BAREKENG: Journal of Mathematics and Its Applications
March 2025 Volume 19 Issue 1 Page 0271–0280
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss1pp0271-0280

KNOT OPTIMIZATION FOR BI-RESPONSE SPLINE NONPARAMETRIC REGRESSION WITH GENERALIZED CROSS-VALIDATION (GCV)

Andre Fajry Al Barra[1*], Dewi Retno Sari Saputro[2]

[1,2]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Sebelas Maret*
Jln. Ir. Sutami No. 36, Surakarta, 57126, Indonesia

*Corresponding author's e-mail: * **barra_12@student.uns.ac.id***

ABSTRACT

Article History:

Received: 28th May 2024
Revised: 30th August 2024
Accepted: 1st September 2024
Published:13th January 2025

Keywords:

Bi-Response;
GCV;
Knot Point;
Nonparametric Regression;
Spline.

Nonparametric regression is a statistical method used to model relationships between variables without making strong assumptions about the functional form of the relationship. Nonparametric regression models are flexible and can capture complex relationships that may not be adequately represented by simple parametric forms. Spline is one of the approaches used in nonparametric regression. Splines have the disadvantage of having to use optimal nodes in the data. Therefore, this article discusses the retrieval of optimal knot points using the generalized cross-validation method in the nonparametric bi-response spline regression model. The research results showed that the generalized-cross validation method is the best method for selecting nodes from other methods such as CV, AIC, BIC, RSS, or a more explicit validation-based approach method because of the development of the Cross Validation (CV) method which automatically selects the optimal number of nodes based on the balance between bias and variance. The process of optimizing knot points with Generalized Cross Validation (GCV) on bi-response spline nonparametric regression is implemented using Python can provide optimization at optimal knot points. Based on the results of the generalized cross-validation model analysis, it is concluded that GCV can effectively optimize knot points for spline fitting, ensuring a balanced and efficient model in capturing data patterns without overfitting.
.

How to cite this article:

A. F. A. Barra and D. R. S. Saputro., "KNOT OPTIMIZATIONFOR BI-RESPONSE SPLINE NONPARAMETRIC REGRESSION WITH GENERALIZED CROSS-VALIDATION (GCV)," *BAREKENG: J. Math. & App.,* vol. 19, iss. 1, pp. 0271-0280, March, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

The industrial revolution 4.0 has changed digital transformation for all groups on earth. This revolution requires humans to continue to develop various aspects of technology, one of which is machine learning. Machine learning is a field of computer science that studies algorithms and techniques to automatically generate solutions to complex and difficult problems [1]. One technique in machine learning is regression analysis. Regression analysis is a statistical method for studying the relationship between several variables and can predict a variable [2].

Regression analysis is generally used for model validation, forecasting, controlling variables, understanding relationships between variables, and others. Regression analysis is used to describe the relationship between independent variables (product quality, price, income, etc.) and dependent variables (time, purchasing decisions, etc.) [3]. Regression analysis itself has several approaches that can be used to make it easier to find out the pattern of variable relationships. Nonparametric regression is an approach used when the shape or pattern of the regression curve between the response variable and the predictor variable is unknown [4].

Nonparametric regression has several techniques for carrying out estimates. One of the regression models with a nonparametric approach that is very often used to estimate regression curves is spline regression [5]. Nonparametric spline regression is an analysis used to obtain an estimated regression curve by estimating data patterns according to their movement. If the regression model consists of two response variables with a relationship between the two response variables, then the model is called bi-response regression. If in the bi-response regression, there is an unknown shape of the regression curve then the analysis used is nonparametric bi-response spline regression [6]. The performance of splines in bi-response nonparametric regression is highly dependent on the location of the knot points used in interpolation. Improper placement of knot points can produce inaccurate or unstable splines [7].

In research [8] the use of generalized cross-validation can select more optimal parameters at knot points. Generalized Cross Validation (GCV) is a statistical method used to select model parameters in the context of regularization and smoothing. Generalized cross-validation has the basic idea of choosing a K value (knot) that minimizes the assumption of a function of the ith observation from the data. Next, a function is obtained from a subsample of size $n-1$ taken from the original data which will later be carried out repeatedly [9]. The advantage of GCV is that it can effectively optimize node points for spline fitting, ensuring a model that is balanced and efficient in capturing data patterns without overfitting.

In another previous study conducted by [10] with the title "Estimates for The Generalized Cross-Validation Function via An Extrapolation and Statistical Approach". The research results can be concluded that generalized cross-validation is the best method for extrapolation and statistical approach. Therefore, this method is the most appropriate method based on relevant previous references. The authors are interested in conducting research with the title "Knot Optimization for Bi-response Spline Nonparametric Regression with Generalized Cross-Validation (GCV)".

2. RESEARCH METHODS

This research uses literature study research methods and reviews of related journals. The Visualization of Similarities Viewer (VosViewer), Nonparametric Spline Bi-response Regression, and Optimal Knot Points will be the three bases on which this research was conducted.

2.1 Visualization of Similarities Viewer (VosViewer)

In research, VosViewer is used for bibliometric analysis, mapping topics to the latest research, searching for the most used references in a particular field, and more [11]. VosViewer has features to make bibliometric data visualization easier to present. VosViewer works to combine several documents and will understand the content, patterns, and trends of these documents [12][13][14]. The mapping produced by VosViewer from bibliometric data related to bi-response nonparametric regression, spline, and generalized cross-validation can be seen in **Figure 1**, which shows the results of bibliometric visualization through extraction in VosViewer with title and abstract field.

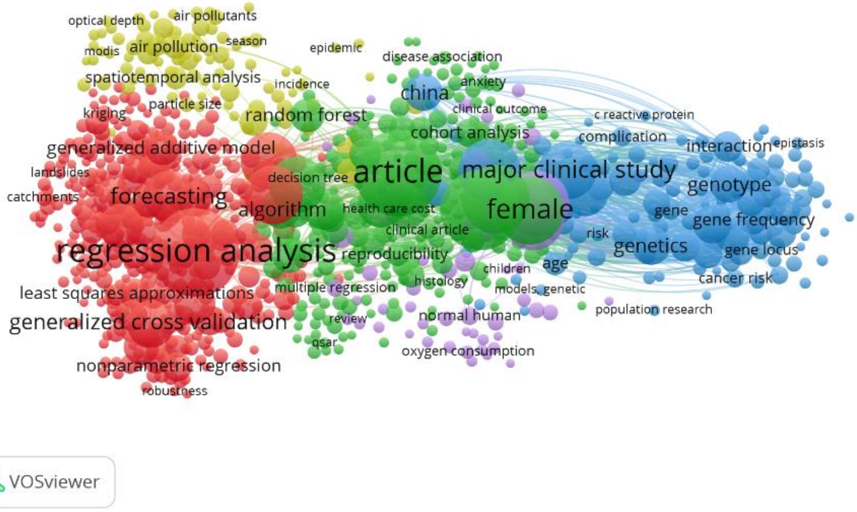


Figure 1. VosViewer with Title and Abstract Visualization

Based on **Figure 1**, visualization of bibliometric data taken from Scopus metadata. A larger circle indicates that the data is widely used. In the visualization, it can be seen that forecasting, algorithms, generalized cross-validation, and regression analysis are widely used in research terms.

2.2 Nonparametric Regression

The model used if the regression curve is not found or does not follow a certain pattern is called a nonparametric regression model **[15]**. Nonparametric regression models are used to estimate regression curves that depend only on observed data. For example, the predictor variable and the response variable respectively (y_i, x_i), then the relationship between x_i and y_i is written as.

$$y_i = f(x_i) + \varepsilon_i, \ \ i = 1,2, \dots, n \tag{1}$$

$f(x_i)$ is the regression function for estimating y_i dan ε_i is the error which is expected to be normally distributed.

2.3 Bi-response Spline Nonparametric Regression

Bi-response regression refers to a regression model in which there is more than one response variable, and between these variables, there is a strong correlation or relationship, both logically and mathematically **[16]**. If the regression curve for bi-response regression is not known, then the approach used is nonparametric bi-response regression. This nonparametric approach uses a spline function involving several predictor variables so the model is called a bi-response spline nonparametric regression model. Based on **[16]**, the function for nonparametric regression is bi-response spline is written with **Equation (2)** and **Equation (3)**.

$$\hat{f}^{(1)}(x_i) = \sum_{j=0}^{m_1-1} \hat{\beta}_j^{(1)} x_i^j + \sum_{j=1}^{R} \hat{\beta}_{j+m_1-1}^{(1)} (x_i - k_j)_+^{m_1-1} \tag{2}$$

$$\hat{f}^{(2)}(x_i) = \sum_{j=0}^{m_2-1} \hat{\beta}_j^{(2)} x_i^j + \sum_{j=1}^{R} \hat{\beta}_{j+m_2-1}^{(2)} (x_i - k_j)_+^{m_2-1} \tag{3}$$

with:

$\hat{f}^{(1)}(x_i)$: Regression function to-*i* the unknown for response 1

$\hat{f}^{(2)}(x_i)$: Regression function to-*i* the unknown for response 2

2.4 Knot Point

One way to choose optimal knot points is the Generalized Cross-Validation (GCV) method **[17]**. The advantage of the GCV method is that it has asymptotically optimal properties that are not found in other methods **[18]**. The formula for calculating GCV values in bi-response regression is written with **Equation (4) [19]**.

$$GCV = \frac{MSE}{[N^{-1}trace(I-H)]^2} \quad (4)$$

$$MSE = N^{-1}\sum_{p=1}^{2}\sum_{i=1}^{n}(Y_i^{(p)} - \hat{Y}_i^{(p)})^2 \quad (5)$$

where $N = 2n$

with:

n : Many observations

I : Identity matrix

H : Hat matrix

3. RESULTS AND DISCUSSION

The regression equation makes it possible to predict the values of a dependent variable from the values of one or more independent variables. An independent variable is a variable whose value is known, while a dependent variable is a variable whose value is not yet known and will be predicted [20]. On the other hand, the use of regression analysis is for description, control, and prediction (forecasting) and is also most widely used for forecasting. Regression is also used to determine the magnitude of the variable contribution [21].

The spline function is the sum of a polynomial function with a truncated function. In this section, we discuss nonparametric regression models, where the estimation of the g curve is carried out using a spline [22]. We will be given paired data (x_1, y_1) and the relationship between the two is assumed to follow a nonparametric regression model with **Equation (6)**.

$$y_i = f(x_j) + \varepsilon_j,\ j = 1,2,\dots,n \quad (6)$$

The regression curve is approximated by a spline function g with r is a K knot points, a spline regression model is written with **Equation (7)**.

$$y_j = g(x_j) + \varepsilon_j,\ j = 1,2,..,n \quad (7)$$

If **Equation (7)** is written in matrix form, it is obtained as.

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} g(x_1) \\ g(x_2) \\ \vdots \\ g(x_n) \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (8)$$

The shape of the regression curve of $g(x_j)$ from **Equation (7)** is assumed to be unknown, while the errors $\varepsilon_j, j = 1,2,..,n$ are mutually independent with zero mean and variance σ^2.

The spline function is a function that can generally be written with **Equation (9)**.

$$g(x_j) = \sum_{i=0}^{m} a_i x_j^i + \sum_{k=1}^{r} \beta_k (x_j - k_k)_+^m \quad (9)$$

with α_i and β_k in **Equation (9)** are real constants. According to [23], if the spline regression model is written in matrix form, it is obtained as:

$$\mathrm{Y} = \mathrm{X}[K_1, K_2, \dots, K_R]\beta + \varepsilon \quad (10)$$

Next, the estimated parameter β in **Equation (10)** is obtained using the Weighted Least Squares (WLS) method in nonparametric bi-response spline regression by minimizing the weighted sum of squared errors which is written with **Equation (11)**.

$$\begin{aligned} \varepsilon^T W \varepsilon &= (Y - X\beta)^T W (Y - X\beta) \\ &= (Y^T - \beta^T X^T)(WY - W\beta X) \\ &= Y^T WY - Y^T WX\beta - \beta^T X^T WY + \beta^T X^T WX\beta \end{aligned} \quad (11)$$

where W is a diagonal matrix.

To determine the value of $\hat{\beta}$ which minimizes the sum of the squares of the weighted errors, a reduction in **Equation (11)** is carried out on β^T and it must be equal to zero, obtained as:

$$\hat{\beta} = (X^T W X)^{-1} X^T W Y \tag{12}$$

From the optimization solution of **Equation (11)**, $\hat{\beta}$ is obtained. Consequently, the estimate for a spline regression curve with K knots is given as.

$$\begin{aligned} \hat{g}(x_j) &= X(k)\hat{\beta} \\ &= X(k)(X(k)^T X(k))^{-1} X(k)^T y \\ &= A(k)y \end{aligned} \tag{13}$$

where $A(k)$ is a function of the knot point K.

From the results obtained, it can be seen that $\hat{g}(x_j)$ is a linear estimator in observations and is very dependent on the knot point. Next, to obtain the best spline, it is necessary to select the optimal knot points using the generalized cross validation method.

The generalized cross-validation method is one method that is often used in selecting optimal knot points. This GCV method is a modification of the Cross-Validation (CV) method. According to **[24]**, the optimal knot point is obtained from the smallest GCV value. The CV function can be written as

$$CV(k) = n^{-1} \frac{\sum_{i=1}^{n} (y_i - \hat{f}(x_i))^2}{[1 - n^{-1} trace(A(k)]^2} \tag{14}$$

$GCV(\tilde{k})$ is a vector that contains the GCV value of the knot point obtained from dividing the sum of the squared residues of $\hat{f}(x)$ by $n\{[1 - n^{-1} trace(A(k)]\}^2$. The general equation of GCV can be written with **Equation (15)**.

$$\begin{aligned} GCV(k) &= \frac{MSE}{[n^{-1} trace(I - A(k)]^2} \\ &= n^{-1} \frac{\sum_{i=1}^{n} (y_i - \hat{f}(x_i))^2}{[n^{-1} trace(I - A(k)]^2} \\ &= n^{-1} \frac{y'(I - A(k))'(I - A(k)]^2}{[n^{-1} trace(I - A(k)]^2} \end{aligned} \tag{15}$$

The minimum MSE value indicates that the optimization value is close to the actual value. Assuming a minimum value of MSE, MSE is divided by the value $[n^{-1} trace(I - A(k)]^2$ with the value of $A(k)$ influenced by k. Optimization of knot points is chosen based on small GCV values. Therefore, to obtain an estimate of the optimal spline function, experiments can be carried out on the value $k(0 < k < 1)$ to obtain the optimal minimum GCV value for spline fitting. Optimizing knot points in spline regression will be optimal if the $GCV(k)$ value is relatively small, meaning that the optimization does not change much, ensuring that the model can be balanced.

The GCV equation includes the term $A(k)$, where $A(k)$ is the effective number of independent knot points that depend on the parameter k. This term acts as a penalty for models that are too complex. The more complex the model (the more degrees of freedom), the value of $A(k)$ will approach 1 so that the GCV value will increase. This can reduce the tendency of complex models to obtain low GCV values thereby encouraging the selection of simpler models that can capture data patterns without overfitting.

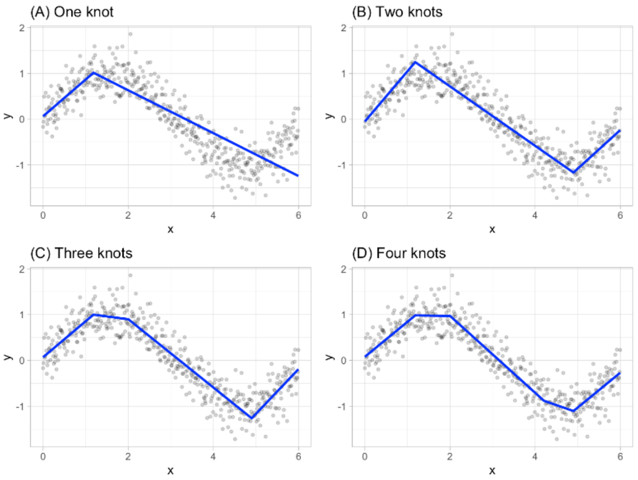


Figure 2. Knot Point Optimization by GCV Method

Knot point optimization by the GCV method in Figure 2 is carried out by calculating the GCV value for each number of knot points in that range and selecting the number of knot points that minimizes the GCV value.

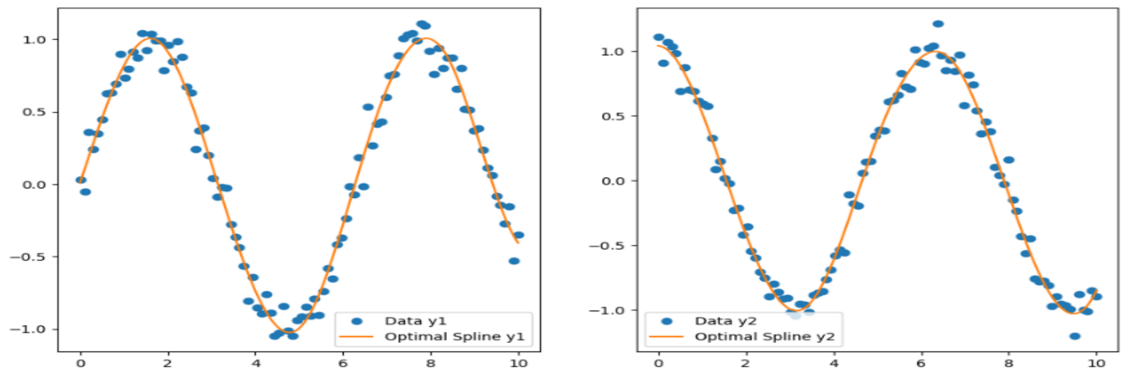
The GCV method will automatically select the optimal number of knot points based on the balance between bias and variance. If the model has too few knot points, the GCV will show a higher value because the model is not flexible enough. If the model has too many knot points, the GCV value will also be high due to overfitting. In other words, the GCV method will ideally find the number of knot points that provide the model with the best predictive performance.

Next, the knot point optimization process with Generalized Cross Validation (GCV) on bi-response spline nonparametric regression is applied using Python to produce the following syntax.

```
import numpy as np
import mathplotlib.pyplot as plt
from scipy.interpolate import LSQUnivariatespline
from sklearn.model_selection import Kfold

# Example of bi-response data
x = np.linspace(0, 10, 100)
y1 = np.sin(x) + 0.1 * np.random.randn(len(x))
y2 = np.cos(x) + 0.1 * np.random.randn(len(x))

# Combine y1 and y2 into one two-dimensional array
y = np.vstack((y1, y2)).T

def. generalized_cross_validation(x, y, num_knots):
    kf = Kfold(n_splits=5, shuffle=True. Random_state=42)
    gcv_errors = [ ]

    for train_index, test index in kf.split(x):
        x_train, x_test = x[train_index], x[test_index]
        y_train, y_test = y[train_index], y[test_index]

        # Determine the knot point (excluding the ends)
        knots = np.linspace(x_train.min ( ), x_train.max( ), num_knots + 2) [1:-1]

        #  Spline fit for each column y
        splines = [LSQUnivariatespline(x_train, y_train[: ,i], t=knots) for i in
                range(y_train.shape[1])]
```

```
            #  Predictions and calculate errors
            y_pred = np.column_stack([spline(x_test) for spline in splines])
            error = np.mean((y_test – y_pred) ** 2)

            gcv_errors.append(error)

        return np.mean(gcv_errors)

# Range of number of knot points to test
num_knots_range = range(5,  16)
gcv_values = [ ]

for num_knots in num_knots_range:
        gcv = generalized_cross_validation(x, y, num_knots)
        gcv_values.append(gcv)

# The number of knot points that minimizes GCV
optimal_num_knots = num_knots_range[np.argmin(gcv_values)]

# Create a spline with the optimal number of knot points
optimal_knots = np.linspace(x.min( ), x.max( ), optimal_num_knots + 2) [1:-1]
optimal_spline = [LSQUnivariatespline(x, y[:, i], t=optimal_knots) for i in range(y.shape[1])]

# Visualization of results
plt.figure(figsize=(12,6))

plt.plot(x, y1, ‘o’, label=’Data y1’)
plt.plot(x, y2, ‘o’, label=’Data y2’)

x_fit = np.linspace(0, 10, 100)
y1_fit = optimal_spliines[0](x_fit)
y2_fit = optimal_splines[1](x_fit)

plt.plot(x_fit, y1_fit, ‘-‘, label=’Optimal Spline y1’)
plt.plot(x_fit, y2_fit, ‘-‘, label=’Optimal Spline y2’)

plt.legend( )
plt.title(f’Optimal Spline with {optimal_num_knots} Knots’)
plt.show( )
```

In previous research entitled "Estimates for The Generalized Cross-Validation Function via An Extrapolation and Statistical Approach", a statistical approach was used to find estimates for GCV. In this article, we develop the GCV method used in nonparametric bi-response spline regression to optimize knot points. Interpretation of the application of knot point optimization is carried out using the Python syntax in the box table above and produces the best knot optimization in the nonparametric bi-response spline regression which can be seen in **Figure 3**.



(a) **(b)**

Figure 3. Optimization of Knot Points on Sin and Cos Graphs, (a) Sin Graph, (b) Cos Graph

Based on **Figure 3**, the GCV method is used to optimize knot points in nonparametric bi-response spline regression (sin and cos graphs). In the graph, it can be seen that the GCV method provides optimal knot points, namely 5 knot points and forms the best model.

According to this explanation, Generalized Cross Validation (GCV) is able to provide optimization at optimal knot points. The GCV process is automatic and does not require complex or manual parameter settings. This makes it easier to implement, especially in settings where determining the optimal number of knot points may be unclear. GCV can effectively optimize knot points for spline fitting, ensuring a model that is balanced and efficient in capturing data patterns without overfitting.

4. CONCLUSIONS

Based on the results and discussion, bi-response spline regression has the weakness that it always depends on optimal knot points. Taking knot points that are not optimal results in the curve resulting in the spline regression model being unable to accurately capture the structure present in the data and can cause poor model performance on unseen test data. Therefore, generalized cross-validation parameter estimation is used to select the most optimal knot points.

ACKNOWLEDGMENT

Thank you to the Softcomputing Mathematics Research Group for their moral support and constructive comments in improving this paper. We also express our gratitude to the Institute for Research and Community Service (LPPM) Universitas Sebelas Maret (UNS) for the research group grant through the letter of approval for research implementation of non-APBN funds No. 228/UN27.22 /PT.01.03/2024.

REFERENCES

[1] B. Mahesh, "Machine Learning Algorithms-A Review," *International Journal of Science and Research*, 2018, doi: 10.21275/ART20203995.

[2] Sifriyani, I.N. Budiantara, and S.H. Kartiko, "A New Method of Hypothesis Test for Truncated Spline Nonparametric Regression Influenced by Spatial Heterogeneity and Application" Hindawi Volume 2018, Article ID 9769150, 13 pages. https://doi.org/10.1155/2018/9769150.

[3] J. Arkes, "Regression Analysis: A Practical Introduction" (2nd ed.), Routledge, 2023, https://doi.org/10.4324/9781003285007

[4] V. Ratnasari, I.N. Budiantara, and A.T.R. Dani, "Nonparametric regression mixed estimators of truncated spline and gaussian kernel based on cross-validation (CV), generalized cross-validation (GCV), and unbiased risk (UBR) methods, Int. J. Adv. Sci. Eng. Inform. Technol. 11 (2021) 2400-2406.".

[5] D. A. Widyastuti, A. A. R. Fernandes, and H. Pramoedyo, "Spline estimation method in nonparametric regression using truncated spline approach," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, May 2021. doi: 10.1088/1742-6596/1872/1/012027.

[6] M. Rosalina, S. Martha, and N. Imro'ah Intisari, "Pemodelan Regresi Nonparametrik Birespon Spline pada Persentase Penduduk Miskin dan Indeks Kedalaman Kemiskinan," 2023.
[7] Y. Guangyu, Z. Baqun, and Z. Min, "Estimation of Knots in Linear Spline Models," Journal of the American Statistical Association, 2021, 118(541), 639–650, https://doi.org/10.1080/01621459.2021.1947307.
[8] D. Krstajic, L. J. Buturovic, D. E. Leahy, and S. Thomas, "Cross-validation pitfalls when selecting and assessing regression and classification models," *J Cheminform*, vol. 6, no. 1, Mar. 2014, doi: 10.1186/1758-2946-6-10.
[9] N. Fitriyani, *Metode Cross Validation dan Generalized Cross Validation dalam Regresi Nonparametrik Spline (Studi Kasus Data Fertilitas di Jawa Timur)*. 2014. [Online]. Available: https://www.researchgate.net/publication/347495233
[10] M. Mitrouli and P. Roupa, "Estimates for the generalized cross-validation function via an extrapolation and statistical approach," *Calcolo*, vol. 55, no. 3, Sep. 2018, doi: 10.1007/s10092-018-0266-3.
[11] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *J Bus Res*, vol. 133, pp. 285–296, Sep. 2021, doi: 10.1016/j.jbusres.2021.04.070.
[12] D. Fitria, A. Husaeni, A. Bayu, and D. Nandiyanto, "Bibliometric Using Vosviewer with Publish or Perish (using Google Scholar data): From Step-by-step Processing for Users to the Practical Examples in the Analysis of Digital Learning Articles in Pre and Post Covid-19 Pandemic," 2021, doi: 10.17509/ijost.v6ix.
[13] A. Kirby, "Exploratory Bibliometrics: Using VOSviewer as a Preliminary Research Tool," *Publications*, vol. 11, no. 1, Mar. 2023, doi: 10.3390/publications11010010.
[14] M. Dubyna, O. Popelo, N. Kholiavko, A. Zhavoronok, M. Fedyshyn, and I. Yakushko, "Mapping the Literature on Financial Behavior: a Bibliometric Analysis Using the VOSviewer Program," *WSEAS Transactions on Business and Economics*, vol. 19, pp. 231–246, 2022, doi: 10.37394/23207.2022.19.22.
[15] W. Haitham, Tuffaha, M. Strong, L. G. Gordon, P. A. Scuffham, "Efficient Value of Information Calculation Using a Nonparametric Regression Approach: An Applied Perspective," Value in Health, Volume 19, Issue 4, 2016, https://doi.org/10.1016/j.jval.2016.01.011.
[16] O. Febriana Rinda Sihotang and S. Prangga, "Aplikasi Regresi Nonparametrik Spline Birespon pada Data Kualitas Air di Das Mahakam", [Online]. Available: http://ejurnal.binawakya.or.id/index.php/MBI
[17] M. Jansen, "Generalized Cross Validation in variable selection with and without shrinkage," *J Stat Plan Inference*, vol. 159, pp. 90–104, Apr. 2015, doi: 10.1016/j.jspi.2014.10.007.
[18] M. Maharani and D. R. S. Saputro, "Generalized Cross Validation (GCV) in Smoothing Spline Nonparametric Regression Models," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1742-6596/1808/1/012053.
[19] Sifriyani, A. R. M. Sari, A. T. R. Dani, and S. Jalaluddin, "Bi-Response Truncated Spline Nonparametric Regression with Optimal Knot Point Selection Using Generalized Cross-Validation in Diabetes Mellitus Patient's Blood Sugar Levels," *Communications in Mathematical Biology and Neuroscience*, vol. 2023, 2023, doi: 10.28919/cmbn/7903.
[20] R. Fenda Refiantoro, C. Rizki Nugroho, and Y. Tri Hapsari, "Analisis Regresi Sederhana Pada Data Nilai UAS Menggunakan Microsoft Excel Dan IBM SPSS Analisis Regresi Sederhana Pada Nilai UAS Menggunakan Microsoft Excel Dan IBM SPSS," Jurnal ARTI : Aplikasi Rancangan Teknik Industri Volume 17 Nomor 2, November 2022. 396-Article Text-2039-1-10-20221129.
[21] P. Laksana and S. Kurnia Viona, "Penggunan Regresi Linear untuk Mengetahui Variabel Pengaruh pada Kekuatan Lereng Highwall Pit Alfa-Pt Arutmin Indonesia Tambang Kintap," *Prosiding Tpt XVIII Perhapi* 2019.
[22] A. Perperoglou, Sauerbrei, M. Abrahamowicz, "A review of spline function procedures in R, " BMC Med Res Methodol 19, 46 (2019). https://doi.org/10.1186/s12874-019-0666-3.
[23] Mariati, N. P. A. Mirah, I. Nyoman Budiantara, and V. Ratnasari, "The Application of Mixed Smoothing Spline and Fourier Series Model in Nonparametric Regression" Symmetry 13, 2021, no. 11: 2094, https://doi.org/10.3390/sym13112094.
[24] I. Nyoman Budiantara, M. Fariz Fadillah Mardianto, M. Putri, I. N. Budiantara, and S. H. Kartiko, "Selection of Optimal Knot Point and Best Geographic Weighted on Geographically Weighted Spline Nonparametric Regression Model Using Generalized Cross Validation and Kernel Function." [Online]. Available: https://ssrn.com/abstract=4811193