International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V68I8P502
Volume 68 Issue 8, 12-20, August 2022

Original Article

The Nonparametric Kernel Method using Nadaraya-Watson, Priestley-Chao and Gasser-Muller Estimators for the Estimation of the Rainfall Data in Lampung

Herawati. N[1*], Sayuti. S.F[1], Nisa. K[1], Setiawan. E[1]

[1]*Department of Mathematics, University of Lampung, Lampung, Indonesia.*

Received: 21 June 2022 Revised: 22 July 2022 Accepted: 02 August 2022 Published: 15 August 2022

Abstract *- There are two methods used in regression analysis, namely parametric and nonparametric. Parametric regression, assumes many assumptions, such as the normality of data, the distribution of data that forms a certain pattern, and others. In nonparametric regression, these assumptions do not have to be fulfilled. To estimate the nonparametric regression function, smoothing techniques are needed, one of which is the kernel method. In this study, the nonparametric regression estimators of Nadaraya-Watson, Priestley & Chao and Gasser & Müller were compared based on the smallest GCV value and the best kernel estimator was determined by the smallest error rate in the rainfall data in the province of Lampung. The results showed that the Nadaraya-Watson estimator was determined as the best estimator with a bandwidth (h) of 5.564 and had the smallest error value compared to the other two methods, namely MSE=5.9773 and MAPE=0.1935.*

Keywords – *Kernel estimator, Bandwidth, Nadaraya-Watson, Priestley-Chao, Gasser-Muller, MSE, MAPE.*

1. Introduction

Regression analysis is statistical technique for analyzing the correlation between predictor variables (X) and response variables (Y) which is explained through a curve called the regression curve. There are two methods used in regression analysis, namely parametric and nonparametric. The parametric models assume many assumptions, such as data normality, data patterns that form certain patterns, the variance was constant, and others. In the parametric regression model, the assumption is often not met, so the researcher looks for another regression model that is more flexible with the assumptions. One method that can be used is the nonparametric regression model [1]. This estimator is a development of the histogram estimator. The difference between the kernel estimator and other nonparametric regression estimators is the kernel has more specialized in bandwidth usage. The kernel method requires optimal kernel function and bandwidth in its operation. One of the methods to find the optimal bandwidth is the Generalized Cross Validation (CGV) method [2].

The use of kernel estimators in many studies has been carried out. one of them is a study conducted by [3] which applied the kernel method to cancer data. In addition, a study by [4] modeled and predicted the number of COVID-19 infections in Iraq using three nonparametric kernel estimators with a Gaussian weighting function. The results showed that the Priestley-Chao model was a suitable model for all sample sizes and other conditions used in the study. The study of [5-10] found that the kernel method has better results than using the classic time series method. In this study, we will compare the nonparametric regression estimators of Nadaraya-Watson, Priestley & Chao and Gasser & Müller based on the smallest GCV value. Determination of the best kernel estimator is done by looking at the smallest error rate in the rainfall data in the province of Lampung.

2. Kernel Estimator

According to [11], nonparametric regression is used when regression curve is not known. Nonparametric regression does not assume the shape of regression curves, regression curve is assumed to be contained within a particular function space. Mathematically, the nonparametric regression model can be written:

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

Where y_i = response variable from the i-th data $(i = 1,2,3, \dots, n)$, $m(x_i)$ = unknown i-th smooth function $(i = 1,2,3, \dots, n)$, ε_i = independent i-th error assumed to spread $N \sim (0, \sigma^2)$ $(i = 1,2,3, \dots, n)$

On this model, the function of unknown regression is $m(x_i)$. In the nonparametric regression model, it depends on the weighted average of the dependent variable with the weight of the observation distance on the independent variable which is measured based on the value of the smoothing parameter [12].

The estimation of nonparametric regression function is based on observational data using smoothing technique. One of the smoothing techniques that is often used is the kernel estimator [13]. Since [14, 15] introduced this estimator, it is known as the Rosenblatt-Parzen kernel density. In the nonparametric approach model, the kernel estimator is very commonly used because this estimator has several advantages:

- It has a flexible form and it is mathematically easy to work with.
- It has a relatively fast average convergence.

In general, the kernel density function of $\hat{f}_h(x)$ is:

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

Then the K kernel with bandwidth (h) is defined as

$$K_h(x - X_i) = \frac{1}{h} K\left(\frac{x - X_i}{h}\right) \qquad (2)$$

for $-\infty < x < \infty$ and $h > 0$. So the kernel density function of $\hat{f}_h(x)$ can be written as:

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \qquad (3)$$

with K is a continuous kernel function. K is usually assumed to satisfy regularity conditions such as constraints [16]. Graphically, a function is said to be continuous at $x\epsilon[a, b]$ if a graph of the function f is not truncated at the point (a, b). While the kernel refiner uses bandwidth (h) which plays a role in defining and determining variations and biases.

According to [17], $x = (x_1, x_2, \dots, x_n)$is a sample which taken randomly from a density distribution f and K a finite and positive function then there are several conditions that must be met, they are:

(i) $K(x) \geq 0$, for all real x
(ii) $\int_{-\infty}^{\infty} K(x)dx = 1$
(iii) $\int_{-\infty}^{\infty} x^2 K(x)dx = \sigma^2 > 0$
(iv) $\int_{-\infty}^{\infty} x_i K(x)dx = \begin{cases} 1, & i = 0 \\ 0, & 1 \leq i < r, \\ \neq 0, & i = r \end{cases}$ for a number r
(v) $K(x)$ is symmetrical around zero.

The shape of the graph is felt around a vertical Y line that forms the bell of some random variable value with a certain distance on one side that is equal to the other side of the value.

2.1. Nadaraya-Watson Estimator

The study of [18] proposed an estimator as a weighted average locale with kernel as a weighting function. According to [19], there are n observational data $\{(X_i, Y_i)\}_{i=1}^{n}$ that satisfies the equation where $X_i \in R$ and $Y_i \in R$, then the estimator $\widehat{m}(x_i)$ is:

$$\widehat{m}(x_i) = E(Y|X = x) = \int \frac{y(X,Y)}{f(X=x)} dy$$

The denominator is estimated by using the kernel density estimator as follows:

$$f_h(x) = \frac{1}{n}\sum_{i=1}^{n}(x - X_i)$$

The joint probability density function is estimated by the kernel product:

$$\acute{f}_{h_1,h_2}(x, y) = \frac{1}{n}\sum_{i=1}^{n} K_{h_1}(x - X_i)\, K_{h_2}(y - Y_i)$$

So, the numerator of the Nadaraya-Watson estimator becomes:

$$\acute{f}_{h_1,h_2}(x,y) = \frac{1}{n}\sum_{i=1}^{n} K_{h_1}(x - X_i) \int y\, K_{h_2}(y - Y_i)dy$$

$$= \frac{1}{n}\sum_{i=1}^{n} K_{h_1}(x - X_i) \int \frac{y}{h_2} K\left(\frac{y-Y_i}{h_2}\right) dy$$

$$= \frac{1}{n}\sum_{i=1}^{n} K_{h_1}(x - X_i) \int (sh_2 + Y_1)K(s)\, ds$$

$$= \frac{1}{n}\sum_{i=1}^{n} K_{h_1}(x - X_i)\, Y_i$$

So that the Nadaraya-Watson estimator can be written:

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

2.2. Priestley-Chao Estimator

According to [20], the Priestley-Chao estimator is an estimator for an unknown regression function. Recall equation (1), with $m(x_i)$ as an unknown function of curve shape and error (ε_i) with a mean of zero and σ^2 which must be constant. It can also be assumed that $(x_1, x_2, \dots, x_n)$ are in the same interval [a, b], so that

$$x_i = i\ x\ \delta \qquad i = 1, \dots, n$$

where $\delta = (b - a)/n$. It is necessary to estimate m nonparametrically using the available data. The estimator is defined as:

$$\widehat{m}(x_i) = \frac{\delta}{h}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) Y_i$$

For $x\epsilon(a, b)$, $K(.)$ is called a symmetric function with zero, such that $\int K(u)^2 du < \infty$ and has a finite second moment that is $\int u^2 K(u)du = \sigma_K^2 < \infty$. In general, we will choose a symmetric probability density function such as standard normal or symmetrical Beta on a finite interval [-1, 1]. The constant h is called the smoothing parameter or bandwidth and controls the kernel function on each x_i. The Priestley-Chao estimator is the weighted average of the response variables $(Y_1, Y_2, \dots, Y_n)$ and the weight is related to Y_i, the weight is $K\left(\frac{x-x_i}{h}\right)$ with the actual result being determined by the proximity of x to x_i that relative to the value of h [21]. It is a linear function of Y_i and therefore called a linear smoothing. If the data does not have the same space or equally-spaced, then the estimator is

$$\widehat{m}^{PC}(x_i) = \frac{1}{h}\sum_{i=1}^{n}(x_i - x_{i-1})K\left(\frac{x-x_i}{h}\right) Y_i \quad (5)$$

2.3. Gasser-Muller Estimator

The unknown regression function $\widehat{m}(x_i)$ can also be estimated by [22] estimator. The kernel functions K as a probability density function and is a bandwidth, so the Gasser and Muller estimators are defined by:

$$\widehat{m}(x_i) = \sum_{i=1}^{n}\left[\int_{s_{i-1}}^{s_i} K_h(x - x_i)dx\right] Y_i$$

For $x \in [0,1]$. Where $s_0 = 0$, $s_n = 1$, $x_i \le s_i \le x_{i+1}$, and $s_i = \frac{x_i + x_{i+1}}{2}$ for $i = 1, 2, \dots, n-1$. Kernel functions are generally expressed as $K_h(x - x_i) = \frac{1}{h} K\left(\frac{x-x_i}{h}\right)$ (Moon, 2011).

According to [23], the Gasser-Muller estimator can be expressed in the form:

$$\widehat{m}^{GM}(x) = \frac{1}{h}\sum_{i=1}^{n}\left[\int_{s_{i-1}}^{s_i} K_h\left(\frac{x-x_i}{h}\right) dx\right] Y_i \quad (6)$$

2.4. Kernel Functions

In kernel regression, there are functions that must be continuous, symetrical, finite, and have real values. The general function of the K kernel with the bandwidth smoothing parameter (h) is defined as [24]:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right), \quad \text{for } -\infty < \mathrm{x} < \infty \text{ and } h > 0$$

with
K = kernel function
h = bandwidth

According to [25], several types of kernel functions are:

a. Uniform: $K(x) = \frac{1}{2} I$, for $|x| \leq 1$

b. Epanechnikov: $K(x) = \frac{3}{4}(1 - x^2) I$, for $|x| \leq 1$

c. Kuartik: $K(x) = \frac{15}{16}(1 - x^2)^2$, for $|x| \leq 1$

d. Triweight: $K(x) = \frac{35}{32}(1 - x^2)^3 I$, for $|x| \leq 1$

e. Cosinus: $K(x) = \frac{\pi}{4} cos\left(\frac{\pi}{2} x\right)$, for $|x| \leq 1$

f. Gaussian: $K(x) = \frac{1}{\sqrt{2\pi}} exp^{\left(-\frac{1}{2}(x^2)\right)}$, for $-\infty < x < \infty$

Where I is an indicator function

$I(u) = \begin{cases} 1, & if\ |x| \leq 1 \\ 0, & if\ |ux| > 1 \end{cases}$

In this case, the Epanechnikov kernel function is used. The general form of a Epanechnikov kernel function is:

$$K(u) = \frac{3}{4}(1 - x^2) I \tag{7}$$

2.5. Optimal Bandwidth Selection

In addition, in kernel estimation, the selection of smoothing parameters (bandwidth) is one of the most crucial things. Bandwidth (h) is a smoothing parameter that functions to control the smoothness of the estimation curve [26]. If the selected bandwidth is too small, it will produce a curve that is less smooth and does not reflect the actual data because it is very volatile. Meanwhile, if the bandwidth used is too wide, it will produce a curve that is not in accordance with the data distribution pattern.Therefore, it is necessary to choose a bandwidth that can predict the data correctly or the difference between the estimated data and the actual data is very small. [27]. According to [28], one method to get the optimal bandwidth value is to use the Generalized Cross Validation (GCV) criteria. To determine the GCV value, the following equation can be used:

$$GCV(h) = \frac{MSE}{\left(1 - \frac{trace\ (W)}{n}\right)^2} \tag{8}$$

where $MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}$ and W is Hat Matrix $(n\ x\ n)$ that fulfills $[m_h(x_1), m_h(x_2), \dots, m_h(x_n)]^t = WY$. The optimal bandwidth value will be obtained if it produces a minimum Generalized Cross Validation.

3. Methods

In this study, researchers used Rainfall data in Lampung from January 2018 to December 2020 obtained from bps.go.id with the predictor variable (X) is the average air humidity and the response variable (Y) is the rainfall amounts. We estimated the data model using kernel estimator. The optimal bandwidth of the three types of estimators (Nadaraya-Watson, Priestley-Chao, and Gasser-Muller) is determined based on the smallest GCV as well as in determining the best estimator for MSE and MAPE values.

4. Results and Discussion

Figure 1 shows the distribution of Rainfall data in Lampung from January 2018 to December 2020.

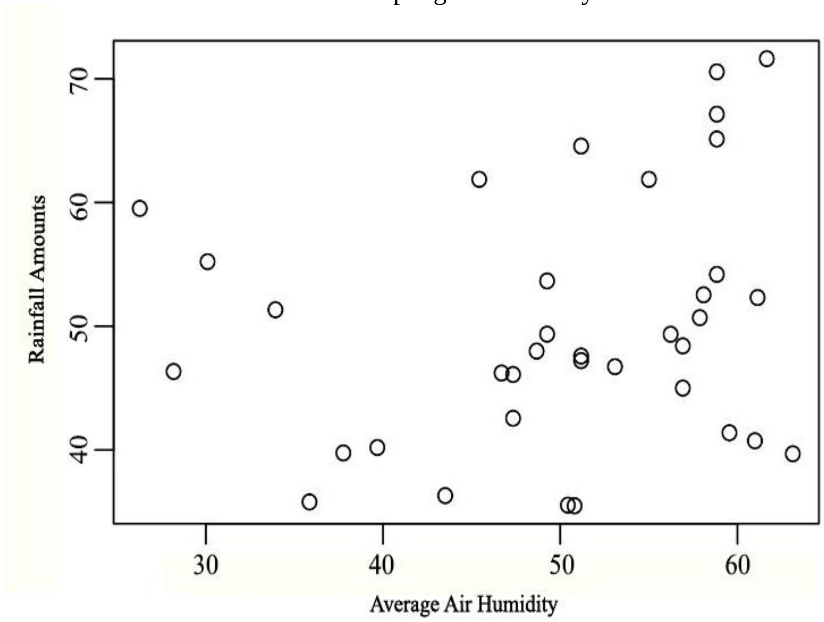

Fig 1. Data Distribution of Rainfall Data in Lampung

Based on Figure 1, it can be seen that the distribution of the data cannot be determined clearly or in other words the data does not form a certain pattern. It is stated that the Y variable has a non-linear relationship (negative correlation) with the X variable. For this reason, further analysis of the distribution of the data is required. whether the data is normally distributed or not so that nonparametric regression can be applied. In this case, the Kolmogorov-Smirnov test was carried out to test the normality of the data. The results of the Kolmogorov-Smirnov test can be seen in Table 1.

Table 1. Kolmogorov-Smirnov test on rainfall data in Lampung Province

Data	D	$p-value$
Rainfall Data	1	2,2e-16

The hypotheses used are H_0: Error normally distributed $(0, \sigma^2)$ and H_1: Error is not normally distributed $(0, \sigma^2)$. Where $H_0$is accepted if $D < D_{table}$ or $p-value > \alpha$ significance value (α) of 0.05 and the test statistic is $D = maks|F_0(x) - S_N(x)|$. The calculation result stated that the value of *D*=1 with *p-value* = 2.2 x 10^{16}. Because the value of $D > D_{table}$ and $p-value > \alpha$ then H_0 is rejected, it means that the data was not normally distributed. Because the assumption of normality is not met, regression curve estimation for data on the amount of rainfall in Lampung Province was carried out using nonparametric regression with a kernel estimator.

Next is the calculation of the optimal bandwidth value (h) using the Generalized Cross Validation (GCV) method with the help of the RStudio application program. The results can be seen in Table 2.

Table 2. Optimal bandwidth selection

no	***h***	**GCV**	**No.**	***h***	**GCV**
1	0,940	166,422	26.	5,069	103,483
2	1,105	161,087	27.	5,234	102,564
3.	1,270	150,460	28.	5,399	101,680
4.	1,435	147,611	29.	**5,564**	**101,054**
5.	1,601	147,177	30.	5,729	101,104
6.	1,766	143,801	31.	5,895	101,246
7.	1,931	138,460	32.	6,060	101,409
8.	2,096	133,193	33.	6,225	101,486
9.	2,261	129,671	34.	6,390	101,585

10.	2,426	125,203	35.	6,555	101,661
11.	2,592	122,198	36.	6,720	101,638
12.	2,757	120,116	37.	6,886	101,759
13.	2,922	118,940	38.	7,051	101,985
14.	3,087	119,779	39.	7,216	102,230
15.	3,252	120,483	40.	7,381	102,634
16.	3,417	120,361	41.	7,546	102,952
17.	3,582	119,435	42.	7,711	103,258
18.	3,748	118,341	43.	7,876	103,541
19.	3,913	117,163	44.	8,042	103,850
20.	4,078	115,196	45.	8,207	104,113
21.	4,243	113,288	46.	8,372	104,186
22.	4,408	110,661	47.	8,537	104,259
23.	4,573	108,040	48.	8,702	104,332
24.	4,739	106,103	49.	8,867	104,342
25.	4,904	104,567	50.	9,033	104,274

As we can see from Table 2 that the smallest GCV value is 101,054 with h=5,564. This bandwidth was used for rainfall data modeling to generate a regression model kernel. By substituting h=5,564 to Epanechnikov kernel function to the Nadaraya-Watson, Priestley-Chao, Gasser-Muller estimator, the resulting kernel regression models are as follows:

1. Nadaraya-Watson Estimator

$$\hat{y}_i = \hat{m}^{NW}(x_i)$$

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

$$= \frac{\sum_{i=1}^{n} \frac{3}{4}\left(1-\left(\frac{x-x_i}{5,564}\right)^2\right) I\, Y_i}{\sum_{i=1}^{n} \frac{3}{4}\left(1-\left(\frac{x-x_i}{5,564}\right)^2\right) I\, Y_i}$$

2. Priestley-Chao Estimator

$$\hat{y}_i = \hat{m}^{PC}(x_i)$$

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

$$= \frac{1}{h}\sum_{i=1}^{n} (x_i - x_{i-1}) \frac{3}{4}\left(1-\left(\frac{x-x_i}{5,564}\right)^2\right) I\, Y_i$$

3. Gasser-Muller Estimator

$$\hat{y}_i = \hat{m}^{GM}(x_i)$$

$$= \frac{1}{h}\sum_{i=1}^{n} \left[\int_{s_{i-1}}^{s_i} K_h\left(\frac{x-x_i}{h}\right) dx\right] Y_i$$

$$= \frac{1}{h}\sum_{i=1}^{n} \left[\int_{s_{i-1}}^{s_i} K_h \frac{3}{4}\left(1-\left(\frac{x-x_i}{5,564}\right)^2\right) I dx\right] Y_i$$

Then we calculated the predicted value using the prediction model that has been obtained. Based on the cal-culations that have been carried out using RStudio, the prediction results of the regression function are obtained.

Table 3. The prediction results

No.	Nadaraya-Watson	Priestley-Chao	Gasser-Muller
1	51,593	0	5,919
2	50,418	29,324	54,3
3	49,001	32,654	52,659
4	45,98	72,398	47
5	44,892	38,303	43,368
6	44,456	41,656	40,486
7	44,716	48,132	39,043
8	46,329	140,39	43,918
9	47,181	85,141	46,137
10	47,706	62,42	47,351
11	47,973	33,906	48,092
12	47,973	0	48,092
13	48,525	76,138	48,923
14	48,778	35,008	49,122
15	48,778	0	49,122
16	49,299	73,832	49,445
No.	Nadaraya-Watson	Priestley-Chao	Gasser-Muller
17	49,474	24,399	49,614
18	49,648	24,133	49,842
19	49,648	0	49,842
20	49,648	0	49,842
21	50,57	132,57	51,246
22	51,48	137,19	52,415
23	52,003	86,988	52,274
24	52,267	48,732	52,172
25	52,267	0	52,172
26	52,593	66,153	51,169
27	52,657	14,366	50,7
28	52,863	49,687	48,23
29	52,863	0	48,23
30	52,863	0	48,23
31	52,863	0	48,23
32	53,027	45,212	44,862
33	53,267	82,121	36,315
34	53,285	8,4358	35,374
35	53,338	27,758	32,035
36	53,413	65,518	22,023

Next, the curves of the estimators were compared using three estimators (Nadaraya-Watson, Priestley-Chao, Gasser-Muller). The results are presented in Figure 2.

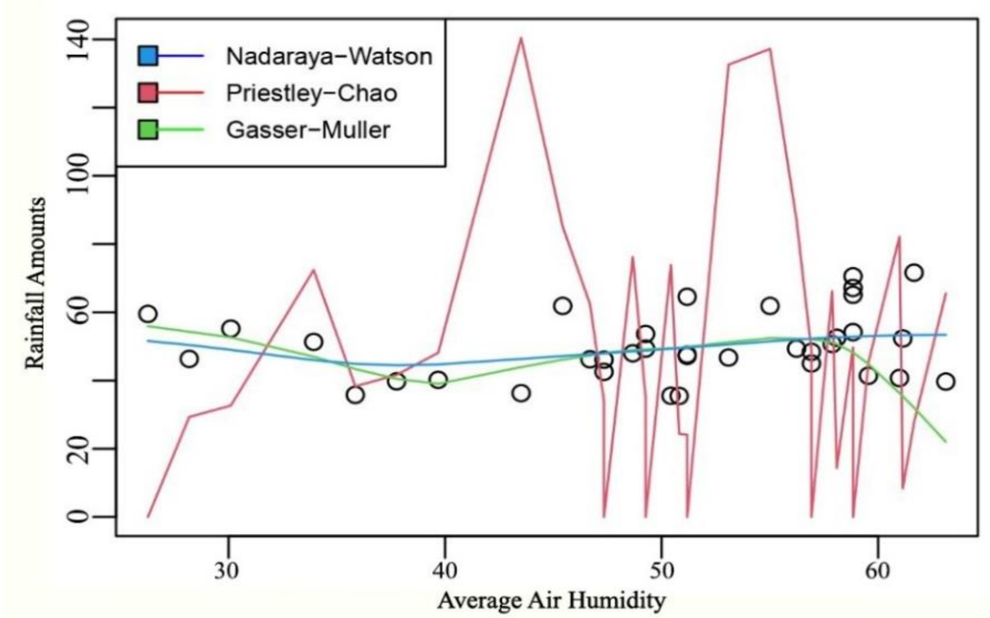

Fig 2. Curve of Nadaraya-Watson, Priestley-Chao, and Gasser-Muller Estimator

Figure 2 shows that the Nadaraya-Watson estimation curve (blue line) is close to some of the original data. This shows that the function of the Nadaraya-Watson estimator is quite good for estimating rainfall data in Lampung Province. While the Gasser-Muller estimator (green line) even though it is close to some of the original data, it is not as good as the Nadaraya-Watson estimator. This shows that the Gasser-Muller Estimator function is not good enough to estimate rainfall data in Lampung. Compared to the two, the Priestley-Chao estimation curve (red line) fluctuates very differently. This shows that Priestley-Chao cannot predict the distribution of the data well.

Based on the comparison of the estimator curve above, to be more convincing, the MSE and MAPE values are calculated as presented in Table 3.

Table 3. MSE and MAPE of the three kernel estimators

No.	Estimator	MSE	$MAPE$
1.	Nadaraya-Watson	5,9773	0,1441
2.	Priestley-Chao	145,238	25,3602
3.	Gasser-Muller	43,6816	0,1935

From Table 3 above, we can compare the three estimators by looking at the smallest guess error value. It can be seen that the Nadaraya-Watson estimator has the smallest error value with MSE = 5.9773 and MAPE = 0.1935 compared to the Priestley-Chao and Gasser-Muller estimators. This shows that the best estimator used in estimating the distribution of rainfall data in Lampung Province is the Nadaraya-Watson estimator.

5. Conclusion

In conclusion, based on the GCV, MSE and MAPE values in the results of this study, it can be stated that the Nadaraya-Watson estimator is able to predict the distribution of rainfall data in Lampung Province very well compared to the other two estimators. estimator with h=5.564, MSE = 5.9773 and MAPE = 0.1935.

References

[1] S. J. Sheather, Density estimation, “Statistical Science,” vol.19,no.4, pp. 588–597, 2004.

[2] S. Halim, and I. Bisono, “Kernel functions on nonparametric regression methods and their applications,” *Journal of Industrial Engineering*, vol. 8, no.1, pp. 73–81, 2006.

[3] N. Herawati, K. Nisa, and E. Setiawan, "The Optimal Bandwidth for Kernel Density Estimation of Skewed Distribution: A Case Study on Survival Time Data of Cancer Patients," *Proc. National Seminar on Quantitative Method*, vol.1, no.1, pp. 380-388, 2017.

[4] M.A. Dakhil and J. N. Hussain, "A Comparative Study of Nonparametric Kernel estimators with Gaussian Weight Function," *Journal of Physics: Conference Series*, vol.18, no.1, pp. 1-10, 2021.

[5] I. Puspitasari, Suparti, and Y. Wilandari, "Analysis of the Composite Stock Price Index (JCI) Using the Kernel Regression Model," *Gaussian Journal*, vol.1, no.1, pp. 93–102, 2012.

[6] N.A.K.. Rifai, Kernel Nonparametric Regression Approach on Composite Stock Price Index Data, STATISTICS," *Journal of Theoretical Statistics and Its Applications*, vol.19, no.1, pp. 53–61, 2019.

[7] D. Suparti, I. P. Safitri, Sari, and A. R. Devi, "Analysis of Inflation Data in Indonesia Using the Kernel Regression Model," *Proceedings of the National Seminar on Statistics*, pp. 499-509, 2013.

[8] Suparti, "Analysis of Inflation Data in Indonesia After the 2013 Increase in TDL and BBM Using the Kernel Regression Model," *Statistical Media*, vol.6, no.2, pp.103-112, 2013

[9] G. Rubio, H. Pomares, L. J. Herrera, and I. Rojas, "Kernel methods applied to time series forecasting," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*, Springer-Verlag, pp.782–789, 2007

[10] S. Saidi, N. Herawati, K. Nisa, and E. Setiawan, "Nonparametric Modeling Using Kernel Method for the Estimation of the Covid-19 Data in Indonesia During 2020," *International Journal of Mathematics Trends and Technology*, vol. 67, no.6, pp. 136-144, 2021.

[11] W. Syisliawati, Wibowo, and I. N. Budiantara, "Regression Spline Truncated Curve in Nonparametric Regression," *Proc. of 3rd International Conference on Research Implementation and Education of Mathematics and Science*, vol.18, no.1 , pp. 115-122, 2016.

[12] M. Y. Mustafa and Z. Y. Algamal, "Smoothing Parameter Selection in Kernel Nonparametric Regression Using Bat Optimization Algorithm," *Journal of Physics: Conference Series*, vol.1897, no.1,pp. 1-9, 1897.

[13] S. Molchanov and S. N. Chiu, "Smoothing Techniques and Estimation Methods for Nonstationary Boolean Models with Applications to Coverage Processes," *Biometrics*, vol.87, no.2, pp. 265-283, 2012.

[14] M. Rosenblatt, "Remarks on Some Nonparametric Estimates of a Density Function," *The Annals of Mathematical Statistics*, vol.33 , pp.832-837, 1956

[15] E. Parzen, "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, vol.33 , pp.1065-1076, 1962.

[16] W.R. Schucany, "On nonparametric regression with higher-order kernels," *Journal of Statistical Planning and Inference*, vol.23, no.2 , pp. 145-151, 1989.

[17] S. Weglarczyk, "Kernel Density Estimation and Its Application," *ITM Web of Conferences*, vol.23, no.374 , pp. 1-8, 2018.

[18] Nadaraya, E. A, "Some new estimates for distribution functions," *Theory of Probability & Its Applications*, vol.9, no.3, pp. 497–500, 1964.

[19] S. Demir and O. Toktamis, On The Adaptive Nadaraya-Watson Kernel Regression Estimators, *Hacettepe Journal of Mathematics and Statistics*, vol.39, no.3, pp. 429-437, 2010.

[20] K. Konecna, The Priestley-Chao Estimator of Conditional Density with Uniformly Distributed Random Design, *statistics*, vol.98, no.3, pp. 283-294, 2019.

[21] M. E. Priestley, and M. T. Chao, "Nonparametric function Fitting, Journal of The Royal Statistical Society, vol.34 ,pp. 385-392, 1972.

[22] T. Gasser, and H.-G. Müller," Kernel Estimation of Regression Functions. In Smoothing Techniques for Curve Estimation, Springer, Berlin, pp.23-68, 1979.

[23] P. Babilua, E. Nadaraya, and G. Sokhadze, "Functionals of Gasser–Muller Estimators," *Turkish Journal of Mathematics*, vol.38, no.1, pp. 1090–1101, 2014.

[24] M. C. Jones, J. S. Marron, and S. J. Sheather, "A Brief Survey of Bandwidth Selection for Density Estimation," *Journal of the American Statistical Association*, vol.91, no.433, pp. 401-407, 1996.

[25] L. R. Cheruiyot, G. O. Orwa, and O. E. Otieno, "Kernel Function and Nonparametric Regression Estimation: Which Function Is Appropriate," *African Journal of Mathematics and Statistics Studies*, vol.3, no.3, pp. 51-59, 2020.

[26] O.M. Eidous, M.A.A.S Maie, and M.A Ebrahem, "A Comparative Study for Bandwidth Selection in Kernel Density Estimation," *Journal of Modern Applied Statistical Methods: JMASM*, vol.9, no.1, pp.263-273, 2015.

[27] H. Dhaker, E. H. Deme, P. Ngom, and M. Mbodj, "New Approach for Bandwidth Selection in The Kernel Density Estimation Based on B-Divergence," *Journal of Mathematical Sciences: Advances and Applications*, vol.51, no.1, pp. 57–83, 2018.

[28] G. Kauermann and J. D. Opsomer, Generalized Cross-Validation for Bandwidth Selection of Backfitting Estimates in Generalized Additive Models, *Journal of Computational and Graphical Statistics*, vol.13, no.1, pp. 66-89, 2014.