Combination of high order kernel estimators for estimation of fruit trees biomass on critical land

Zulfikar¹, Munawarah², Ambar Susanti³

¹Informatics Science Study, Faculty of Informatics Technology, University of KH. A. Wahab Hasbullah, Jl. Garuda no. 09 Tambakberas Jombang, East Java, Indonesia
²Informatical System Study, Faculty of Informatics Technology, University of KH. A. Wahab Hasbullah, Jl. Garuda no. 09 Tambakberas Jombang, East Java, Indonesia
³Agroecotechnology Study, Agriculture Faculty, University of KH. A. Wahab Hasbullah, Jl. Garuda no. 09 Tambakberas Jombang, East Java, Indonesia

E-mail : unwaha@unwaha.ac.id

Abstract. This study aims to estimate fruit tree biomass with a high-order Gaussian kernel estimator. The source of the data was obtained by estimating the relationship between tree diameter (x) and tree biomass (y) which was obtained through the allometric equation \( y = 0.11x^{2.62} \) for branched trees. Data samples were taken as many as 250 obtained from 5 sub-district locations in critical land on longan and mango fruit tree vegetation. The relationship between x and y is then estimated in the form of an equation \( y_i = m(x_i) + e_i \), where, \( i = 1, 2, n \). The results of the analysis using the kernel estimator showed that the tree biomass estimation resulted in a smooth regression curve at an optimal bandwidth (h) score of 1.3 (GCV = 122.92) longan and 2.6 (GCV = 3523.25) mango. The relationship between MSE and bandwidth shows a very strong negative correlation with a score of \( r = -0.959 \) (p.value = 0.04) in the estimated longan biomass and \( r = -0.964 \) (p.value = 0.03) in the estimated mango biomass and is significant (\( \alpha = 0.05 \)). This means that a parallel increase in the bandwidth score will decrease the MSE score when GCV is controlled.

1. Introduction

Indonesia has been designated as the country with the second-largest vegetation diversity in the world [1]. As a tropical area, Indonesia is a potential source of new species that have more than half the diversity of flora and fauna in the world, so it is suspected as the richest source of discovery of new types of macro-organisms [2]. At this time, the condition of land resources and the environment in Indonesia is increasingly alarming, indicated by the wide area of critical land that has an impact on the decline in vegetation diversity. Critical lands are generally more fragile, prone to erosion, less productive, and not easy to manage, thus leading to a decline in land quality [3]. The Ministry of Environment and Forestry (2018) reports that the area of critical and very critical land in Indonesia in 2015 without DKI Jakarta was approximately 24,303,294 ha consisting of 19,564,911 ha critical and 4,738,384 ha [4]. The expansion of critical land is caused by several things, including forest destruction, expansion of agricultural areas that are not in accordance with the carrying capacity of the environment, increasing population pressure, and uncontrolled fires [5].

The development of a conservation success model must be able to take an approach to save the environment from ecological damage, one of which is by estimating tree biomass. The concept of
developing a conservation success model that combines the architectural characteristics of trees with the biomass potential of each individual tree so that with a practical method of measuring biomass estimation, the potential of the land can be known quickly so that efforts to save the environment are immediately resolved. Estimation of tree biomass can be done using the analytical method of a combination of a High Order Kernel estimator with dendrometric analysis. The dendrometric analysis method is the concept of measuring tree biomass using tree sample measurements on the trunk, branches, and canopy [6]. This method is expected to provide more accurate and practical tree biomass estimation results and not leave damage to trees and forests. The conservation success model in critical land that was built is to establish a relationship between fruit tree productivity and growth limiting factors in the environmental and a biotic and socio-economic factors and culture in the research area so that it is hoped that a conservation success model will be formed as a recommendation for stakeholders to make policies for the program conservation in the future.

Measurement of tree biomass, especially to measure tree biomass components such as leaves, branches, and roots so that it does not take a long time and reduces costs. Empirical relationships can be used to estimate the total biomass of biometric variables such as diameter at breast height or tree height [7], where the empirical relationship between the components of tree biomass forms an allometric equation. Because the measurement of tree biomass in the field takes a long time, especially to measure tree biomass components such as leaves, branches, and roots, and is not small, empirical relationships can be used to estimate total biometric biomass. Variables such as diameter at breast height or tree height [8]. Therefore, it is natural that the estimation of the tree and forest biomass has become a topic of long-term research such as that of Kunze in 1873 and Burger in 1929 [9]. Krisnawati et al. has compiled 807 biomass allometric models and tree volume allometric models in several types of forest ecosystems, of which 437 allometric models are used for estimating tree biomass components and 370 allometric models for estimating several types of tree volume [10]. Almost all types of major forest ecosystems in Indonesia are available with allometric models of tree biomass and/or volume, although they are not evenly distributed across major islands in Indonesia. Thus, research on the formulation of an allometric equation model for estimating the biomass of a tree species in an ecosystem type is still needed to enrich the existing data throughout the archipelago.

Almost all estimates of biomass studies focused on the application of linear and nonlinear regression models [11-13]. However, the mapping between ground surface parameters and SAR imagery is always very complex because of the strong nonlinearity. Regression models based on real data measurements cannot provide a clear enough relationship. The traditional allometric method in the form of the Schumacher-Hall equation provides less accurate estimates of biomass, where the logarithmic transformation has a weakness that cannot increase the validity of the estimation [14]. Therefore, an estimation method is needed that is able to provide more accurate results with large amounts of data, besides that the estimation technique must be flexible and does not require regression assumptions. The kernel estimator is a regression approximation technique that does not require an assumption of normality. Zulfikar states that the High Order Kernel Estimator has a smaller MSE value and is an indicator for choosing the best estimator [15].

Growth factors are simpler than constructing and implementing, but allometric models are preferred because of the increased flexibility to describe tree architectural variations and biomass compartmentalization [16]. An important aspect in the quantification of individual tree biomass is the large natural variability in the data, especially for species native to tropical and subtropical regions. A single mathematical formulation may not be able to reproduce such a great natural variation. These factors affect the quality of the fit model and can give incorrect estimates. Another feature of the allometric model is that when using the regression technique several assumptions must be reached. These assumptions are as follows: additive and linearity, residual independence, homoscedasticity, and residual normality [17]. Related to this, this study aims to estimate fruit tree biomass on critical land using a high-order kernel estimator based on stem diameter to the amount of plant biomass to obtain a regression curve estimator obtained by estimating these parameters [18].
2. Research Method
2.1. Study Area
This research was conducted in Jombang Regency, East Java Province, from Mei - July 2021. The study location was determined by purposive sampling, which is one of the techniques using certain considerations [19]. Based on that, the researchers determined the location by considering the conditions of vegetation density and geographical site, by making indirect visual observations using the Google Earth application, while also conducting surveys and direct visualizations. The Research Map Design was generated with the R version 3.6.1 application program, as shown in Figure 1.

Figure 1. Map of Research Locations in Jombang Regency, East Java Province on 5 sub-district research sites, namely A = Plandaan, B = Kabuh, C = Ploso, D = Kudu and E = Ngusikan. Each site was taken as much as 50 data samples at the village level.
2.2. Data source and research variable
The data used in this study is primary data in Jombang Regency. The regression model that was built was to obtain the form of the relationship between trunk diameter (X) and tree biomass (Y). Plant parameters measured were stem diameter at a height of 1.3 meters above ground level and stem diameter in centimetres. The fruit trees studied were longan (Dimocarpus longan L) and mango (Mangifera indica L), each of which took 250 data samples. Measurement of tree biomass using the branching tree allometric equation formula [20]:

\[ Y = 0.11 \rho X^{2.62} \]  

(1)

Where:
Y = aboveground biomass (Kg/tree)  
\( \rho \) = density of wood species (g.cm\(^{-2}\))  
X = tree diameter (cm)

The fruit trees whose biomass was measured were longan and mango with wood densities of 0.91 and 0.58, respectively [21].

2.3. Analysis method
The steps taken in this study were to perform the estimation stage, then analyze the kernel regression model using the application program of R 3.61 version. Biomass estimation using nonparametric regression model based on local polynomial kernel estimator with estimation steps as follows:

a. observational data are obtained \((y_i, x_i)\) that meet the nonparametric regression;

\[ Y_i = m(x_i) + \epsilon_i, \quad i = 1, 2, 3, \ldots, n \]  

(2)

b. plotting data in pairs: \((y_i, x_i)\), \( i = 1, 2, \ldots, n \)
c. determines the type and weighted function of the Gaussian Kernel;
d. determine the matrix \( A(h) \) of size N x N;
e. chooses the order of the polynomial \( p \) and minimizes the optimal bandwidth value;

\[ GCV = \frac{n^{-1} \sum_{i=1}^{n} [y_i - \hat{y}_i]^2}{(n^{-1} tr[1 - A(h)])^2} \]  

(3)
f. simultaneously models the local polynomial order \( p \) and the optimal bandwidth value from step (4);
g. calculates the average value of the square of the error;

\[ MSE(h) = n^{-1} \sum_{i=1}^{n} (y_i - \hat{y}_i) \]  

(4)
h. gets a biomass estimation model with a local polynomial kernel estimator.

3. Research Results
To apply the Gaussian kernel estimator, research data from the results of measuring the diameter of the longan and mango tree trunks are used, each with 250 sample data. It will be seen the relationship between trunk diameter (x) and the amount of tree biomass (y) based on equation (2) with m regression curve. The plot between x and y is given in Figure 2 & 3.
From Figure 2 & 3 above is can be seen that there is no clear pattern regarding the relationship between x and y. Furthermore, nonparametric regression, especially Kernel order 2 estimators, is used to estimate m. The kernel estimate is given by:

\[
\hat{m}(x, 0, h) = \frac{\sum_{i=1}^{n} K_h(x_i - x) y_i}{\sum_{i=1}^{n} K_h(x_i - x)} 
\]  

(5)

With the Gaussian kernel. The second order Gaussian kernel form is obtained by substituting the kernel function:

\[
K(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2 / 2) 
\]  

(6)

Into equation (5).

First, the optimal bandwidth selection for each kernel is given by using the GCV method. The GCV function is given by:

\[
\text{GCV}(h) = n^{-1} \sum_{i=1}^{n} \frac{\left\{ y_i - \hat{m}(x_i) \right\}^2}{\left\{ 1 - n^{-1} \text{tr}(A(h)) \right\}^2}
\]  

(7)

With A (h) obtained from the equation:

\[
\hat{m}(x, p = 0, h) = A(h) y
\]

3.1. Kernel regression function

Kernel order 2 regression used for the estimation of longan biomass was built with the approach of minimizing the GCV (h) function to obtain the optimum bandwidth value. For some bandwidths, the GCV (h) values are listed in Table 1 and 2 below:

| No. | GCV(h)   | Bandwidth(h) | No. | GCV(h)   | Bandwidth(h) |
|-----|----------|--------------|-----|----------|--------------|
| 1.  | 126.4935 | 0.5          | 10. | 122.9487 | 1.4          |
| 2.  | 125.8920 | 0.6          | 11. | 123.0502 | 1.5          |
| 3.  | 125.1507 | 0.7          | 12. | 123.2061 | 1.6          |
| 4.  | 124.4403 | 0.8          | 13. | 123.4125 | 1.7          |
| 5.  | 123.8527 | 0.9          | 14. | 123.6715 | 1.8          |
| 6.  | 123.4140 | 1.0          | 15. | 123.9822 | 1.9          |
| 7.  | 123.1241 | 1.1          | 16. | 124.3464 | 2.0          |
| 8.  | 122.9667 | 1.2          | 17. | 124.7719 | 2.1          |
| 9.  | 122.9177 | 1.3          | 18. | 125.2594 | 2.2          |
Table 1 shows that the optimum bandwidth value is 1.3 with a minimum GCV value of 122.9177 in the estimated longan tree biomass. The same thing is also shown in Table 2, where the optimum bandwidth value is 2.6 with a minimum GCV value of 3523.248 in the estimated mango tree biomass as shown in Table 2. The optimum h in kernel regression is used to estimate the regression curve obtained by estimating these parameters.

**Table 2. Bandwidth estimator value kernel order 2 on the measurement of mango tree biomass.**

| No. | GCV(h)   | Bandwidth(h) | No. | GCV(h)   | Bandwidth(h) |
|-----|----------|--------------|-----|----------|--------------|
| 1.  | 3631.274 | 1.4          | 10. | 3527.719 | 2.3          |
| 2.  | 3611.305 | 1.5          | 11. | 3525.106 | 2.4          |
| 3.  | 3593.866 | 1.6          | 12. | 3523.632 | 2.5          |
| 4.  | 3578.582 | 1.7          | 13. | 3523.248 | **2.6**      |
| 5.  | 3565.415 | 1.8          | 14. | 3523.755 | 2.7          |
| 6.  | 3554.208 | 1.9          | 15. | 3525.064 | 2.8          |
| 7.  | 3544.951 | 2.0          | 16. | 3527.008 | 2.9          |
| 8.  | 3537.562 | 2.1          | 17. | 3529.537 | 3.0          |
| 9.  | 3531.791 | 2.2          | 18. | 3532.424 | 3.1          |

The GCV (h) function given in Figure 2 shows that the GCV (h) function forms a curve at the minimum value to obtain the optimum bandwidth value. This curve image emphasizes that minimizing the score of the GCV (h) function will get the optimum bandwidth (h) value, both shown in the estimated biomass of longan and mango.

**Figure 4.** Stages of forming a smooth regression curve on the kernel order 2 estimators for the estimation of longan biomass. Regression curves of various values of bandwidth (h).

**Figure 5.** Minimizing the GCV(h) function.

Figure 4 shows that the formation of the kernel regression curve for the estimation of longan tree biomass shows that at h = 0.1 it forms a rough curve, while at h = 1.5 the regression curve begins to look smooth. In order to obtain a smooth curve with the optimum h value, it is necessary to minimize the GCV value as shown in Figure 5, where the minimum GCV value in this kernel estimate is obtained with the optimum h value = 1.3.
Figure 6. Kernel regression of longan biomass estimation with optimum bandwidth (h = 1.3).

Furthermore, the kernel regression curve with the optimum h is obtained as shown in Figure 6 and becomes the ideal kernel regression function for estimating the amount of longan tree biomass. The same thing is also shown in Figure 7, which is the formation of a smooth kernel regression curve at the optimum bandwidth of 2.6.

Figure 7. Stages of forming a smooth regression curve on the kernel estimator order 2 for mango biomass estimation regression curves of various values of bandwidth (h).

Figure 8. Minimize the GCV(h) function.

In figure 9, it can be seen that the stages of forming a smooth regression curve in the kernel order estimator 2 for mango biomass estimation. Formation of regression curves from various values of bandwidth (h) at scores of 0.5, 2.0 and 5.0. Minimizing the GCV (h) function, the optimum bandwidth (h) is obtained with a score of 2.6. 2nd order kernel regression function with optimum bandwidth (h).
Figure 9. The 2nd order kernel regression function with the optimum bandwidth (h = 2.6).

3.2. Validation Test

The minimum GCV (h) function for determining the optimum bandwidth will get a smooth kernel regression function, which will then form a relationship between MSE and bandwidth. In figure 3 it can be seen that the greater the bandwidth score, the smoother the curve where the limit of the h value cannot be determined to what extent the kernel regression curve is optimal. As a form of validation of the results of using high-order kernel regression, MSE scores are needed and compare them with bandwidth and GCV values.

Table 3. MSE and GCV values and bandwidth (h) Kernel Estimator on Longan and Mango Biomass Estimation.

| Biomass Estimation | GCV        | Bandwidth | MSE         |
|--------------------|------------|-----------|-------------|
| a. Longan          | 151.6033   | 0.1       | 88.83703    |
|                    | 126.4935   | 0.5       | 48.88685    |
| **122.9177**       | **1.3***   | **43.09379** |
|                    | 123.0502   | 1.5       | 41.66774    |
|                    | 131.7924   | 3.0       | 28.41932    |
| b. Mango           | 4166.082   | 0.5       | 8448.196    |
|                    | 3523.632   | 2.5       | 7753.965    |
| **3523.248**       | **2.6***   | **7670.941** |
|                    | 3673.627   | 5.0       | 7591.417    |
|                    | 4273.334   | 7.5       | 6988.180    |

Note: * = h optimum

The relationship between MSE and bandwidth to see the validation of the kernel regression used is then carried out by a partial correlation test using the Pearson method where GCV is the control variable as shown in Tables 4 and 5.
Table 4. Partial correlation between MSE and Bandwidth with GCV as a control in the estimation of longan biomass using the Pearson method.

|        | Bandwidth | GCV  | MSE     |
|--------|-----------|------|---------|
| Bandwidth | 1.000     | 0.893| -0.960* |
| p-value  | 1.000     | 0.107| 0.040   |
| GCV      | 0.893     | 1.000| 0.961   |
| p-value  | 0.107     | 1.000| 0.038   |
| MSE      | -0.960*   | 0.961| 1.000   |
| p-value  | 0.040     | 0.038| 1.000   |

The significance of the partial correlation at the level of P-value < (α = 0.05)

The partial correlation between bandwidth and MSE at a score of -0.9598964 is a strong negative correlation. This shows that the parallel increase in the bandwidth value will decrease the MSE score when GCV is controlled. The amount of p-value in this partial correlation is 0.04010363 which is statistically significant at = 0.05. The same thing is shown from the results of the partial correlation test between bandwidth and MSE at a score of -0.9642895 which is a strong negative correlation. A parallel increase in the score in bandwidth value will decrease the MSE score when GCV is controlled with a p-value of 0.03571046 which is statistically significant at = 0.05. The level of difference in MSE, GCV, and bandwidth values can be seen in Figure 5.

Table 5. Partial correlation between MSE and Bandwidth with GCV as a control on mango biomass estimation using Pearson's method.

|        | Bandwidth | GCV  | MSE     |
|--------|-----------|------|---------|
| Bandwidth | 1.000     | 0.651| -0.964* |
| p-value  | 1.000     | 0.349| 0.035   |
| GCV      | 0.651     | 1.000| 0.613   |
| p-value  | 0.349     | 1.000| 0.387   |
| MSE      | -0.964*   | 0.613| 1.000   |
| p-value  | 0.036     | 0.387| 1.000   |

The significance of the partial correlation at the level of P-value < (α = 0.05)

In Figure 10 and 11, it can be seen that the bar graph formed between MSE, GCV, and bandwidth forms the same pattern both in the estimation of longan and mango biomass. Relationship between MSE, bandwidth and GCV, where MSE and bandwidth showed a strong negative correlation (r = -0.82) for the estimated longan biomass and (r = -0.94) for mango. While the minimum GCV value to show the optimum bandwidth value. The MSE score pattern shows a decrease as the bandwidth value increases, while the GCV score shows a decreasing pattern and then increases with increasing bandwidth scores and decreasing MSE scores. Determining the optimum bandwidth at the minimum GCV value finally determines the MSE value as a valid score for the kernel regression model that formed.
4. Discussion
The main obstacle in estimating tree biomass is the use of an estimator that is not careful in choosing the method of analysis, where estimation data is often found using the classical regression approach. Moreover, many ecologists conducting studies of almost all biomass estimates focused on applying linear and nonlinear regression models [11-13]. It is also added that the results of previous studies show that biomass estimation has parameters that deal with complex environments where the mapping between soil surface parameters and SAR images is always very complex because of the strong nonlinearity. Regression models based on real data measurements cannot provide a clear enough relationship. The traditional allometric method in the form of the Schumacher-Hall equation provides less accurate estimates of biomass, where the logarithmic transformation has a weakness that cannot increase the validity of the estimation [14]. Consequently, the use of nonparametric models, which have several advantages for example no need of normality, should be used within the forestry and environmental purposes of obtaining robust and understandable models [23]. Higher-order kernels have been suggested for use in nonparametric curve estimation, it is seen that they work well when the curvature of the target curve is roughly constant, and work poorly when there are abrupt changes in curvature on neighborhoods that are the size of the bandwidth [24].

The biomass estimation method using a high-order kernel estimator is very helpful in overcoming these obstacles because the variable relationship model formed is flexible without the need to predict the form of the relationship between the two variables. The results of this study indicate that the kernel estimator is able to produce tree biomass estimates for both longan and mango by estimating the regression form by minimizing the GCV score and succeeding in obtaining the optimum bandwidth with 1.3 for longan biomass estimation and 2.6 for mango. The use of optimum h in kernel regression is used to estimate the regression curve obtained by estimating these parameters [18]. The bandwidth selection approach has been found that the GCV method is asymptotically optimal for the mean value in the case of data independence [22]. Hall et al. (1995) added that subsequent studies of asymptotic properties at optimal bandwidth under different levels of dependencies [25]. The direct plug-in approach, where unknown functional that appear in expressions for the asymptotically optimal bandwidths are replaced by kernel estimates, is used [26,27].

The validation of the regression method built with this approach is used by measuring the MSE, which can show a low MSE score, where the use of the Kernel Estimator estimator has a smaller MSE value and is an indicator for choosing the best estimator [17]. The relationship between Bandwidth and MSE partially with controlled GCV shows that a strong negative correlation is obtained, where the greater the bandwidth, the lower the MSE score. The role of bandwidth as a curve smoothing with an increasing score will be controlled by decreasing the MSE value, thus the role of GCV will show the limit of its optimum bandwidth value. In this study, it has been shown that the relationship between bandwidth and MSE has a strong and significant negative correlation at the (α) 0.05 level, with an r-value of -0.960 in the estimated longan biomass and an r of -0.964 for mangoes. Kim et al. 2016 add
that the use of low-scoring MSE for regression model validation is more recommended than the standard method. Another advantage of using a high-order kernel estimator is that the method steps are more concise because there is no need to use the classical requirements as in the use of parametric regression. In this nonparametric regression estimation, the regression formed is more indicated in the formation of a regression curve from a rough shape to a smooth limit that is controlled by GCV.

5. Conclusion
The use of a high-order kernel estimator is able to overcome the difficulties in estimating tree biomass for both longan and mango. The results of the biomass estimation show that the resulting regression curve forms a smooth curve at the optimal bandwidth with a score of 1.3 for longan and 2.6 for mango from the results of minimizing the GCV(h) function. The best estimator validation used the selection of the MSE score and its relationship with Bandwidth on the controlled GCV value. The results show that there is a strong negative correlation where the greater the bandwidth value, the smaller the MSE.

Acknowledgments
Researchers say thank you for the help of multi-year grants for Applied Research in 2021-2022, of Research and Community Service to the Directorate General of Strengthening Research and Development, Ministry of Research, Technology and Higher Education. Thanks are also conveyed to the Jombang Regency Environmental Service (DLH) for the collaborative research carried out so far, the Agriculture Service, and the entire academic community of KH University. A. Wahab Hasbullah.

References
[1] Whitten A, Anwar J, Damanik S, and Hisyam N 1984 The Ecology of Sumatra Oxford University Press 512 pp. https://doi.org/10.1046/j.1420-9101.1989.2020154.x
[2] Gandjar, Indrawati and Wellyzar Sjamsuridzal 2006 Mikologi Dasar dan Terapan Jakarta: Yayasan Obor
[3] Naidu, LGK, S Dharumaraja, M Lalitha, S Srinivas, V Ramamurthy and S K Singh. 2014 Categorization and delineation of prime and marginal lands of Andhra Pradesh for different uses. Agropedology, 24 (02), 253-261
[4] Kementerian Lingkungan Hidup Kehutanan. 2018 Rencana Strategis 2015-2019 Kementerian Lingkungan Hidup dan Kehutanan: Jakarta
[5] Muharam 2011 Pengembangan Model konservasi Lahan dan Sumberdaya Air dalam rangka Pengentasan Kemiskinan Solusi Unisika 10 (20) - Ed. Sept-Nop
[6] Fernadez-Puratich, Jose V Oliver-Villanueva, David Alfonso-Solar and Elisa Penalvo-Lopez 2013 Quantification of Potential Lignocellulosic Biomass in Fruit Trees Grown in Mediterranean Regions BioResources 8 (1), 88-103
[7] Pilli R, Anfodillo T, and Carrer M 2006 Towards a functional and simplified allometry for estimating forest biomass For Ecol Manage 237: 583–593
[8] Cordero, L D P & M Kanninen 2003 Aboveground biomass of Tectona grandis plantation in Costa Rica Journal of Tropical Forest Science 15 (1): 199-213
[9] Fehrmann, L & C Kleinn 2006 General considerations about the use of allometric equations for biomass estimation on the example of Norway spruce in central Europe Forest Ecology and Management 236: 412-421
[10] Krisnawati H, W C Adinugroho & R Imanuddin 2012 Monograf: Model-model allometrik untuk pendugaan biomassa pohon pada berbagai tipe ekosistem hutan di Indonesia Pusat Penelitian dan Pengembangan Konservasi dan Rehabilitasi. Badan Penelitian dan Pengembangan Kehutanan,Bogor. Indonesia.
[11] Kasischke E S, Christensen N L and Bourgeau-Chavez L L, 1995 Correlating radar backscatter with components of biomass in loblolly pine forest. IEEE Trans. Geosci.
[12] Polatin P F, Sarabandi K and Ulaby F T, (1994). An iterative inversion algorithm with application to the polarimetric radar response of vegetation canopies. *IEEE Trans. Geosci. Remost Sensing* 32, pp. 643–659

[13] Rignot E, Way J B, Williams C and Viereck L, (1994). Radar estimates of aboveground biomass in boreal forests of interior Alaska. IEEE Trans. Geosci. Remost Sensing 32, pp. 62–71

[14] Sanquetta C R, Jaime Wojciechowski, Ana P Dalla Corte†, Alexandre Behling†, Sylvio Pélico Netto†, Aurélio L. Rodrigues*†and Mateus N I Sanquetta†. 2015. Comparison of data mining and allometric model in estimation of tree biomass. *BMC Bioinformatics*, 16:247

[15] Zulfikar 2010 Kernel Order Tinggi untuk Estimasi Value at Risk (VaR) Manajemen Resiko Tenaga Kerja, Proseding Seminar Nasional Manajemen Teknologi XII, MMT ITS, Surabaya. Hal : C 1 – 6

[16] Petrokofsky G, Kanamaru H, Achard F, Goetz S J, Joosten H, Holmgren P, et al. 2012. Comparison of methods for measuring and assessing carbon stocks and carbon stock changes in terrestrial carbon pools. How do the accuracy and precision of current methods compare? A systematic review protocol. *Environmental Evidence*. 1:1–22

[17] Osborne J, and Waters E 2002 Four assumptions of multiple regression that researchers should always test *Pract Assessment Res Evaluation*; 8 (2):1–8

[18] Hardle, W 1991 *Applied nonparametric regression* Cambridge: Cambridge University Press

[19] Fachrul, M F 2007 *Metode Sampling Bioekologi Bumi Aksara*, Jakarta, Indonesia

[20] Hairiah J, Sitompul S M, van Noordwijk M, Palm C A 2001 *Methods for sampling carbon stocks above and below ground* Bogor: ICRAF

[21] Baruna T, Dwi A, dan Herlina D, 2017 Pendugaan stok karbon di kawasan hutan Pendopo Gubernur Kota Pontianak *Jurnal Hutan Lestari* Vol. 5 (2): 547–555

[22] Hart J D & Vieu P (1990) Data-driven bandwidth choice for density estimation based on dependent data *The Annals of Statistics* 18, 873–890

[23] Dario Domingo, María Teresa Lamelas-Gracia, Antonio Luis Montalegre Gracia & Juan de la Riva-Fernández (2017) Comparison of regression models to estimate biomass losses and CO2 emissions using low-density airborne laser scanning data in a burnt Aleppo pine forest, *European Journal of Remote Sensing*, 50:1, 384-396, DOI: 10.1080/22797254.2017.1336067.

[24] Marron J (1994) Visual Understanding of Higher-Order Kernels *Journal of Computational and Graphical Statistics, 3*(4), 447-458. doi:10.2307/1390905

[25] Hall P, Lahiri S N & Truong Y K (1995) On bandwidth choice for density estimation with dependent data. *The Annals of Statistics* 23, 2241–2263

[26] Sheather S J and M C Jones (1991) A reliable data-based bandwidth selection method for kernel density estimation. *J. Roy. Statist. Soc. B*, 683-690

[27] Wand M P and Jones M C (1995) *Kernel Smoothing*. Chapman and Hall, London