Engineering and Technology Journal e-ISSN: 2456-3358
Volume 09 Issue 10 October-2024, Page No.- 5285-5292
DOI: 10.47191/etj/v9i10.04, I.F. – 8.227

Estimation of the Unemployment Rate in Indonesia Using Nonparametric *Spline Regression*

Muhammad Marizal[1], Widya Tamaka Yofi[2]
[1,2] Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau
Jl. HR. Soebrantas No. 155 Simpang Baru, Panam, Pekanbaru, 28293

ABSTRACT: The COVID-19 pandemic impact is not only on health but on all aspects of life, one of the socio-economic aspects is unemployment. The linear regression be used as a method in modeling the unemployment rate frequently. Whereas linear regression is inappropriate for data that have trends or irregular relationship patterns such as unemployment rate data. Therefore, this research used of the nonparametric spline regression method. It can produce more flexible curves according to the flatline of data based on optimum knots at the optimum value of the cross validation. There are several factors suggested that affect the Unemployment Rate (TPT), including the number of positive COVID cases, labor force participation rate, economic growth and human development index. The purpose of this study is to find out if there is a difference between the open unemployment rate (TPT) in Indonesia before and during the pandemic. The Data used are secondary data obtained from Badan Pusat Statistika (BPS). The calculations results show that there are differences between the models of the unemployment rate before and during the pandemic. Generally there are three significant variables in the before the pandemic model: variables of economic growth (PE) and the human development index (IPM), while during the pandemic the significant variable are the labor force participation rate (TPAK) and economic growth (PE).

KEYWORDS: COVID-19, Generalized Cross Validation, Open unemployment rate, Optimum Knot, Spline Nonparametrics Regression.

1. Introduction

The impact of the COVID-19 pandemic is not only about health but all aspects of life, one of which is on the socio-economic aspect. One of the impacts of the COVID-19 pandemic on this aspect is the increase in the unemployment rate. In research [1] said that one in four workers in Indonesia is vulnerable to becoming unemployed due to COVID-19. This can also be proven from the results of a survey by the Central Statistics Agency (BPS) [2] which shows that the increase in Indonesia's Open Unemployment Rate (TPT) in August 2020 was 7.07 percent, an increase of 1.84 percent compared to August 2019.

By analyzing the available data, many researchers try to reveal what factors affect the unemployment rate. However, it is very unfortunate if the number of COVID-19 cases is not reviewed as one of the factors. For example, in research [3], the factors studied are economic growth and population growth. Likewise in research [4] and [5], does not make the number of COVID-19 cases one of the variables. Even though the addition of COVID-19 variables is urgently needed [6] Because it is considered to have a large percentage as the cause of the high unemployment rate.

In addition to variables, the methods used also need special attention. The use of the right method can minimize the resulting errors. Linear regression is a method that is often used to get a model of a problem, including unemployment. Like research [7] Using linear regression analysis among the research methods used to find out the factors that affect the unemployment rate in the five ASEAN countries. However, linear regression is not appropriate to use for data that has an irregular trend or relationship pattern such as unemployment rate data.

Therefore, researchers are interested in using the spline nonparametric regression method. Splines are polynomial pieces that have the property of being divided into several parts and continuous (*truncated*). Nonparametric regresi *Spline* is used because it tends to follow where the data pattern is. This advantage occurs because there are knots that indicate a change in the data pattern around the point [8]. With this, nonparametric regression *Spline* It is more flexible because it can adjust to the characteristics of the data.

Based on the explanation above, the author is interested in conducting research on the factors that affect the unemployment rate in Indonesia before and during the COVID-19 outbreak using the spline nonparametric regression method. Therefore, in general, the author raises the title "Estimation of the Unemployment Rate in Indonesia Using Nonparametric *Spline* Regression". Researchers hope that these results can be additional information in setting

policies to reduce Indonesia's unemployment rate in the era of the pandemic crisis.

2. RESEARCH METHODS

2.1 Theoretical Foundations

a. Compare *Mean* Analysis

Average comparison (*compare mean*) used to compare averages between two or more data samples. In the case of nonparametric, one of the tests for average comparison is the sign test. The mark test is a test to find out the average difference between two samples by paying attention to whether there is a difference or not from an observation [9]. With the following conditions [10]:

a. Hypothesis formulation

H_0: The average in each experiment was the same.

H_1: The average in each experiment is different.

b. Statistical Test

$$z_{count} = \frac{(x \pm 0{,}5) - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \quad (1)$$

where if used when and used when $.x + 0{,}5x <$ $\frac{1}{2}\, nx - 0{,}5x >$ $\frac{1}{2}\, n$

c. Results

If the value $Z_{count} > Z_{table}$is then subtracted. On the other hand, if the value $H_0 Z_{Count} >$ Z_{table}then accept. H_0

b. Regresi Nonparametric *Spline*

Nonparametric regresi *spline truncated* is a regression that produces a segmented/truncated polynomial function [11]. It is said to be segmented because there are knots that divide functions according to the pattern. This segmented nature provides more flexibility than ordinary polynomials, making it possible to adapt more effectively to the local characteristics of a function or data [8].

The general form of the p-degree *spline* function is as follows:

$$y = \sum_{j=0}^{p} \beta_j x_i^j + \sum_{k=1}^{r} \beta_{p+k}(x_i - K_k)_+^p + \varepsilon_i \,; i = 1,2,\dots,n. \quad (2)$$

by being a real constant, andβ_j

$$(x_i - K_k)_+^p = \begin{cases} (x_i - K_k)^p ; x_i \geq K_k \\ 0 \quad\quad\quad ; x_i < K_k \end{cases} \quad (3)$$

The knot point is represented by K_k. If p = 1 then it is said to be a linear *spline*, p = 2 then it is said to *be a quadratic spline and p = 3 then it is said* to be a cubic *spline*.

Parameter stimulation on nonparametric regression methods *Spline* using the least squares method or *Ordinary Least Square* (OLS), i.e. minimizing the number of residual squares [12]. The OLS formula is as follows.

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{y} \quad (4)$$

After obtaining the parameter estimation, the next step is to determine the optimal knot point. It is said to be a nonparametric regression model *Spline* best if it has an optimal knot point. Method *Generalized Cross Validation* (GCV) is used to select the optimal knot point by looking at the smallest value of the method's results. The formula of the GCV method, namely [13]:

$$GCV(k) = \frac{MSE(k)}{[n^{-1}trace(\boldsymbol{I}-\boldsymbol{A})]^2} \quad (5)$$

where n is the number of observations, I is the identity matrix, is the knot points,$k = (k_1, k_2, \dots, k_r)$ $MSE(k) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2$ and.$A = \boldsymbol{X}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}$

c. Signifikan test

1. Simultaneous Testing

Parameter testing is carried out simultaneously to see the influence of independent variables simultaneously/as a whole on dependent variables [14]. The test was carried out with the following provisions.

a. Hypothesis formulation

$H_0: \beta_1 = \beta_2 = \cdots = \beta_{p+k} = 0.$

H_1: at least one $\beta_{p+k} \neq 0$.

where the value of p + k is the number of independent variables in the nonparametric regression *of the spline.*

b. Statistical Test

$$F_{hitung} = \frac{Rataan\ Kuadrat\ Regresi\ (RKR)}{Rataan\ Kuadrat\ Eror\ (RKE)} \quad (6)$$

with $RKR = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{p+k}$ and $RKE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-(p+k)-1}$.

c. Withdrawal of decision

Reject the if or $H_0 F_{hitung} > F_{tabel(n-(p+k)-1)}$p-value$< \alpha$, so it is concluded that there is at least one predictor variable that has a significant effect on the response variable.

2. Partial Testing

Partial testing was carried out to see the significance of the influence of individual independent variables on dependent variables [14]. Partial testing is carried out under the following conditions:

a. Hypothesis formulation

$H_0 : \beta_j = 0$

H_1 : with.$\beta_j \neq 0 j = 1, 2, \dots,$ p + k

b. Statistical Test

$$t_{hitung} = \frac{b_j - \beta_j}{SE(b_j)} \quad (7)$$

where is the j-th variable coefficient, is the hypothetical j-parameter of the hypothesis, and is the standard error.$b_j \beta_j SE(b_j) b_j$

c. Withdrawal of Decision

If the value or $t_{hitung} > t_{tabel}$p-value $< \alpha$can be withdrawn, a rejection decision can be

drawn. On the other hand, if the value or $H_0 t_{hitung} < t_{tabel}$p-value $> \alpha$then accept.H_0

d. Coefficient of determination

The determination coefficient is a quantity to measure the feasibility of the model (*goodness of fit*). The value of R2 has an interval from 0 to 1 ($0 \leq R^2 \leq 1$). The lower the value, the more the independent variable cannot explain the variability and dependent variable [12]. Here is the formula to get the value R2.

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{Jumlah\ Kuadrat\ Regresi\ (JKR)}{Jumlah\ Kuadrat\ Total\ (JKT)} \quad (8)$$

with y is the observation value, is the average value of the observation and is the prediction value using the regression model.$\bar{y}\hat{y}y$

2.2 Research Methods/Methodologies

The method used in this study is a quantitative method with a secondary data approach where the author only processes the data that is already available. Data was obtained from the website and publications of the Central Statistics Agency (BPS) from each province. In addition, the author also obtained data from the monthly publication of the COVID-19 Task Force. The authors use data from 2019 and 2020 as research data to represent the years before and during the pandemic crisis.

The sampling technique used is Total Population Sampling or saturated sampling, which is a sampling technique that makes all members of the population as samples [15]. The population as well as the research sample is all provinces in Indonesia as many as 34 provinces. In this study, there are four variables that are suspected to affect the Open Unemployment Rate (TPT), namely the Number of COVID-19 Cases (COV), Labor Force Participation Rate (TPAK), Economic Growth (PE), and Human Development Index (HDI).

The data that has been compiled is then analyzed *compare mean*. Then continued with modeling of the nonparametric regression method *spline.* As for the classical assumption test, the nonparametric regression method does not have assumptions that must be met [11]. Steps of analysis of nonparametric regression methods *Spline* are as follows:

1. Model the dependent variable using the spline nonparametric regression method with 1, 2, 3 and a combination of knot points.
2. Selects the optimal knot point based on the lowest GCV value.
3. Obtain the best spline *regression model* with optimal knot point.
4. Conduct parameter significance tests.
5. Calculate the determination coefficient.
6. Explain the interpretation of the model and draw conclusions.

3. RESULTS AND DISCUSSION

3.1 Descriptive statistics

Below is a table that provides information about descriptive statistics from research variables in the year before the pandemic (2019) and during the pandemic (2020).

Table 1. Descriptive Statistics of Research Variables Before & During the Pandemic

	Variable	*Minimum*	*Maximum*	*Mean*	*Std. Deviation*
2019	**TPT**	1.52	8.11	4.7691	1.50853
	TPAK	62.90	76.92	67.6506	3.10996
	ON	26561	1836198,49	325584,0962	469131.6831
	IPM	60.84	80.76	71.0403	3.91320
2020	**TPT**	3.32	10.95	6.0335	2.01313
	COV	1868	175926	20981,32353	34680.13216
	TPAK	63.40	74.32	68.0432	2.91219
	ON	27868	1792794.59	319195.8515	457943.7623
	IPM	60.44	80.77	71.0809	3.90188

From Table 1, there are several points obtained. The variables of the Open Unemployment Rate (TPT), Labor Force Participation Rate (TPAK) and Human Development Index (HDI), both in 2019 and 2020, had a standard deviation value smaller than the average value. This indicates that the variable tends to be homogeneous so that the average value can be used as a representation of the entire data.

Different variables such as the Number of COVID-19 Cases and Economic Growth, with a standard deviation value greater than the average means that the data is heterogeneous. It is undeniable that the number of COVID-19 cases in each region is different, the value of the number will depend on the large number of residents, the community's obedience to health protocols and government policies. Meanwhile, the data on economic growth variables are heterogeneous because each region has a different pattern of economic growth [16].

3.2 *Compare Mean*

To find out whether the Open Unemployment Rate (TPT) data in 2019 and 2020 has a difference, a sign test is carried out. With the help of SPSS (*Statistical Package*

Social Science) Software, the following results were obtained.

Table 2. Marking Test Results

	N
Negative Changes (TPT2020 < TPT2019)	0
Positive Change (TPT2020 > TPT2019)	34
Ties (TPT2020 = TPT2019)	0
Total	34
z-value	-5.659
Asymp. Sig. (2-tailed)	0.000

Asymp *Value.Sig*. in Table 4.3. can be a benchmark in making decisions. With an *Asymp value. Sig*. which is less than 0.05 () means that there is a difference between the Open Unemployment Rate (TPT) data in 2019 and 2020.$\alpha = 5\%$

3.3 Regresi Nonparametric *Spline*

To visualize the data, the author uses a scatter plot (*Scatter Plot*). A scatter chart is a graph used to see a pattern of relationships between 2 variables.

a. Pre-Pandemic Data Analysis

The selection of the best model is done by looking at the lowest GCV value. Here is a table containing the GCV values of *the single-knot, two-knot, three-knot and knot-combination* spline nonparametric regression models.

Table 3. Optimal Knot Point Based on Pre-Pandemic GCV Value

Number of Knotts	GCV	Knot Point		
		x_1	x_2	x_3
1	1.00833	66.90571	543.6003	66.53143
2	0.80685	65.4751	358.94340	64.49878
		66.04735	432.80615	65.31184
3	0.73275	65.18898	322.01202	64.09224
		65.47510	358.94340	64.49878
		65.76122	395.87477	64.90531
Combination	0.70210	65.4751	322.012	64.49878
		66.04735	358.9434	65.31184
			395.8748	

Based on Table 3, it was obtained that the model with the knot point combination had the lowest GCV value. Therefore, the best model is a spline nonparametric regression model with a combination of knots (2,3,2).

As described in Section 2.1, there are two tests to determine the significance of the parameters, namely the overall and partial tests. With the help of *R Software* , the results for the overall test are as follows.

Table 4. ANOVA Data Before the Pandemic

Source of Variation	*df*	*SS*	*MS*	F_{hitung} $(pvalue)$
Regression	10	63.30144	6.330144	12.24045 (0.0000005)
Error	23	11.89444	0.5171495	
Total	33	75.19588		

Based on Table 4, it is seen that the value is 12.24045, the value $F_{hitung} p - value < 0.05$, so it is concluded that it is negative. This means that there are variables that have a significant effect on Indonesia's TPT before the pandemic.H_0

Table 5. Partial Testing of Pre-Pandemic Data

Variable	Parameter	Estimator	t_{hitung}	$pvalue$	Information
Konstan	β_0	0.285537	6.53532	0.000001	Signifikan
x_1	β_{11}	-0.32138	-1.20621	0.239998	Insignificant
	β_{12}	-1.90115	-1.99273	0.058288	Insignificant

	β_{13}	1.942445	2.41089	0.024303	Signifikan
x_2	β_{21}	0.003443	1.56312	0.13168	Insignificant
	β_{22}	-0.28397	-2.38325	0.025805	Signifikan
	β_{23}	0.581074	2.53010	0.018707	Signifikan
	β_{24}	-0.30119	-2.68190	0.013314	Signifikan
x_3	β_{31}	0.461407	1.69429	0.10371	Insignificant
	β_{32}	-5.07054	-3.74689	0.001052	Signifikan
	β_{33}	4.634455	3.84540	0.000825	Signifikan

Table 5. shows several parameters that are not significant to the open unemployment rate in Indonesia during the pandemic due to $p - value > 0.05$the value of . In general, there are two significant variables, namely the Economic Growth (PE) variable and the Human Development Index (HDI).

b. Pre-Pandemic Data Analysis

The selection of the best model is done by looking at the lowest GCV value. Here is a table containing the GCV values of *the single-knot, two-knot, three-knot and knot-combination* spline nonparametric regression models.

Table 6. Optimal Knot Point Based on GCV Value During Pandemic

Number of Knotts	GCV	Knot Point			
		x_1	x_2	x_3	x_4
1	2.36552	5.42020	63.62286	63.88691	60.85490
2	2.13049	90.6731	68.97143	928.34075	70.81245
		119.09073	70.75429	1216.49203	74.13163
3	1.54239	16.07682	64.29143	171.94364	62.09959
		19.62902	64.51429	207.96255	62.51449
		23.18122	64.73714	243.98146	62.92939
Combination	1.35319	16.07682	64.29143	171.9436	70.81245
		19.62902	64.51429	207.9625	74.13163
		23.18122	64.73714	243.9815	

Based on Table 6, it was obtained that the model with the knot point combination had the lowest GCV value. Therefore, the best model is a spline nonparametric regression model with a knot combination (3,3,3,2).

As described in Section 2.1, there are two tests to determine the significance of the parameters, namely the overall and partial tests. With the help of *R Software* , the results for the overall test are as follows.

Table 7. ANOVA Data During the Pandemic

Source of Variation	df	SS	MS	F_{hitung} (*pvalue*)
Regression	15	119.4533	7.963554	9.976823 (0.0000007)
Error	18	14.3677	0.7982054	
Total	33	133.821		

Based on Table 7. It is seen that the value is 9.976823 and the value, so it is concluded that the rejection . This means that there are variables that have a significant effect on Indonesia's TPT during the pandemic.$F_{hitung} pvalue < 0.05 H_0$

Table 8. Partial Testing of Data During the Pandemic

Variable	Parameter	Estimator	t_{hitung}	$pvalue$	Information
Konstan	β_0	-1.9474	-2.86086	0.01039	Signifikan
x_1	β_{11}	0.053039	0.845072	0.40916	Insignificant
	β_{12}	1.025495	1.556474	0.13700	Insignificant
	β_{13}	-2.18317	-1.93335	0.06908	Insignificant
	β_{14}	1.149414	2.24875	0.03729	Signifikan
x_2	β_{21}	0.246464	2.684907	0.01513	Signifikan
	β_{22}	13.45078	2.736336	0.01356	Signifikan

	β_{23}	-25.5035	-2.86651	0.01026	Signifikan
	β_{24}	11.44518	2.544867	0.02032	Signifikan
x_3	β_{31}	0.001795	0.314349	0.75687	Insignificant
	β_{32}	0.901881	2.310633	0.03291	Signifikan
	β_{33}	-1.82581	-2.33936	0.03105	Signifikan
	β_{34}	0.921091	2.345373	0.03067	Signifikan
x_4	β_{41}	-0.12286	-1.4238	0.17161	Insignificant
	β_{42}	0.737548	1.911052	0.07206	Insignificant
	β_{43}	-0.8893	-1.61283	0.12418	Insignificant

Table 8. shows several parameters that are not significant to the open unemployment rate in Indonesia during the pandemic due to the value of $pvalue > 0.05$. A variable that did not have a significant effect on the open unemployment rate in Indonesia during the pandemic was the Human Development Index (HDI).

Based on the calculation above, the best model for both years is as follows.

$$TPT = 0.285537 - 0.32138TPAK - 1.90115(TPAK - 65.4751)_+^1 + 1.942445\,(TPAK - 66.04735)_+^1 + 0.003443PE - 0.28397(PE - 322.012)_+^1 + 0.581074(PE - 358.9434)_+^1 - 0.30119(PE - 395.8748)_+^1 + 0.461407IPM - 5.07054(IPM - 64.49878)_+^1 + 4.634455(IPM - 65.31184)_+^1 \quad (9)$$

$$TPT = -1.9474 + 0.053039COV + 1.025495(COV - 16.07682)_+^1 - 2.18317(COV - 19.62902)_+^1 + 1.149414(COV - 23.18122)_+^1 + 0.2464TPAK + 13.45078(TPAK - 64.29143)_+^1 - 25.5035(TPAK - 64.51429)_+^1 + 0.001795PE + 0.901881(PE - 171.9436)_+^1 - 1.82581(PE - 207.9625)_+^1 + 0.921091(PE - 243.9815)_+^1 - 0.12286IPM + 0.737548(IPM - 70.81245)_+^1 - 0.8893(IPM - 74.13163)_+^1 \quad (10)$$

The value of the determination coefficient (R^2) for the pre-pandemic model was 0.8418206 or 84.18206 %. This indicates that the model explains 84.18206% of the variation of TPT or in other words, independent variables can explain TPT in Indonesia before the pandemic of 84.18206% while the remaining 15.81794% is explained by other variables. As for the value of the determination coefficient (R^2) for the model during the pandemic of 0.892635 or 89.2635% by interpreting it the same as R^2 the pre-pandemic value.

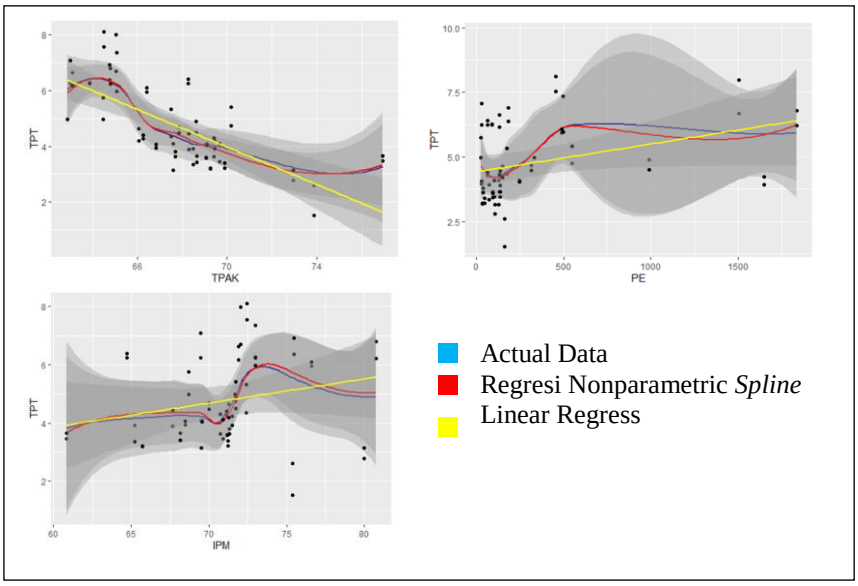

Figure 1. Visualization of Actual Data and Estimates Before the Pandemic

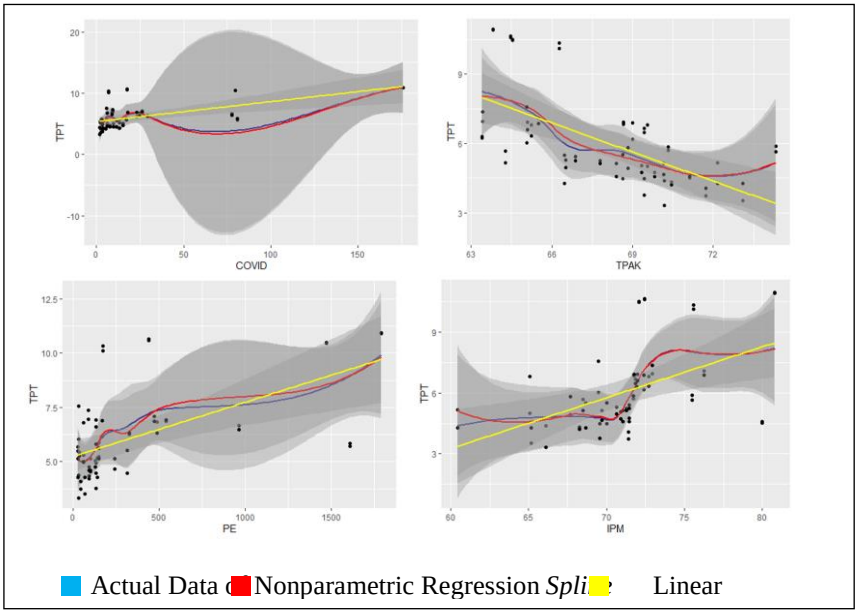

Figure 2. Visualization of Actual Data and Estimates During the Pandemic

4. CONCLUSION

Based on the results of the comparison *mean analysis*, it was found that the average Open Unemployment Rate (TPT) in Indonesia before the pandemic was different from the average Open Unemployment Rate (TPT) during the pandemic, which means that there is a difference in the TPT value in Indonesia before and during the pandemic. From the results of the selection of optimal knots from data before and during the pandemic, the minimum GCV value is found in the number of combined knots. This means that the best modeling with the nonparametric regression method *of the spline* produced is a *spline* with a combination of knots (2,3,2) for before the pandemic and a combination of knots (3,3,3,2) for during the pandemic. So that the best model for the year before and during the pandemic is equations (9) and (10).

Based on the results of the research that has been discussed in Chapter IV, of the three independent variables in the case before the pandemic that have been tested, in general there are two significant variables, namely the Economic Growth (PE) variable and the Human Development Index (HDI), while in the case of the pandemic the significant variables for TPT in Indonesia are the variables of Labor Force Participation Rate (TPAK) and Economic Growth (PE).

For future research, to see the influence of the variable number of COVID positives on the Open Unemployment Rate (TPT), an analysis method can be used that can reveal the form of direct or indirect relationship. Furthermore, by adding independent variables that are more influential and can add knots to make the model better.

REFERENCES

1. R. R. Samudra and D. Setyonaluri, "Inequitable Impact of COVID-19 in Indonesia: Evidence and Policy Response," *Unesco*, p. 15, 2020.
2. BPS, "The State of Indonesia's Employment in August 2019," *Badan Pus. Stat.*, vol. XXiI, 05 N, no. 91, pp. 1–20, 2019.
3. H. Taime and P. N. Djaelani, "The Effect of Economic Growth and Population Growth on the Unemployment Rate in Mimika Regency," *J. Econ. Reg. Sci.*, vol. 1, no. 1, pp. 54–66, 2021.
4. K. Y. Mahendra, M. Susilawati, and N. L. P. Suciptawati, "Modeling the Open Unemployment Rate in Indonesia," *E-Journal Mat.*, vol. 10, no. 1, p. 20, 2021.
5. S. A. Wahyuni, R. Ratnawati, I. Indriyani, and M. Fajri, "Spline Regression Analysis to Modelling The Open Unemployment Rate in Sulawesi," *Nat. Sci. J. Sci. Technol.*, vol. 9, no. 2, pp. 2–7, 2020.
6. L. E. Agustiana, "The Effect of the Covid-19 Outbreak on the Open Unemployment Rate in Affected Sectors in Indonesia," *J. Management Science. Apply.*, vol. 1, no. 6, pp. 546–556, 2020.
7. A. Aljileedi Mustafa Rayhan, R. Rusdarti, and H. Yanto, "Factors Influencing Unemployment Rate: A Comparison Among Five Asean Countries," *J. Econ. Educ.*, vol. 9, no. 1, pp. 37–45, 2020.
8. R. Hidayat, Yuliani, and M. Sam, "Nonparametric regression model with spline truncated approach," *Pros. Semin. Nas.*, vol. 3, pp. 203–210, 2015.
9. S. Siegel, *Nonparametic Statistics for the Social Sciences*. Jakarta: PT. Gramedia, 1992.
10. B. Suharjo, *Applied Regression Analysis with SPSS*First. Yogyakarta: Graha Ilmu, 2013

11. R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. New York: Marcel Dekker. 1999.
12. L. P. S. Pratiwi, "Selection of Optimal Knot Points of Truncated Spline Model in Multivariable Nonparametric Regression with GCV," *J. Mat.*, vol. 10, no. 2, p. 78, 2020.
13. W. Sulaiman, *Regression Analysis Using SPSS, Case Examples & Solutions*First. Yogyakarta: Andi. 2013.
14. M. F. Qudratullah, *Applied Regression Analysis: Theory, Case Studies, and Applications With SPSS*First. Yogyakarta: Andi.
15. Sugiyono, *Qualitative and R&D Quantitative Research Methods*. Bandung: Alfabeta., 2005.
16. R. Watrianthos, *Indonesia's Economic System*. Medan: Yayasan Kita Menulis, 2021.