

Kernel Regression on Reflectance of Lithium Niobate in various Concentrations of Ruthenium Oxide

Anne Mudya Yolanda¹, Muhammad Nur Aidi², Indahwati³, Irzaman⁴
{anne_mudya@apps.ipb.ac.id¹, nuraidi@yahoo.com²,
indahwati.stk@gmail.com}

Student of Department of Statistics, IPB University, Bogor, 16680, Indonesia¹
Lecturer of Department of Statistics, IPB University, Bogor, 16680, Indonesia^{2,3}
Department of Physics, IPB University, Bogor, 16680, Indonesia⁴

Abstract. Research on LiNbO_3 doped with RuO_2 has been widely developed. This study aimed to use kernel regression to measure the influence of wavelength to the percentage of reflectance of LiNbO_3 doped with RuO_2 and compare estimated results among all concentrations. A hundred candidates of bandwidth were tried to find the optimum bandwidth. For the percentage of reflectance with wavelength around 450.2 to 900.9, the results show that the optimum bandwidth is 11.56. The kernel regression performs smoothing and estimating at each data point based on the optimum bandwidth used so that the fitted data is closely following the observed data. The kernel regression model produced adjusted R-square as 0.9629, 0.9590, 0.9871, and 0.9840, respectively for LiNbO_3 doped with concentration 0, 2, 4, and 6%. For the same wavelength, the percentage of reflectance of a material made of LiNbO_3 doped with various concentration is higher than material made of LiNbO_3 only.

Keywords: kernel, LiNbO_3 , reflectance, RuO_2

1 Introduction

Lithium Niobate is synthesis material, compound of niobium, lithium, and oxygen. Lithium Niobate as one of the ferroelectric materials has many benefits. It has been thought to be one amongst promising platforms for the multi-purpose optical elements and photonic circuits [1] and the most important materials of electronic engineering [2]. That is why research on lithium niobate has been widely developed, including when it doped with Ruthenium Oxide. It is synthesis material that has an awfully smart candidate for its good electrochemical science capacitance performance and high specific capacitance [3]. The doping process is a form of treatment to optimize the benefits of lithium Niobat and for reducing cost.

In its application, the lithium performance is affected by wavelength, considering that each tool can only work effectively at certain wavelengths. The relationship of wavelength with Lithium Niobate could be seen from its optical properties, such as reflectance. Reflectance was taken from the results of the research using a UV-VIS spectrophotometer at Lithium Niobate for a range of wavelengths. Therefore, regression modeling is appropriate, but based on data patterns and assumptions, data from the research can be studied by kernel regression. Using bandwidth as the parameter of smoothing in kernel regression, the estimated reflectance would be approached specifically based on the wavelength of Lithium Niobate.

Kernel regression estimates each observed data based on bandwidth. Thus, the result of kernel regression is better than other analysis.

The previous research was carried out using the Univariate and Multivariate ARIMA to estimate absorption pattern in Strontium Titanate dop variation. The result showed that ARIMA model is more appropriate to approach the absorption of strontium titanate with variation of dop [4]. Another researcher studied the characteristics modeling of optical properties on barium titanate using ARIMA and VARIMA, concluded that the best model for the absorption variable is VARIMA (1,1,0), reflectance variables is VARIMA (1,1,0), and the transmittance variable is VARIMA (1,1,0) [5]. Arima Analysis on Lithium Niobate revealed that FTIR and XRD value on the heated LiNbO3 thin films could be modelled well by ARIMA [6].

Based on these descriptions, that studies only learn the optical properties itself, without adding the influence of wavelength. Therefore this regression model is needed. Research with a regression model on Barium Titanate has been done by forming a regression model according to the data pattern, linear or quadratic [7]. In this study, Lithium Niobate is analyzed by a non-parametric approach. Doping process is carried out at concentrations of 2, 4 and 6%. The selection of the four concentrations refers to the previous research, which states that there is an effect that results from the process of Lithium Niobate and Ruthenium Oxide doping [8].

This study aimed to use kernel regression to measure the influence of wavelength to the percentage of reflectance of Lithium Niobate doped with ruthenium oxide and compare estimated results among all concentrations.

2 Materials

The data in this study are primary data from research conducted by a team from Department of Physics and Department of Statistics, IPB University. The data in this study are part of the Competency Grant project 2018 entitled "High Accuracy ARIMA Model and Rapid Prediction in Analyzing XRD Spectral Data" funded by the Ministry of Research, Technology, and Higher Education. The research was carried out by four variations of doping, i.e. 0%, 2%, 4%, and 6% with reflectance (in percentage) as the response variable and wavelength as the explanatory variable (450.3 to 900.9 nm).

3 Methods

Kernel models is flexible, simple mathematical form, and relatively fast convergent [9]. Based on the non-parametric regression model, the estimator of kernel regression or often called the Nadaraya-Watson estimator [10] can be defined as

$$\hat{f}(x_i) = \frac{\frac{1}{n} \sum_{i=1}^n K_h(x - x_i) Y_i}{\frac{1}{n} \sum_{i=1}^n K_h(x - x_i)} \quad (1)$$

where $\sum_{i=1}^n K_h(x - x_i) = \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)$ so

$$\hat{f}(x_i) = \frac{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)} \quad (2)$$

for $i = 1, 2, 3, \dots, n$ with
 n = number of observation,
 $K_h(\cdot)$ = kernel K for bandwidth “ h ”,
 x_i = i -th explanatory variable,
 Y_i = i -th respons variable.
or could be written as follow

$$\widehat{f}(x_i) = \sum_{i=1}^n w_h(x) Y_i \quad (3)$$

with $w_h(x)$ is a function given in equation (4),

$$w_h(x) = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)} \quad (4)$$

The often used kernel function is the Gaussian

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(-u^2)\right) \quad (5)$$

for $-\infty < u < \infty$

Then, the Nadaraya-Watson estimation obtained by the Gaussian Kernel function as the estimator of regression is as follows

$$w_h(x) = \frac{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-x_i}{h}\right)^2\right)}{\sum_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-x_i}{h}\right)^2\right)} \quad (6)$$

The difference between the kernel estimator and other estimators is bandwidth. Bandwidth in the kernel functions is similar to the width of the interval on a histogram [10]. In kernel analysis, bandwidth is used as the smoothing parameter. In kernel regression, choosing an optimum bandwidth is more important than kernel function. The optimum window has the smallest *Generalized Cross Validation (GCV)*.

$$GCV = \frac{MSE}{(n^{-1} \text{tr}[\mathbf{I} - \mathbf{H}])^2} \quad (7)$$

where

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \widehat{Y}_i)^2 \quad (8)$$

and

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T + n\mathbf{h}\mathbf{I})^{-1} \mathbf{X}^T \quad (9)$$

with

\mathbf{X} = explanatory variable matrix in kernel regression,
 n = number of observation,
 h = bandwidth,
 \mathbf{I} = matrix I.

In this study, bandwidth and kernel regression were analyzed by PLRModels [11] in the R program.

4 Results and Discussion

The optical properties of reflectance are influenced by the wavelength. Figure 1 shows that the fluctuation of reflectance is not too large. Each concentration has a different pattern. This shows that the doping process has an effect on the reflectance produced.

The optical properties of reflectance are influenced by the wavelength. Figure 1 shows that the fluctuation of reflectance is not too large. Each concentration has a different pattern. This shows that the doping process has an effect on the reflectance produced.

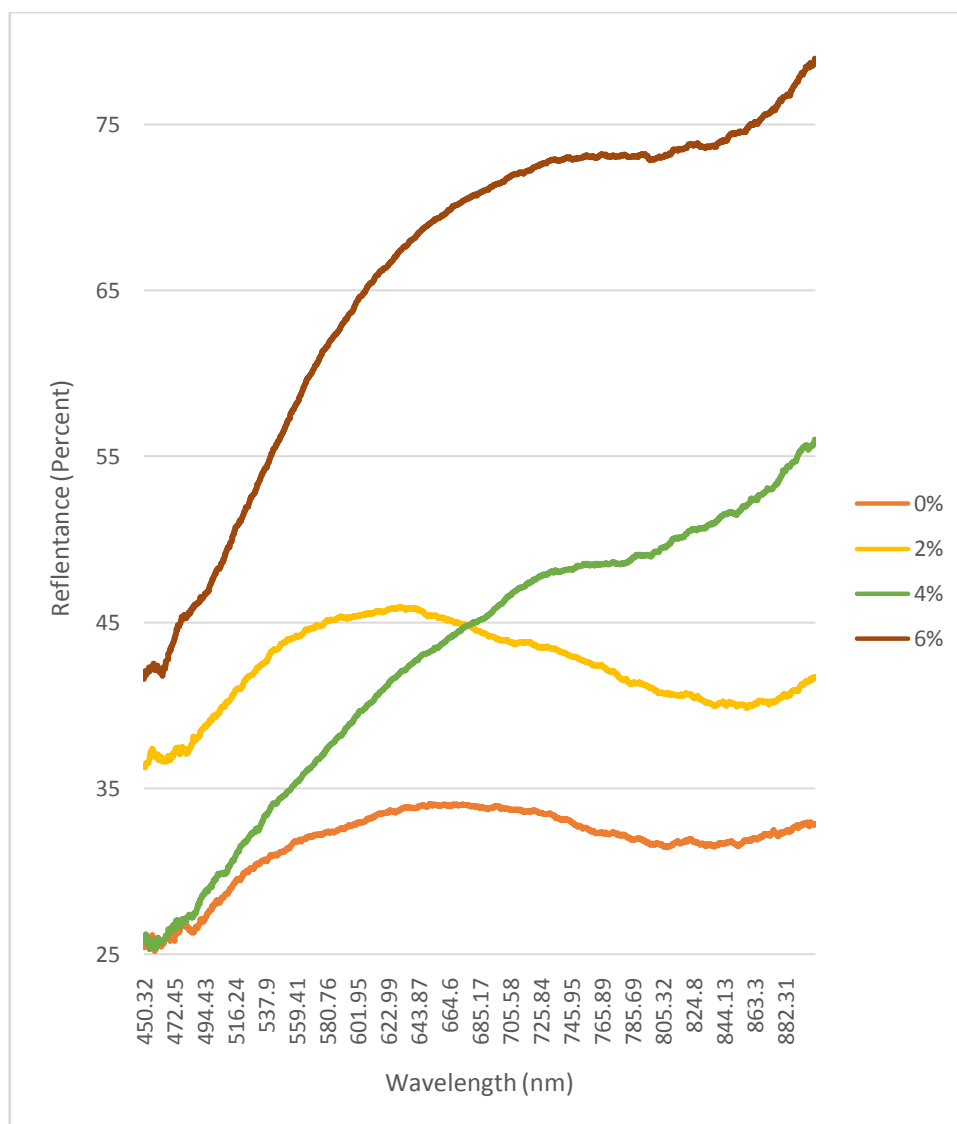


Fig.1 Plot of Lithium Niobate doped with Ruthenium Oxide

Figure 1 shows the plot of each concentration which is very difficult to estimate if a parametric regression is used because the curves do not form specific patterns. That is why the regression curve will be estimated using a nonparametric regression, namely kernel regression.

Determining bandwidth is very important in kernel regression. If the bandwidth used is very small, then the resulting regression curve is very rough. It is because kernel regression performs smoothing one by one according to the bandwidth.

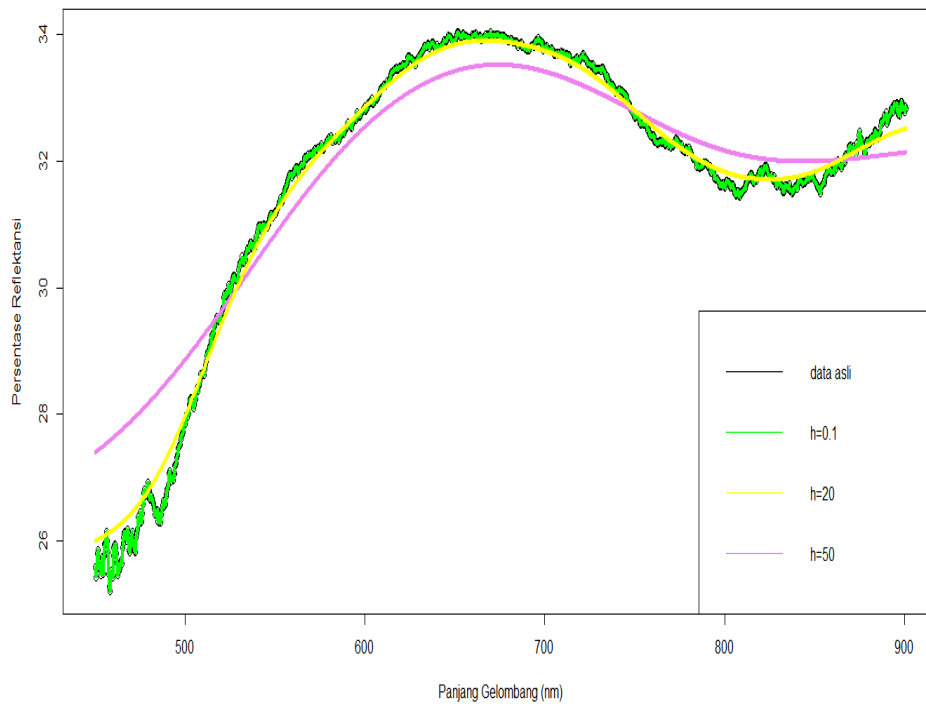


Fig.2 The kernel regression curve for several candidate of bandwidths

Figure 2 shows the use of several candidates of bandwidths in the observed data for Lithium Niobate doped with Ruthenium Oxide concentration 0%. The smaller the bandwidth used, the smaller the bias obtained, but the greater the variance produced. Conversely, if the selected bandwidth is too large, the regression curve formed looks very smooth but does not match the data pattern. The greater the bandwidth chosen, the greater the bias obtained, even though the variance gets smaller. This condition provides evidence that the choice of bandwidth greatly determines the goodness of the kernel regression curve.

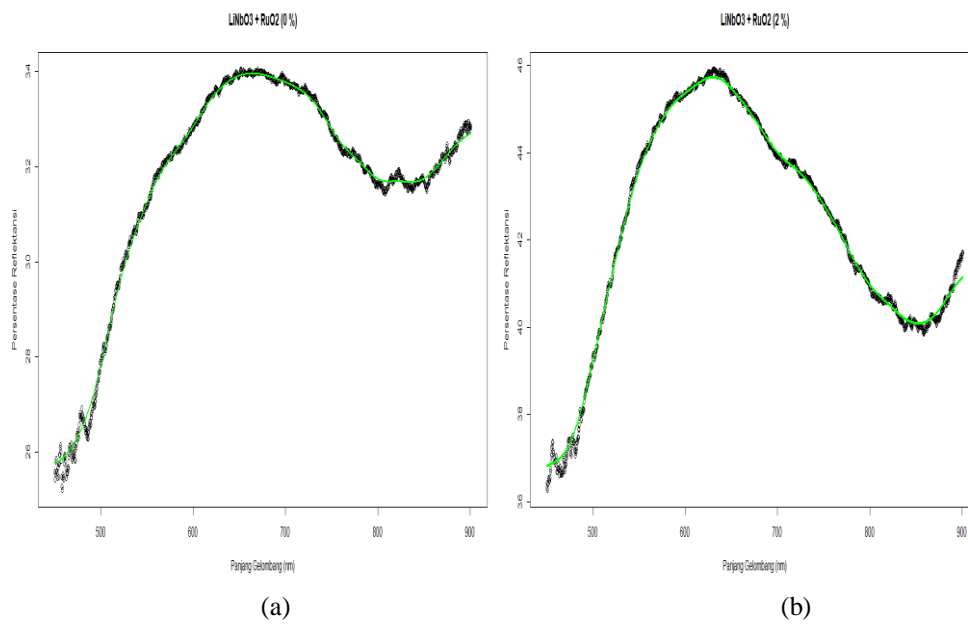
Table 1 presents the GCV of optimum bandwidth at each concentration. In this study, as many as 100 bandwidth candidates were tried using the Gaussian Kernel Function with the Nadaraya-Watson estimator.

Table 1. GCV of Optimum Bandwidth

Concentratio n	Bandwid th	GCV
-------------------	---------------	-----

0%	11.56	0.0143
2%	11.56	0.0185
4%	11.56	0.0304
6%	11.56	0.0511

The analysis finds that the optimum bandwidth for all concentrations is 11.56. The kernel regression performs smoothing and estimating at each data point based on the optimum bandwidth used, hence at optimum bandwidth, that's why the fitted data is closely following the observed data. Optimum bandwidth describes the limitation of the wavelength range which has a considerable influence on other wavelengths around it. The regression curve of various concentrations of Lithium Niobate and Ruthenium Oxide are given in Figure 3.



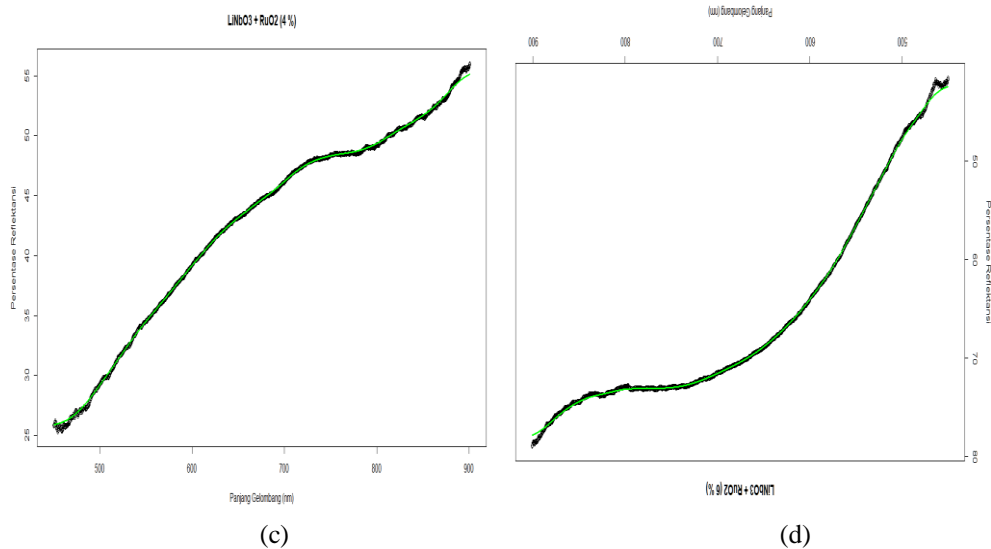


Fig.3 Kernel Regression Curve of Lithium Niobate doped with Ruthenium Oxide
 (a) 0% (b) 2% (c) 4% (d) 6%

Figure 3 (a), (b), (c), (d) show the observed and fitted data for concentrations 0, 2, 4, and 6%, respectively Lithium Niobate doped with Ruthenium Oxide. Figure 3 shows the contribution of kernel regression term of wavelength to the percentage of reflectance. The variable in non-parametric model usually displayed graphically, because it can't be expressed as parametric. The non-parametric model only have function. As shown in Figure 3, the fitted is fluctuate like the observed data, means that fitted data can predict the observed data very accurate. Summary of fitted and observed data is shown in Table 2.

Table 2. Summary of Fitted and Observed Data

Statistic	Concentration			
	0%		2%	
	Observed	Fitted	Observed	Fitted
Min	25.20	25.76	36.27	36.82
Q1	31.58	31.68	40.42	40.43
Median	32.24	32.22	42.40	42.37
Mean	31.76	31.76	42.27	42.27
Q3	33.45	33.36	44.54	44.51
Max	34.05	33.96	45.91	45.74
Statistic	Concentration			
	4%		6%	
	Observed	Fitted	Observed	Fitted
Min	25.31	25.88	25.31	25.88

Q1	36.29	36.35	36.29	36.35
Median	45.13	45.20	45.13	45.20
Mean	42.85	42.85	42.85	42.85
Q 3	49.05	49.15	49.05	49.15
Max	56.00	55.11	56.00	55.11

Based on Table 2, for the same wavelength, reflectance after doping process is greater than the reflectance of Lithium Niobate without doping. The more Ruthenium Oxide concentration is added, the greater the reflectance produced. Therefore, it is known that, for the same wavelength, the ability of a material to reflect is higher if it is made of Lithium Niobate doped with various concentration of Ruthenium Oxide. The adjusted R-Square for each concentration are given in Table 3.

Table 3. Adjusted R-Square for each Concentration

Concentration	Adjusted R-square
0%	0.9629
2%	0.9590
4%	0.9871
6%	0.9840

The kernel regression model produces a high number of adjusted R-square. It can reveal that the model is adequate for estimating the percentage of reflectance for various concentration of Lithium Niobate doped with Ruthenium Oxide. Kernel regression approaches segmentation with the bandwidth as a parameter, but it does not immediately appear in the function generated for estimation. If desired gain a model that can estimate for a specific range of wavelength, then in the next study it can apply spline regression or non-parametric regression by smoothing the kernel and spline simultaneously.

5 Conclusion

Based on the analysis, it can be concluded that the optimum bandwidth for the percentage of reflectance with a wavelength around 450.2 to 900.9 nm is equal to 11.56 with adjusted R-square is 0.9629, 0.9590, 0.9871, and 0.9840, respectively for 0, 2, 4, and 6%. The kernel regression performs smoothing and estimating at each data point based on the bandwidth used so that the fitted data is closely following the observed data. It is also known that for the same wavelength, the percentage of reflectance of a material made of Lithium Niobate doped with various concentration is higher than material made of Lithium Niobate only.

Acknowledgments. We thank our lecturers and colleagues from the Department of Statistics, IPB University for sharing insight and comments that greatly improved the manuscript, although any errors are our own and should not tarnish the reputations of these esteemed persons. We are also immensely grateful to the Ministry of Research and Technology of Higher Education for funding this research as Competency-based Research Grants number 1611/IT3.11/PN/2018.

References

- [1] B. Gao, M. Ren, W. Wu, H. Hu, W. Cai, and J. Xu: Lithium Niobate Metasurfaces. *Laser Photonics Rev.*, vol. 13, no. 5 (2019)
- [2] N. V. Sidorov, M. N. Palatnikov, N. A. Teplyakova, A. V. Syuy, and D. S. Shtarev: Structural features and optical properties of lithium niobate crystals in *Advanced Materials* (2018)
- [3] P. S. Joshi and S. D.S.: Ruthenium Oxide : Thin Film and Electrochemical Properties. *Int. J. Res. Appl. Sci. Eng. Technol.*, Vol. 4, no. VIII, pp. 765–772 (2016)
- [4] D. Oktasari, M. N. Aidi, and I. M. Sumertajaya: A Comparison of Univariate ARIMA and Multivariate to Estimate Absorption Pattern in Stronsium Tittanate Dop Variation. *IJSRSET*, Vol. 4, no. 9, pp. 356–360 (2018)
- [5] L. Hariningrum, M. N. Aidi, I. M. Sumertajaya, and Irzaman: Characteristics Modeling of Optical Properties (Absorption, Reflectance, Transmittance) on Barium Titanate Using ARIMA and VARIMA. *Int. J. Sci. Eng. Res.*, Vol. 9, no. 12, pp. 29–32 (2018)
- [6] M. N. Aidi, A. A. Setiawan, M. Zuhri, H. Alatas, and Irzaman: ARIMA analysis of ferroelectric lithium niobate (LiNbO₃) thin films. *Int. J. Sci. Res. Sci. Eng. Technol.*, Vol. 4, no. 4, pp. 823–848 (2018).
- [7] M. N. Aidi, B. Susetyo, H. Hardhienata, and Irzaman: Regression and arima residual have high accuracy to estimate percent adsorption spectral based on wavelength at barium titanate. *Int. J. Sci. Res. Sci. Eng. Technol.*, Vol. 4, no. 9, pp. 585–602 (2018)
- [8] M. Zuhri, Novitri, A. A. Setiawan, H. Alatas, and Irzaman: Optical properties of lithium niobate (LiNbO₃) thin film doped with ruthenium oxide. *IOP Conf. Ser. Earth Environ. Sci.*, Vol. 187, no. 012020 (2018)
- [9] I. N. Budiantara and Mulianah: Pemilihan bandwidth dalam regresi semiparametrik kernel dan aplikasinya. *SIGMA ITS Surabaya*, Vol. 10, no. 2, pp. 159–166 (2007)
- [10] I. G. D. E. Sukarsa, I. G. A. M. Srinadi, and N. L. A. P. Lestari: Perbandingan estimator kernel dan estimator spline dalam model regresi nonparametrik. *KNM XVI 3-6 Juli 2012 UNPAD*, no. July 2012, pp. 581–590 (2015)
- [11] G. A. Perez and A. L. Cheda: Package ‘PLRModels’: Statistical inference in partial linear regression models. R package version 1.1. [diunduh 2019 Maret 20]. (2015)