SAINTIFIK:
Jurnal Matematika, Sains, dan Pembelajarannya **Vol.12, No.2, Juli 2026, pp.215~220**
ISSN 2407-4098 (print) DOI:10.31605/saintifik.v12i2.668
ISSN 2622-8904 (online)

Estimation of Rayleigh Distribution Parameters Using the Maximum Likelihood Estimation Method on Wind Speed Data in Jayapura

Safitri Rahayu[1], Agustinus Langowuyo[2*]
[1,2]Department of Mathematics, Universitas Cenderawasih, Indonesia
Email : agustinus.langowuyo@fmipa.uncen.ac.id

Abstract

This study aims to analyze the characteristics of wind speed data and test its suitability to the Rayleigh distribution using the Maximum Likelihood Estimation *(MLE) method. The data used consists of 66 wind speed observations. The analysis begins with descriptive statistics showing an average value of 7.21 km/h, a median of 7.3 km/h, a mode of 4.9 km/h, a variance of 5.73, and a standard deviation of 2.39 km/h. The histogram results show a unimodal distribution pattern with data concentration in the range of 6–9 km/h, while the skewness value of -0.333889 indicates a relatively low negative skew so that the data distribution is still close to symmetrical. The estimation of the Rayleigh distribution parameters using the MLE method produces a scale parameter (σ) value of 5.3702, indicating a fairly large variation in wind speed data. Next, a distribution suitability test was carried out using the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests. The KS test results produced a statistic of 0.2041 with* a p-value *of 0.0069, thus rejecting the null hypothesis and declaring the data not to follow the Rayleigh distribution. Conversely, the AD test produced a statistic of 0.5278 with* a p-value of *0.1718, thus failing to reject the null hypothesis and declaring the Rayleigh distribution still appropriate for modeling the data. The difference in the results of the two tests is caused by the difference in sensitivity of the methods in detecting distribution deviations. Overall, the Rayleigh distribution can still be used to describe the probability characteristics of wind speed data, although there are some deviations between the theoretical distribution and the actual data.*

Keywords: *Rayleigh Distribution, Wind Speed, Maximum Likelihood Estimation, Kolmogorov-Smirnov, Anderson-Darling, Goodness-of Fit Test.*

1. INTRODUCTION

Wind speed is a meteorological element that plays a crucial role in various aspects of life, such as air and sea transportation, building construction planning, hydrometeorological disaster mitigation, and the development of new and renewable energy sources, particularly wind power plants. Information on wind speed characteristics is also essential for regional planning and decision-making related to atmospheric conditions. Therefore, statistical analysis of wind speed data is crucial for developing models that accurately depict data distribution patterns (Wilks, 2019; BMKG, 2024). Jayapura City, the capital of Papua Province, has a unique geographic location, situated on the northern coast of Papua Island and directly bordering the Pacific Ocean. This geographical setting influences atmospheric dynamics through the interaction between land, sea, and varying topography, causing wind speed characteristics in the region to fluctuate according to weather conditions and seasons. Consequently, information regarding the distribution of wind speed in Jayapura City is essential to support transportation, infrastructure development, and renewable energy resource management (BMKG, 2024).

In statistics, phenomena containing elements of uncertainty are generally modeled using probability distributions. A probability distribution is a mathematical function that describes how the probability of a random variable is distributed. Probability distributions are classified into discrete and continuous distributions. Continuous probability distributions are commonly used to model variables whose values lie within certain intervals, including meteorological variables such as wind speed. Various continuous probability distributions have been widely applied to meteorological data, including the Normal, Gamma, Weibull, Exponential, and Rayleigh distributions. However, each distribution possesses different characteristics; therefore, selecting an appropriate probability distribution is essential because it directly affects the results of statistical analysis and interpretation (Johnson & Wichern, 2019).

Among continuous probability distributions, the Rayleigh distribution has been extensively used to model positive-valued continuous data, particularly wind speed observations. Introduced by Lord Rayleigh in 1880, this distribution has found broad applications in meteorology, oceanography, communications, engineering, health sciences, and biology. In meteorological studies, the Rayleigh distribution is widely employed because it is capable of representing the distribution pattern of wind speed data effectively. Wind speed itself is an important meteorological variable that describes the movement of air masses from areas of higher pressure to lower pressure. Its magnitude is affected by atmospheric pressure differences, temperature, topography, altitude, and surface obstacles such as vegetation and buildings. Owing to these characteristics, the Rayleigh distribution has become one of the most frequently used statistical models for analyzing wind speed data and supporting climatological analysis, wind energy planning, weather forecasting, and environmental studies.

To obtain an appropriate probability model, the parameters of the selected distribution must be estimated accurately. One of the most widely used parameter estimation methods is Maximum Likelihood Estimation (MLE). This method determines parameter values that maximize the likelihood function, thereby producing parameter estimates that have the highest probability of generating the observed data. MLE possesses several desirable statistical properties, including consistency, efficiency, and asymptotic normality, making it one of the most reliable estimation techniques for probability distributions (Casella & Berger, 2002; Pawitan, 2001). However, parameter estimation using MLE cannot always be solved analytically because the resulting likelihood function is frequently nonlinear. Consequently, numerical optimization techniques, such as the Newton–Raphson algorithm, are commonly employed to obtain convergent parameter estimates efficiently (Nocedal & Wright, 2006).

Various previous studies have demonstrated that the Rayleigh distribution provides satisfactory modeling results for wind speed data in different regions. Nevertheless, the suitability of this distribution strongly depends on the characteristics of the observed data, implying that findings obtained in one region cannot necessarily be generalized to another. Therefore, an investigation of wind speed data in Jayapura City remains necessary to evaluate whether the Rayleigh distribution adequately represents its statistical characteristics. Based on this background, this study aims to estimate the parameters of the Rayleigh distribution using the Maximum Likelihood Estimation (MLE) method for wind speed data in Jayapura City. Furthermore, the estimated parameters are used to evaluate the suitability of the Rayleigh distribution through statistical analysis and goodness-of-fit tests, namely the Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) tests. The results of this study are expected to provide useful information regarding the characteristics of wind speed distribution in Jayapura City and serve as a reference for future studies in statistics, climatology, and meteorological data modeling.

2. METHODS

This study employed a quantitative research approach using statistical analysis to estimate the parameters of the Rayleigh distribution and evaluate its suitability for modeling wind speed data in Jayapura City. The research utilized secondary wind speed data obtained from the Meteorology, Climatology, and Geophysics Agency (BMKG) through the official website (https://www.bmkg.go.id/cuaca/prakiraan-cuaca/91.71.01.1002) or other related meteorological sources within a specified observation period. The dataset consisted of 66 wind speed observations, expressed in kilometers per hour (km/h).

The data analysis was conducted in several stages. First, descriptive statistical analysis was performed to identify the characteristics of the wind speed data, including the mean, median, mode, variance, standard deviation, skewness, and kurtosis. Histogram visualization was also employed to illustrate the distribution pattern of the observed data.

The second stage involved estimating the parameter of the Rayleigh distribution using the Maximum Likelihood Estimation (MLE) method. The MLE method estimates the parameter by maximizing the log-likelihood function, thereby obtaining the parameter value that has the highest probability of generating the observed data. For the Rayleigh distribution, the scale parameter (σ) was estimated using the maximum likelihood estimator derived from the corresponding log-likelihood function.

Finally, the suitability of the estimated Rayleigh distribution was evaluated using two goodness-of-fit tests, namely the Kolmogorov–Smirnov (KS) test and the Anderson–Darling (AD) test at a significance level of 0.05. The hypotheses tested were as follows:

- H_0: Wind speed data in Jayapura City follow the Rayleigh distribution.
- H_1: Wind speed data in Jayapura City do not follow the Rayleigh distribution.

The Rayleigh distribution was considered appropriate if the resulting *p*-value exceeded 0.05, indicating that the null hypothesis could not be rejected. Conversely, a *p*-value below 0.05 indicated that the Rayleigh distribution was not suitable for modeling the observed wind speed data.

3. RESULTS DAN DISCUSSION

3.1 Descriptive Analysis of Wind Speed Data

Descriptive analysis was conducted to provide an overview of the characteristics of the wind speed data before further statistical analysis. The dataset consisted of 66 wind speed observations with a mean of 7.21 km/h, a median of 7.3 km/h, and a mode of 4.9 km/h. The variance and standard deviation were 5.72888 km/h and 2.39351 km/h, respectively, while the skewness and kurtosis values were −0.33389 and −0.36551. These results indicate that the wind speed data are relatively balanced and suitable for subsequent statistical modeling.

Table 1. Results of Descriptive Analysis of Wind Speed Data in Jayapura City

Number of Data Samples (n)	66 km/h
Mean	7.213636 km/h
Median	7.3 km/h
Mode	4.9 km/h
Variance	5.72888 km/h
Standard Deviation	2.39351 km/h
Skewness	(-0.33389) km/h
Kurtosis	(-0.36551) km/h

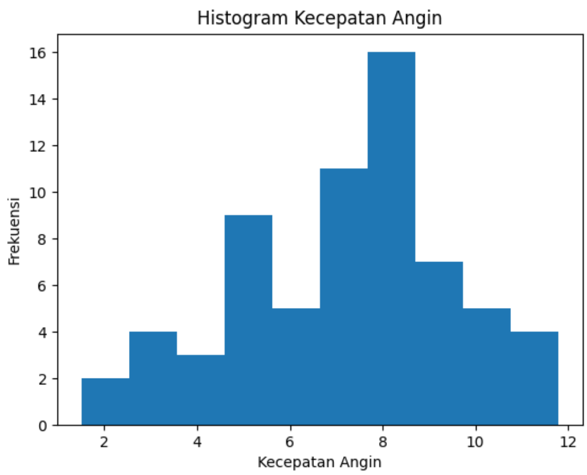

Figure 1. Histogram of Wind Speed Data

3.2 Estimation of Rayleigh Distribution Parameters

The scale parameter of the Rayleigh distribution was estimated using the Maximum Likelihood Estimation (MLE) method. Based on the estimation procedure, the obtained scale parameter (σ) was **5.3702**. This estimated value was subsequently used to construct the Rayleigh distribution model and perform the goodness-of-fit analysis.

In the Rayleigh distribution, the parameter being estimated is the scale parameter (σ), which indicates the degree of data dispersion. Parameter estimation is performed using the Maximum Likelihood Estimation (MLE) method. The probability density function of the Rayleigh distribution is given by:

$$f(x;\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x > 0$$

Based on this function, the log- *likelihood function is obtained* as follows:

$$\ell(\sigma) = \sum_{i=1}^{n} \ln(x_i) - 2n\ln(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2$$

likelihood function with respect to $\sigma$is equated to zero so that the estimator of the scale parameter of the Rayleigh distribution is obtained as

$$\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} x_i^2}$$

Based on 66 wind speed observations, the estimated scale parameter value was 5.3702. This value indicates that the Rayleigh distribution used to model wind speed data in Jayapura City has a scale parameter of 5.3702, which is then used in the formation of the distribution model and data fit testing.

3.3 Goodness-of-Fit Test

Distribution fit tests were conducted using the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests to determine the suitability of the Rayleigh distribution to wind speed data in Jayapura City. Testing was conducted at the significance level $\alpha = 0{,}05$.

Based on the results of the Kolmogorov-Smirnov (KS) test, the test statistic value was 0.2041 with *a p-value* of 0.0069. Because the *p-value* is less than 0.05, the null hypothesis is rejected.(H_0) rejected. These results indicate that the Rayleigh distribution does not provide a good fit to the wind speed data based on the KS test. Meanwhile, the Anderson-Darling (AD) test results show a test statistic value of 0.5278 with *a p-value* of 0.1718. Because the *p-value* is greater than 0.05, the null hypothesis (H_0)is not rejected. Thus, based on the AD test, the distribution can still be accepted as an appropriate model to describe wind speed data in Jayapura City.

The difference in results between the KS and AD tests can occur because the two tests have different characteristics. The KS test measures the maximum difference between the empirical and theoretical distributions, making it more sensitive to differences in the middle of the distribution. Conversely, the AD test places greater weight on the tails of the distribution, making it more sensitive in detecting deviations in those tails. Therefore, although the KS test rejects the fit of the Rayleigh distribution, the AD test results indicate that the distribution is still capable of representing the characteristics of the wind speed data quite well.

3.4 Discussion

3.4.1 Characteristics of Wind Speed Data

The descriptive statistics indicate that the wind speed data exhibit a relatively balanced distribution, as evidenced by the close agreement between the mean and median values. In addition, the relatively small negative skewness suggests that the distribution only slightly deviates from symmetry. These characteristics indicate that the dataset possesses sufficient stability for statistical modeling using continuous probability distributions. Since wind speed is a positive continuous variable, the Rayleigh distribution is considered one of the appropriate probabilistic models for representing its behavior, consistent with the characteristics of continuous distributions described by Johnson and Wichern (2019).

3.4.2 Rayleigh Parameter Estimation Using MLE

The estimated Rayleigh scale parameter obtained through the Maximum Likelihood Estimation method was 5.3702, indicating the dispersion level of the observed wind speed data. The MLE method was selected because it provides parameter estimates with desirable statistical properties, including consistency, efficiency, and asymptotic normality (Casella & Berger, 2002; Pawitan, 2001). These properties enable MLE to produce reliable parameter estimates for probability distribution models. Consequently, the estimated scale parameter obtained in this study provides the basis for constructing the Rayleigh distribution model and conducting subsequent goodness-of-fit analysis.

3.4.3 Evaluation of Distribution Suitability

The goodness-of-fit analysis produced different conclusions between the KS and AD tests. The KS test rejected the Rayleigh distribution because the obtained p-value (0.0069) was below the significance level, whereas the AD test accepted the distribution with a p-value (0.1718). This discrepancy is attributable to the different sensitivities of the two statistical tests. The KS test measures the maximum difference between the empirical and theoretical cumulative distribution functions and is generally more sensitive to deviations occurring around the center of the distribution. In contrast, the AD test places greater emphasis on observations in the tails of the distribution, making it more sensitive to deviations in extreme values. Therefore, the different statistical decisions observed in this study reflect the distinct evaluation criteria employed by the two goodness-of-fit tests rather than inconsistencies in the estimated distribution itself. These findings suggest that assessing the suitability of a probability distribution should not rely on a single statistical test but should consider multiple goodness-of-fit procedures to obtain a more comprehensive evaluation of model performance.

4. CONCLUSION

Based on the results of research on the estimation of Rayleigh distribution parameters using the Maximum Likelihood Estimation (MLE) method on wind speed data in Jayapura City, it was found that the wind speed data had relatively balanced characteristics with an average value of 7.21 km/h, a standard deviation of 2.39 km/h, and a skewness value of -0.33389 which indicated a negative skew but still tended to be close to symmetry. The estimation results using the MLE method produced a Rayleigh distribution scale parameter value of 5.3702, which was then used in the formation of the Rayleigh distribution model.

Based on the Goodness-of-Fit Test, different results were obtained between the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests. The KS test produced *a p-value* of 0.0069, indicating that the Rayleigh distribution does not fit the wind speed data. Meanwhile, the AD test produced *a p-value* of 0.1718, indicating that the Rayleigh distribution is still acceptable for modeling wind speed data in Jayapura City. These differences in results indicate that the suitability of the Rayleigh distribution depends on the characteristics of the test used, so model evaluation should consider more than one fit test model.

BIBLIOGRAPHY

Casella, G., & Berger, R.L. (2002). Statistical Inference (2 nd ed.). Duxbury Press.

Hapsah, S., Somayasa, W., & Mukhtar, N. (2025). Rayleigh Distribution Parameter Estimation on Type II Censored Samples. *Bakti Cendekia: National Journal of Multidisciplinary Research Results* , 2(1), 145-166.

Johnson, R. A., & Wichern, D. W. (2019). Applied Multivariate Statistical Analysis (7th ^{ed} .). Pearson.

Langowuyo, A., Yokhu, S., & Reba, F. (2025). Parameter Estimation and Analysis of Average Years of Schooling in Merauke District with Birnbaum-Saunders Distribution Approach. *KUBIK: Jurnal Publikasi Ilmiah Matematika*, 10(1), 48-55.

Nocedal, J., & Wright, S. J. (2006). Numerical Optimization (^{2nd} ed.). Springer.

Pawitan, Y. (2001). In All Likelihood: Statistical Modeling and Inference Using Likelihood. Oxford University Press.

Simbolon, RU (2021). *Estimation of Rayleigh Distribution Parameters Using the Maximum Likelihood Method and the Bayes Method on Type II Sensor Data* (Bachelor's Thesis, University of Lampung).

Sroyer, A. M., Tandiangnga, T., Reba, F., & Beno, I. S. (2018). Comparison of Parameter Estimation Methods to Determine the Frequency Data Magnitude of Aftershock in Nabire, Papua. *Journal of Science & Science Education*, 2(2), 17-21.

Wilks, D. S. (2019). Statistical Methods in the Atmospheric Sciences (4 th ed.). Academic Press.