Parameter: Journal of Statistics
Available at https://bestjournal.untad.ac.id/index.php/parameter
Vol. 1 No. 2 2021, 31 - 37
DOI: *https://doi.org/10.22487/27765660.2021.v1.i2.15427*
eISSN: 2776-5660

SPLINE NONPARAMETRIC REGRESSION MODEL FOR LOCAL REVENUE IN CENTRAL SULAWESI

Mohammad Fajri[1], Ayu Anggraeni[2], Eka Rizky Wulansari[3], Mufitatul Annisa[4]
[1,2,3,4]Tadulako University, Kota Palu, Sulawesi Tengah

***e-mail**: rizkywulansarieka@gmail.com*

ABSTRACT

Local Own-source Revenue (LOR) is all regional revenue that comes from the region's original economic resources. It is very important to identify it by researching and determining the Regional Local Own-source Revenue (LOR) by properly researching and managing the source of revenue so as to provide maximum results. Central Sulawesi Province itself has Local Own-source Revenue (LOR) in the Regional Revenue and Expenditure Budget of the 2018 Budget Year has reached Rp1 trillion. The increase or decrease in growth of local revenue is influenced by the amount and type of tax, levies collected by local governments and the lack of incentives for the management apparatus to carry out tax collection and levies. This study uses spline regression analysis because the data of the Local Own-source Revenue (LOR) in Central Sulawesi in 2018 does not have a pattern so that it fits perfectly with that method. Then after processing the data obtained the results of spline nonparametric regression modeling using the optimal knots point obtained from the minimum GCV value. The best spline nonparametric regression model is written as follow $\hat{y} = -232.1386 + 78280.27_{x_1} - 49019(x_1 - k_1)_+ + 24556.22\,(x_1 - k_1)_+ - 31100.06(x_1 - k_1)_+ + 3076.348_{x_2} - 48203.6(x_2 - k_1)_+ + 15226.63(x_2 - k_1)_+ - 23529.5(x_2 - k_1)_+ + 15226.63(x_2 - k_1)_+ - 23529.5(x_2 - k_1)_+ + 21463.37_{x_3} - 49358.29(x_3 - k_1)_+ + 48935.15(x_3 - k_1)_+ - 31516.8(x_3 - k_1)_+ - 2323.598_{x_4} + 2463.922(x_4 - k_1)_+ - 91.16495(x_4 - k_1)_+ - 91.69352(x_4 - k_1)_+ + \varepsilon.$ *It can be concluded that in Central Sulawesi in 2018 the lowest Local Own-source Revenue (LOR) value was Banggai Laut Regency with 21,776 billion rupiahs and the highest Local Own-source Revenue (LOR) value was Palu City at 267,402 billion rupiahs.*

Keywords: *Central Sulawesi, GRDP, LOR, Regional Expenditure, Tax.*

Cite: Fajri, M., Anggraeni, A., Wulansari, R. E., & Annisa, M. (2021). Spline Nonparametric Regression Model for Local Revenue in Central Sulawesi. *Parameter: Journal of Statistics, 1*(2), 31-37.

*Received: Jan 2021; Reviewed: Jan 2021; Published: Jun 2021

INTRODUCTION

Based on law Number 32 of 2004 concerning regional autonomy, the rights, authorities and obligations of the autonomous region to regulate and manage government affairs and the interests of local communities in accordance with statutory regulations, then amended for the third time by law number 12 years. 2008, the amendment was made to anticipate developments in the implementation of local governments to follow up on the decisions of the constitutional court.

This is in line with the meaning of fiscal decentralization itself, which implies that regions are given the authority to utilize their own financial sources derived from Local Own-source Revenue (LOR), whose main source is regional taxes and regional levies while still based on reasonable limits. Apart from that, it must also be supported by a financial balance between the center and the regions. This fiscal decentralization expects local governments to have greater independence in regional finances. Therefore, the role of LOR, which is part of Regional Revenue, greatly determines regional financial performance.

Local Own-source Revenue (LOR) according to Halim (2001) is revenue obtained by a region from sources within its own territory which is collected based on regional regulations in accordance with applicable laws and regulations. LOR consists of four types of income, namely local taxes, regional levies, proceeds from regional owned companies and proceeds from the management of separated regional assets and other legal LOR. Local Own-source Revenue (LOR) is revenue from the region itself which consists of; (1) proceeds from regional taxes, (2) proceeds from regional retribution, (3) proceeds from regional-owned companies and proceeds from the management of separated regional assets, (4) other legitimate regional original revenues.

Local Own-source Revenue (LOR) is expected to be the main capital for governance and development, at this time the condition is still inadequate, meaning that the proportion that ROI contributes to Total Regional Revenue (TRR) is still relatively low.

Central Sulawesi Province itself has Local Own-source Revenue (LOR) in the Regional Budget for Fiscal Year 2018 which has reached IDR 1 trillion. The increase or decrease in the growth of original regional income is influenced by the amount and type of taxes, levies collected by local governments and the lack of incentives for the management apparatus in carrying out tax and levies collection. If the number of types of local taxes and levies that are collected is increasing, then the consequence is that the local revenue will be higher (Hibzon, 2007).

Research on Regional Original Income has been carried out by several previous researchers, including Priyono (2016) showing that only the Gross Regional Domestic Product (GRDP) variable has an effect on the realization of Local Own-source Revenue (LOR). For research using spline regression, Wahyuni et al. (2020) used Spline Nonparametric Regression to model the open unemployment rate in Sulawesi showing that the optimal knot is a combination of knots. In this study, it will be investigated about the factors that are thought to affect the Local Own-source Revenue (LOR) in Central Sulawesi using the spline model. The spline method is very good at modeling data that has a changing pattern at certain sub-intervals. Spline is a model that has statistical interpretation and visual interpretation and has a very good ability to be generalized to complex and complicated statistical modeling (Budiantara, 2015). Wahba (1990) provides a method for selecting optimal smoothing parameters in spline estimators, namely Generalized Cross Validation (GCV). There are two problems in this study, namely how the characteristics of Local Own-source Revenue (LOR) in Central Sulawesi and the factors that are thought to influence and how is the modeling of Local Own-source Revenue (LOR) in Central Sulawesi using nonparametric spline regression.

MATERIALS AND METHODS

1. Data Sources

The data used in this research is secondary data obtained from the publication of BPS 2018 , with the research units being observed are districts / cities in Central Sulawesi, which consists of 13 administrative regions with 12 districts and 1 city.

2. Research Variables

The research variables used include Local Own-source Revenue (y) as a response variable and Constant Gross Regional Domestic Product Price (x_1), Current Gross Regional Domestic Product (x_2), Area (x_3), and Population (x_4) as predictor variables.

3. Methods

The method used in this research is Spline Nonparametric Regression. Spline is a segmented polynomial which has flexibility properties. The spline is very dependent on the point of the knots. Knots point is a joint point where there is a pattern of changing the behavior of a function at different intervals (Hardle, 1990). In general, the function G in a spline space of order m with knots k_1, k_2,…, k_j is any function that can be expressed as Equation (2).

$$G(x_i) = \sum_{j=0}^{m} \beta_j x_i^j + \sum_{k=1}^{J} \beta_{k+m}(x_i - K_k)^m$$

With,

$$(x_i - K_k)_+^m = \begin{cases} (x_i - K_k)^m, & x_i \geq K_k \\ 0, & x_i < K_k \end{cases}$$

β is the model parameters where m is the spline order. The estimation of the spline regression model parameters can be stated as follows.

$$y_i = G(x_i) + \varepsilon_i$$

4. Data analysis

Data analysis in this research is as follows:

1. Make the descriptive statistics of each variable to determine the characteristics of each regency/city in Sulawesi Tengah.
2. Make a scatter plot between Local Own-source Revenue (y) with each predictor variable.
3. Modeling the Local Own-source Revenue in the Central Sulawesi by using a linear spline with a few knots point. Which the modeling is

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{1x_1} + \hat{\beta}_2(x_1 - k_1)_+ + \hat{\beta}_{3x_2} + \hat{\beta}_4(x_2 - k_2)_+ + \hat{\beta}_{5x_3} + \hat{\beta}_6(x_3 - k_3)_+ + \hat{\beta}_{7x_4} + \hat{\beta}_8 (x_3 - k_3)_+ + \varepsilon$$

4. Selecting the optimal knots point based on the minimum GCV. The formula of GCV is

$$GCV\left(\vec{k}\right) = \frac{MSE\left(\vec{k}\right)}{\left(n^{-1}tr\left(I - S\left(\vec{k}\right)\right)\right)^2}$$

5. Modeling the Local Own-source Revenue in Central Sulawesi with variables predictor using a spline with knots optimal.
6. Testing the significance of parameters. There are two tests used, namely simultaneous test and individual test.
   a. Simultaneous Test

$$F_{hit} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2/p}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2/(n-p-1)}$$

   b. Individual Test

$$t_{hit} = \frac{\left(\hat{\beta}_j\right)^2}{SE\left(\hat{\beta}_j\right)^2}$$

7. Calculate The R^2.
8. Make a conclusion.

RESULTS AND DISCUSSION

1. The characteristics of the Local Own-source Revenue (LOR) and Suspected Factors

The characteristics of the Local Own-source Revenue (LOR) along with the factors which allegedly affect in the Central Sulawesi Province is shown in table 1.

Table 1. Local Own Income (PAD) and the Factors Suspected to Affect

Variable	Means	Varians	Minimum	Maximum
y	96341153.85	4.76248E+15	21776000	2.67E+08
x_1	11552.62846	55817724.42	2202.56	26876.73
x_2	8087.711538	26408761.39	1633.8	18372.03
x_3	4764.714615	7935114.651	395.06	10004.28
x_4	231572.5385	14229623585	73697	482794

2. Scatterplot between Local Own-source Revenue (LOR) and the suspected factors

The pattern of the relationship that is formed between Local Own-source Revenue (LOR) (y) which is the response variable with x_1, x_2, x_3, and x_4 is visualized in Figure 1.Figure 1. shows the pattern of the relationship formed between Local Own-source Revenue (LOR) and the four variables does not form a particular pattern. This indicates that there is a nonparametric component where the function of the regression curve is unknown.

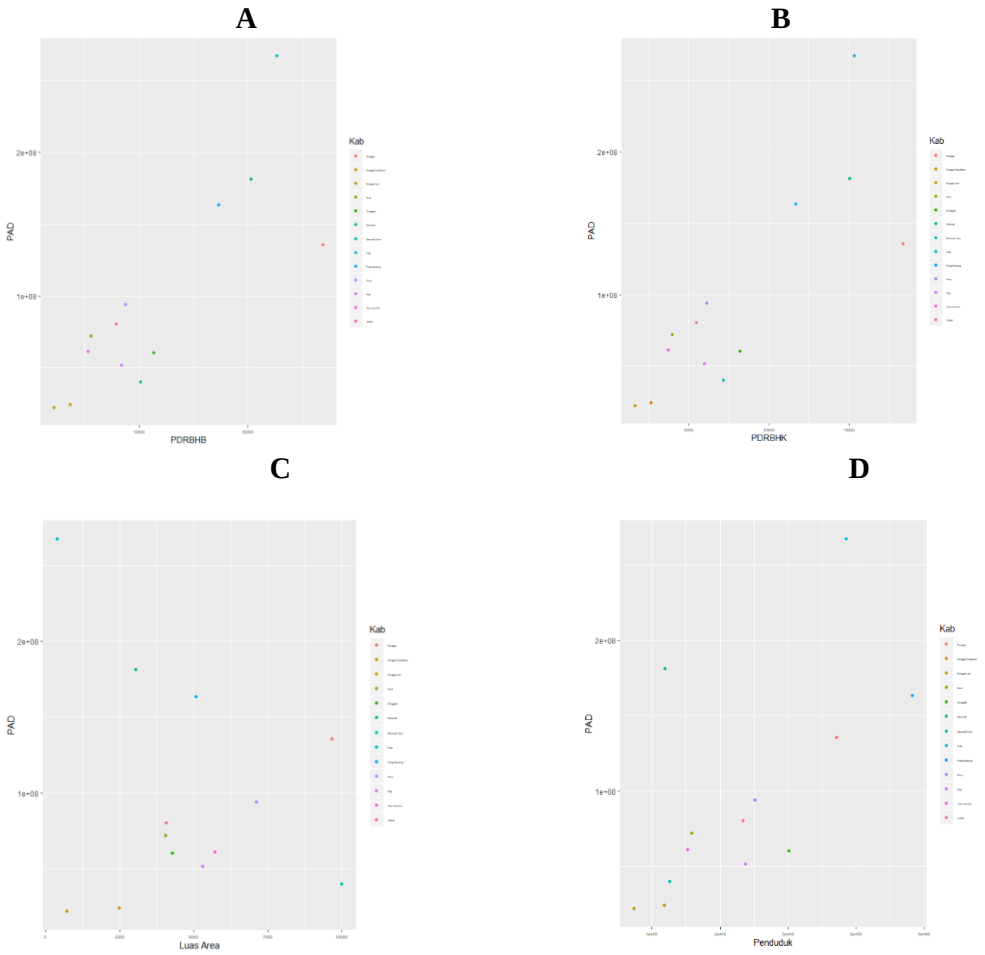

Figure 1. Scatterplot Between Local Own-source Revenue (LOR) with Predictor Variables X_1, X_2, X_3, and X_4. **A**. Scatterplot Between Local Own-source Revenue (LOR) with Predictor Variable X_1; **B**. Scatterplot Between Local Own-source Revenue (LOR) with Predictor Variable X_2; **C**. Scatterplot Between Local Own-source Revenue (LOR) with Predictor Variable X_3; **D**. Scatterplot Between Local Own-source Revenue (LOR) with Predictor Variable X_4.

3. Spline Nonparametric Regression Model

The nonparametric spline regression model for 1 knot point is shown in Equation (1).

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{1x_1} + \hat{\beta}_2(x_1 - k_1)_+ + \hat{\beta}_{3x_2} + \hat{\beta}_4(x_2 - k_2)_+ + \hat{\beta}_{5x_3} + \hat{\beta}_6(x_3 - k_3)_+ + \hat{\beta}_{7x_4} + \hat{\beta}_8(x_3 - k_3)_+ + \varepsilon \quad (1)$$

4. The Selection Of The Optimal Knots Point

The selection of the optimal knots point is done by finding the value of the GCV the best that is produced. GCV, which is produced by using 1 point knots, 2 knots point, 3 point knots, and the combination of knots are shown in Table 2.

Table 2. GCV For 1 Point Knot, 2 Point Knots, 3 Point Knots, And Combination Of Knots

Number	Knot	GCV
1	1 Point Knot	4.54169E+14
2	2 Point Knots	5.9707E+12
3	3 Point Knots	-1.66878E+34
4	Combination of Optimum Knots and Minimum GCV	-1.66878E+34

Table 2 shows that the minimum GCV value generated when using a combination of optimum knots and minimum GCV is -1.66878E+34.

5. Combination of Optimum Knots and Minimum GCV

The minimum GCV value from the values of GCV using 1 point knot, 2 point knots, 3 point knots, and the combination of knots is generated when using the combined knot which is -1.66878E + 34. Modeling of Local Own-source Revenue using the optimal knot point is shown in Equation (2) below.

$$\hat{y} = -232.1386 + 78280.27_{x_1} - 49019(x_1 - k_1)_+ + 24556.22(x_1 - k_1)_+ - 31100.06(x_1 - k_1)_+ + 3076.348_{x_2} - 48203.6(x_2 - k_1)_+ + 15226.63(x_2 - k_1)_+ - 23529.5(x_2 - k_1)_+ + 21463.37_{x_3} - 49358.29(x_3 - k_1)_+ + 48935.15(x_3 - k_1)_+ - 31516.8(x_3 - k_1)_+ - 2323.598_{x_4} + 2463.922(x_4 - k_1)_+ - 91.16495(x_4 - k_1)_+ - 91.69352(x_4 - k_1)_+ + \varepsilon \quad (2)$$

6. Test Parameters

There are two test parameter estimation is performed, namely test parameters simultaneously and test individually. The results of the test estimation of the parameters simultaneously is presented in the form of a table which can be seen in Table 3.

Table 3. ANOVA

Source of Variation	df	SS	MS	F
Regression	16	6.191221E+16	3869512818494835	6956287455.1782
Residual	-4	-2225045	556262.2016	
Total	12	6.191221E+16		

The statistical value of the F test is 6956287455.1782. When the F-count is compared with the F-table, a decision can be made to reject H_0. So it can be concluded that there is at least one variable that has a significant influence on the model. The occurrence of reject H_0 indicates that an individual test is needed to determine which variables have a significant effect on the model. The individual test results are presented in Table 4.

Table 4. Individual Test

Variable	Parameters	Estimation	t_{count}	*P-value*
Constant	β_0	-232.1386	-3780.12	7.912805E-38
x_1	β_1	78280.27	36641.85	1.149928E-49
	β_2	-49019.15	-8752.196	3.334223E42
	β_3	24556.22	4089.943	3.074613E-38
	β_4	-31100.06	-9379.064	1.453735E-42
x_2	β_5	3076.348	915.0695	1.953961E-30
	β_6	-48203.6	-7898.982	1.141683E-41
	β_7	15226.63	2695.552	4.577447E-36
	β_8	-23529.5	-2854.424	2.302357E-36
x_3	β_9	21463.37	3335.291	3.554612E-37
	β_{10}	-49358.29	-6900.353	5.780208E-41
	β_{11}	48935.15	27750.54	3.229549E-48
	β_{12}	-31516.8	-32603.51	4.669028E-49
x_4	β_{13}	-2323.598	-21177.62	8.277008E-47
	β_{14}	2463.922	21901.86	5.52876E-47
	β_{15}	-91.16495	-1339.922	2.011076E-32
	β_{16}	-91.69352	-884.3396	2.943972E-30

Table 4 shows that there are 16 parameters that produce a p-value less than the significance level used at 0.05, namely the parameters of the variable Gross Regional Domestic Product at Constant Prices (PDRBHK), Gross Regional Domestic Product current Prices (PDRBHB), Area and Population. So it can be concluded that all of these variables have a significant effect on the model.

7. Coefficient of Determination

The coefficient of determination reached 99.18%. This shows that the model formed is suitable to be used to model data patterns. This means that 99.18% of the variable Gross Regional Domestic Product at Constant Prices, Gross Regional Domestic Product current Prices, Area and Population can explain the variable Local Own-source Revenue (LOR).

CONCLUSION

The results of spline nonparametric regression modeling using the optimal knot point were obtained from the minimum GCV value. The best spline nonparametric regression model is written as follows

$$\begin{aligned}\hat{y} = -232.1386 &+ 78280.27_{x_1} - 49019(x_1 - k_1)_+ + 24556.22(x_1 - k_1)_+ - 31100.06(x_1 - k_1)_+ \\ &+ 3076.348_{x_2} - 48203.6(x_2 - k_1)_+ + 15226.63(x_2 - k_1)_+ \\ &- 23529.5(x_2 - k_1)_+ + 21463.37_{x_3} - 49358.29(x_3 - k_1)_+ + 48935.15(x_3 - k_1)_+ \\ &- 31516.8(x_3 - k_1)_+ - 2323.598_{x_4} + 2463.922(x_4 - k_1)_+ - 91.16495(x_4 - k_1)_+ \\ &- 91.69352(x_4 - k_1)_+ + \varepsilon\end{aligned}$$

It can be concluded that in Central Sulawesi in 2018 the lowest Local Own-source Revenue (LOR) value was Banggai Laut Regency with 21,776 billion rupiahs and the highest Local Own-source Revenue (LOR) value was Palu City at 267,402 billion rupiahs.

REFERENCES

Halim, A. (2001). *Bunga rampai manajemen keuangan daerah*. *First edition.* Yogyakarta: UPP AMP YKPN.

Badan Pusat Statistik (BPS). (2018). Sulteng Dalam Angka 2018. Publikasi BPS, Sulawesi Tengah.

Budiantara, I. N., Ratnasari, V., Ratna, M., & Zain, I. (2015). The combination of spline and kernel estimator for nonparametric regression and its properties. *Applied Mathematical Sciences*, *9*(122), 6083-6094.

Härdle, W. (1990). *Applied nonparametric regression* (No. 19). Cambridge university press.

Hibzon, H. (2007). Analisis Faktor-faktor yang Mempengaruhi Pendapatan Asli Daerah Sumatera Selatan. *Jurnal Ekonomi Pembangunan*, *5*(2), 106-116.

Priyono, M. M. (2016). *Metode penelitian kuantitatif*. Taman Sidoarjo: Zifatama.

Undang-Undang Republik Indonesia Nomor 32 Tahun 2004 tentang Pemerintahan Daerah.

Undang-Undang Republik Indonesia Nomor 12 Tahun 2008 tentang perubahan Kedua Atas UU Nomor 32 Tahun 2004 tentang Pemerintahan Daerah.

Wahba, G. (1990). *Spline models for observational data*. Society for industrial and applied mathematics.

Wahyuni, S. A., Ratnawati, R., Indriyani, I., & Fajri, M. (2020). Spline Regression Analysis to Modelling The Open Unemployment Rate in Sulawesi. *Natural Science: Journal of Science and Technology*, *9*(2), 34-39.