada kurva parabola terbuka ke atas dan terbuka ke bawah. Hal tersebut terjadi pada setiap kuartil yang terpilih yaitu 0,25; 0,50, dan 0,75. Kenaikan trombosit yang tinggi hanya terjadi pada kuartil 0,75 untuk hematokrit di bawah 30,30%. Adapun pada interval dan kuartil lainnya, kenaikan trombosit sangat kecil sehingga tim medis sangat perlu untuk aktif memantau kondisi pasien DBD meskipun memiliki hematokrit normal.

### Ucapan Terima Kasih

Penelitian ini dibiayai oleh LPPM Universitas Hasanuddin atas Hibah Internal Penelitian untuk Skim Penelitian Dasar Unhas Tahun Anggaran 2022 dengan Nomor Kontrak : 1474/UN4.22/PT.01.03/2022 tanggal 9 Juni 2022.

### Referensi

- [1] F. Yanuar, H. Yozza, and I. Rahmi, "Penerapan Metode Regresi Kuantil pada Kasus Pelanggaran Asumsi Kenormalan Sisaan," *Eksakta*, vol. 1, pp. 33–37, 2016.
- [2] M. Furno, "Parameter instability in quantile regression," *Statistical Modelling*, vol. 7, no. 4, pp. 345–362, dec 2007, doi: 10.1177/1471082X0700700405.
- [3] M. Y. Matdoan and A. M. Balami, "Estimasi Parameter Regresi Kuantil dengan Fungsi Spline Truncated Pada Kasus Demam Berdarah Dengue di Kota Surabaya," *Jurnal MSA ( Matematika dan Statistika serta Aplikasinya )*, vol. 7, no. 1, p. 44, aug 2019, doi: 10.24252/msa.v7i1.7511.
- [4] N. Idris, R. Rais, and I. T. Utami, "Aplikasi Regresi Kuantil pada Kasus DBD di Kota Palu Sulawesi Tengah," *Jurnal Ilmiah Matematika dan Terapan*, vol. 5, no. 1, pp. 108–117, 2018.
- [5] W. N. A. Puteri, A. Islamiyati, and A. Anisa, "Penggunaan Regresi Kuantil Multivariat pada Perubahan Trombosit Pasien Demam Berdarah Dengue," *ESTIMASI: Journal of Statistics and Its Application*, vol. 1, no. 1, pp. 1–9, feb 2020, doi: 10.20956/ejsa.v1i1.9224.
- [6] Z. A. Ramdhani, A. Islamiyati, and R. Raupong, "Hubungan Faktor Kolesterol Terhadap Gula Darah Diabetes dengan Spline Kubik Terbobot," *ESTIMASI: Journal of Statistics and Its Application*, vol. 1, no. 1, pp. 32–39, jan 2020, doi: 10.20956/ejsa.v1i1.9252.
- [7] N. Chamidah and B. Lestari, "Estimation of covariance matrix using multi-response local polynomial estimator for designing children growth charts: A theoretically discussion," *Journal of Physics: Conference Series*, vol. 1397, no. 1, p. 012072, dec 2019, doi: 10.1088/1742-6596/1397/1/012072.
- [8] N. Chamidah and T. Saifudin, "Estimation of children growth curve based on kernel smoothing in multi-response nonparametric regression," *Applied Mathematical Sciences*, vol. 7, no. 37, pp. 1839–1847, 2013, doi: 10.12988/ams.2013.13168.
- [9] M. F. F. Mardianto, E. Tjahjono, and M. Rifada, "Statistical modelling for prediction of rice production in Indonesia using semiparametric regression based on three forms of fourier series estimator," *ARPN Journal of Engineering and Applied Sciences*, vol. 14, no. 15, 2019.
- [10] A. Islamiyati, Raupong, A. Kalondeng, and U. Sari, "Estimating the confidence interval of the regression coefficient of the blood sugar model through a multivariable linear spline with known variance," *Statistics in Transition New Series*, vol. 23, no. 1, pp. 201–212, mar 2022, doi: 10.2478/stattrans-2022-0012.
- [11] A. Islamiyati, A. Kalondeng, N. Sunusi, M. Zakir, and A. K. Amir, "Biresponse

- nonparametric regression model in principal component analysis with truncated spline estimator," *Journal of King Saud University - Science*, vol. 34, no. 3, p. 101892, apr 2022, doi: 10.1016/j.jksus.2022.101892.
- [12] A. Islamiyati, F. Fatmawati, and N. Chamidah, "Use of Two Smoothing Parameters in Penalized Spline Estimator for Bi-Variate Predictor Non-Parametric Regression Model," *Journal of Science, Islamic of Republic Iran*, vol. 31, no. 2, pp. 175–183, 2020, doi: <https://doi.org/10.22059/jsciences.2020.286949.1007435>.
- [13] B. Lestari, Fatmawati, I. N. Budiantara, and N. Chamidah, "Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches," *Journal of Physics: Conference Series*, vol. 1397, no. 1, p. 012064, dec 2019, doi: 10.1088/1742-6596/1397/1/012064.
- [14] N. Jao, A. Islamiyati, and N. Sunusi, "Pemodelan Regresi Nonparametrik Spline Poisson pada Tingkat Kematian Bayi di Sulawesi Selatan," *Estimasi: Journal of Statistics and Its Application*, vol. 3, no. 1, pp. 14–22, 2022, doi: <https://doi.org/10.20956/ejsa.vi.11997>.
- [15] R. A. Marrie, N. V. Dawson, and A. Garland, "Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables," *Journal of Clinical Epidemiology*, vol. 62, no. 5, pp. 511–517.e1, may 2009, doi: 10.1016/j.jclinepi.2008.05.015.
- [16] B. Aprilia, A. Islamiyati, and A. Anisa, "Platelet Modeling Based On Hematocrit in DHF Patients with Spline Quantile Regression," *International Journal of Academic and Applied Research (IJAAAR)*, vol. 3, no. 12, pp. 51–54, 2019.
- [17] A. Islamiyati, "Regresi Spline Polynomial Truncated Biprediktor untuk Identifikasi Perubahan Jumlah Trombosit Pasien Demam Berdarah Dengue," *Al-Khwarizmi: Jurnal Pendidikan Matematika dan Ilmu Pengetahuan Alam*, vol. 7, no. 2, pp. 97–112, oct 2019, doi: 10.24256/jpmipa.v7i2.799.
- [18] S. F. Nisai and I. N. Budiantara, "Analisis Survival dengan Pendekatan MARS pada Kasus DBD," *Jurnal Sains dan Seni ITS*, vol. 1, no. 1, pp. 318–323, 2012, doi: <https://doi.org/10.24252/msa.v5i2.4511>.



This article is an open-access article distributed under the terms and conditions of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). Editorial of JJoM: Department of Mathematics, Universitas Negeri Gorontalo, Jln. Prof. Dr. Ing. B.J. Habibie, Moutong, Tilongkabila, Kabupaten Bone Bolango, Provinsi Gorontalo 96554, Indonesia.