CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 10 (2) (2025), Pages 929-940
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Time-varying Distribution Analysis for Rainfall and Temperature Data in Jakarta in Response to Future Climate Change

Suci Nur Setyawati, Sri Nurdiati*, I Wayan Mangku, and Mohamad Khoirun Najib

Department of Mathematics, IPB University, Bogor, Indonesia

Abstract

Jakarta is one of the major cities in Indonesia that is vulnerable to the impacts of climate change, particularly in terms of rainfall and air temperature variability. This study aims to model time-varying statistical distributions of rainfall and air temperature using the maximum likelihood estimation (MLE) method and the fminsearch algorithm. Three types of distributions are evaluated: normal, generalized extreme value (GEV), and lognormal distributions. The MLE method is applied to stationary models, while fminsearch is used for both stationary and nonstationary models. Model performance is assessed based on the Akaike Information Criterion (AIC). The results show that the nonstationary lognormal distribution is the most suitable for rainfall data, while the nonstationary GEV distribution is the most appropriate for air temperature data. These findings indicate an increasing variability in rainfall and a rising average air temperature over time. The results are expected to contribute to the development of climate change mitigation and adaptation strategies in Jakarta.

Keywords: air temperature, fminsearch algorithm, maximum likelihood estimation, parameter estimation, rainfall.

1 Introduction

Climate change has become a global issue with significant impacts on social, economic, and environmental systems [1]. Indonesia, as a tropical archipelago, experiences unique weather and climate characteristics that are not yet fully understood [2]. The two primary variables that define climate are precipitation and air temperature. Indonesia is a country that is highly vulnerable to the impacts of climate change [3]. [4] stated that climate change could cause increasingly severe implications due to increased Earth's temperature by 1.1°C. In the next two decades, the rise in Earth's temperature is expected to reach 1.5°C. The effect of global warming due to climate change, including rainfall and temperature, can cause hydrometeorological hazards, such as increasing the frequency of extreme events and their severity over a long period [5].

Rainfall and temperature patterns can be analyzed to predict climate change in a region. Climate change can often create extreme events that can impact life. Proper risk assessment of extreme events due to climate change can help develop effective mitigation strategies. One commonly used method is hydrological frequency analysis based on stationary assumptions, where

***Corresponding author**. E-mail: nurdiati@apps.ipb.ac.id

Submitted: April 28, 2025 Reviewed: June 26, 2025 Accepted: September 26, 2025
DOI: https://doi.org/10.18860/cauchy.v10i2.32780

conditions are considered unchanging over time. However, this approach is being questioned with the actual impacts of climate change, and nonstationary distribution models are now considered more relevant to anticipate dynamic climate change patterns [6].

Internationally, nonstationary hydrological frequency analysis has become a significant concern as public awareness of climate change impacts increases [7] [8] [9]. In recent years, time-based approaches have often been used to accommodate changing probability parameters to describe patterns of extreme hydrometeorological variables in univariate analysis [10] [11] [12] [13] emphasized the importance of nonstationary hydrological hazard analysis to address climate change and population growth, highlighting the need for adaptive risk management strategies. Furthermore, research by [14] in the Heihe River Basin provides insights into the influence of extreme rainfall on uncertainty in water resources management, which is increasingly relevant under nonstationary conditions due to climate change. In addition, [15] researched how temperature and precipitation fluctuations in China impact water availability. This research focused on the importance of low-flow frequency analysis in a nonstationary context.

In Indonesia, research on using time-varying hydrological models still needs to be conducted. The research that has been conducted includes the study by [16], which discusses the analysis of extreme rainfall using the Extreme Value Theory (EVT) with stationary and nonstationary approaches in Surabaya and Mojokerto. The results show that the nonstationary method provides a more accurate return level estimate for data with trend patterns than the stationary method. The most studies in Indonesia focus on estimating stationary distribution parameters, i.e., the data pattern is considered unchanging over time. For example, research conducted by [17] discussed a copula-based joint distribution model to analyze the relationship between climate conditions and forest fire hotspots in Kalimantan, using rainfall, drought, and ENSO conditions. In addition, similar studies discuss using stationary distributions, namely research conducted by [18], which examines the joint distribution analysis between rainfall and burned area in Southern Sumatra by considering global climate phenomena and applies a stationary distribution approach in his research.

Before developing a time-varying multivariate analysis, an essential first step is to conduct a univariate analysis of time-varying climate data. Jakarta, as Indonesia's capital and most densely populated urban area, is highly vulnerable to climate-related hazards such as flooding, extreme rainfall, and rising temperatures. These conditions make Jakarta a strategic case study for understanding localized climate variability and supporting adaptive urban planning.

Therefore, the author is interested in discussing Jakarta's distribution analysis of time-varying climate change patterns in Jakarta using rainfall and air temperature data as the main variables. The fminsearch algorithm is used to optimize the distribution parameters of climate change patterns. Parameter estimation of stationary and nonstationary distributions is performed with this method on several types of distributions, including the normal distribution, the generalized extreme value (GEV) distribution which models extreme or maximum values, and the lognormal distribution which is used for positively skewed data. The parameter estimation results of stationary distributions are compared with other methods, such as maximum likelihood estimation (MLE). Thus, this study aims to estimate the time-varying distribution parameters of the normal, generalized extreme value (GEV), and lognormal distributions using the maximum likelihood estimation method and the fminsearch algorithm on rainfall and air temperature data in Jakarta, as well as visualize and analyze the best distribution obtained.

This research is expected to provide a more in-depth picture of climate change patterns in Jakarta as well as the contribution of the nonstationary approach in anticipating and predicting the impact of climate change in the future. This article is organized systematically to facilitate readers' understanding of the research process. The second section provides an explanation of the data used and the methods applied, including the selection of distribution types, parameter estimation techniques, and optimization algorithms. The third section presents the results of the analysis of stationary and non-stationary distributions of rainfall and air temperature data in

Jakarta, accompanied by a discussion linking the findings to previous studies and the context of climate change. Finally, the fourth section contains the conclusions of the research results along with recommendations that can be used as a reference for future research.

2 Methods

This section presents the methodology of the study, starting from data description and univariate distribution selection, followed by parameter estimation using the maximum likelihood estimation (MLE) and the fminsearch algorithm. Model performance is then evaluated using the Akaike Information Criterion (AIC), and the best-fitting distributions are visualized to interpret climate variability patterns in Jakarta.

2.1 Data

The author used annual maximum data from monthly rainfall (mm) and air temperature (°C) in Jakarta. The data period covers the years 1864 to 2020. Data from 1864 to 1990 were obtained from KMNI. Meanwhile, data from 1991 to 2020 were sourced from the Meteorology, Climatology, and Geophysics Agency (BMKG). These data were used to support analysis and parameter distribution estimation.

2.2 Univariate Distribution

Distributions are used to describe how values in the rainfall and air temperature datasets are distributed in a specific pattern. The distributions used in this study include the normal distribution, the Generalized Extreme Value (GEV) distribution, and the lognormal distribution. The normal distribution is a continuous variable distribution, so its probability is calculated by determining the area under the curve. the probability density function of the normal distribution is as follows [19]

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty. \tag{1}$$

The generalized extreme value (GEV) is a distribution developed to study extreme events. The extreme data used is within the maximum range of a certain period, such as on a daily, monthly, and annual scale. The probability density function of the GEV distribution is as follows [19]

$$f(x|\mu,\sigma,\xi) = \begin{cases} \frac{1}{\sigma}\exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right)\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1-\frac{1}{\xi}}, & \xi \neq 0 \\ \frac{1}{\sigma}\exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right)\exp\left(-\frac{x-\mu}{\sigma}\right), & \xi = 0. \end{cases} \tag{2}$$

The GEV distribution belongs to the continuous random distribution and is suitable for modeling extreme data.

The lognormal distribution is used to describe data with a favorable and unsymmetrical distribution when the logarithm of the data has a normal distribution. Thus, the lognormal distribution is often used to describe data with exponential properties on a logarithmic scale. The probability density function of the lognormal distribution is as follows

$$f(x|\mu,\sigma) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log(x)-\mu)^2}{2\sigma^2}\right), \quad x,\mu,\sigma > 0. \tag{3}$$

The distributions whose parameters are estimated consist of stationary and nonstationary distributions. A stationary distribution is a distribution that does not depend on the time variable. A nonstationary or time-varying distribution is the probability distribution of a random variable in a time series or changes over time. All parameters in the distribution will be given a time-dependent trend. There are two kinds of location parameter functions plus stationary situations that can be presented as follows

$$\mu_t = \begin{cases} \text{constant}, & \text{if stationary} \\ \mu_0 + \mu_1 t, & \text{if nonstationary} \end{cases} \tag{4}$$

where t denotes time; μ_0 and μ_1 are the new parameters for the time-varying distribution.

2.3 Parameter Estimation

Parameter estimation is used to predict the parameter values of rainfall and air temperature data. The parameter estimation is conducted through maximum likelihood estimation (MLE) and the fminsearch algorithm.

2.3.1 Maximum Likelihood Estimation (MLE)

The parameter values obtained through maximum likelihood estimation (MLE) are stationary distribution estimates that will be compared accurately with the parameters generated by the fminsearch algorithm. According to [19], MLE estimates parameters by maximizing the likelihood function of the probability function of each distribution.

1. Take a random sample size $n; x_1, x_2, ..., x_n$
   $x_i \sim f(x_i, \theta), \quad i = 1, 2, ..., n.$
2. Formulate the likelihood density function of each distribution as the multiplication of the likelihood function of each distribution's likelihood density function as the likelihood function's multiplication.
$$L(\theta|x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i, \theta). \tag{5}$$
3. Maximize the likelihood function. To maximize the likelihood function, derive the likelihood function equation concerning the parameter θ, then solve the first derivative to zero to get the estimate of the parameter θ.
$$\frac{dL(\theta|x_1, x_2, ..., x_n)}{d\theta} = 0 \to \hat{\theta}. \tag{6}$$
   If the likelihood function is difficult to derive, then perform a log transformation on the likelihood function first
$$\log(L(\theta|x_1, x_2, ..., x_n)). \tag{7}$$
4. Test the maximum of the likelihood function using the second derivative test. $\hat{\theta}$ maximizes the likelihood function if its second derivative is less than zero.
$$\frac{d^2L(\theta|x_1, x_2, ..., x_n)}{d\theta^2} < 0. \tag{8}$$
   This step verifies that $\hat{\theta}$ is a local maximum and represents the maximum likelihood estimate.

2.3.2 Fminsearch Algorithm

Furthermore, the fminsearch algorithm estimates the parameters of stationary and nonstationary distributions. This algorithm is a specialized version of the Nelder-Mead Simplex method implemented in Matlab software. [20] stated that this method is a numerical method used to find the minimum value of a function with many variables, especially when the function's derivative is difficult to find. The fminsearch algorithm iteratively modifies the simplex according to specific procedures.

1. Let z_i denote the list of points in the current simplex, $i = 1, 2, \ldots, n, n+1$.
2. Sort the points in the simplex from the lowest function value $(f(z_1))$ to the highest function $f(z_{n+1})$. At each step in the iteration, the algorithm discards the current worst point z_{n+1} and accepts another point into the simplex.
3. Generates bounce points (z_r), which are the values of the side opposite the corner point with the worst value, $z_r = 2m - z_{n+1}$, with $m = \frac{1}{n}\sum_{i=1}^{n} z_i$, and calculate the value of $f(z_r)$.
4. Comparing the $f(z_r)$ value with $f(z_1)$, $f(z_n)$, and $f(z_{n+1})$.
   a) If $f(z_1) \leq f(z_r) < f(z_n)$, accept z_r as the new simplex and end this iteration. *Reflect.*
   b) If $f(z_r) < f(z_1)$, calculate the expansion point z_s $(z_s = m + 2(m - z_{n+1})$, with $m = \frac{1}{n}\sum_{i=1}^{n} z_i)$, and calculate the value of $f(z_s)$. If $f(z_s) < f(z_r)$, accept z_s as the new simplex and end this iteration. *Expand.* Otherwise, accept z_r as the new simplex and end this iteration. *Reflect.*
   c) If $f(z_r) \geq f(z_n)$, a contract between m and z_{n+1} or z_r, depending on which one has the lower objective function value. If $f(z_r) < f(z_{n+1})$, calculate the value of z_c, $z_c = m + \frac{z_r - m}{2}$, and calculate $f(z_c)$. If $f(z_c) < f(z_r)$, accept z_c as the new simplex and end the iteration. *Contract outside.* If $f(z_r) \geq f(z_{n+1})$, calculate the value of z_{cc}, $z_{cc} = m + \frac{z_{n+1} - m}{2}$, and calculate $f(z_{cc})$. If $f(z_{cc}) < f(z_{n+1})$, accept z_{cc} as the new simplex and end the iteration. *Contract inside.* Otherwise, calculate the new points $v_i = z_1 + \frac{z_i - z_1}{2}$ and calculate the value of $f(v_i)$, $i = 2, \ldots, n+1$. In this case, the next iteration of the simplex is $z_1, v_2, ..., v_{n+1}$. Shrink.
5. Checking the domain convergence of the iteration results.
   a) If the distance between all pairs of simplex corner points is less than 0.001, then the simplex is small enough, and it can be said that the simplex is the optimum result of the fminsearch algorithm.
   b) If there is a distance between pairs of simplex corner points of more than 0.001, then the result is not yet optimum, and the next iteration needs to be done. The next iteration means taking the values of $f(z_1), f(z_{n-1})$, and $f(z_n)$ that are new according to the results of the previous iteration.

2.3.3 Comparison of Parameter Results and Visualization of the Best Distribution

Akaike Information Criterion (AIC) is used to find the best distribution of parameter estimation results using the fminsearch algorithm. According to [21], the Akaike Information Criterion (AIC) method is used to select the best model based on the smallest AIC value, and the value is calculated using a specific formula.

$$AIC = -2\ln(L) + 2k \tag{9}$$

where
k = number of parameters estimated in the model;
L = likelihood of the model (the most significant likelihood of the model against the data).

The best distribution obtained with the lowest AIC value from rainfall and air temperature data is visualized using probability density function (pdf) plots of the distribution, along with quartile and mean graphs of the distribution. The visualization results are analyzed and interpreted in relation to the rainfall and air temperature data.

3 Results and Discussion

This section presents the analysis results and discussion of rainfall and air temperature patterns, including descriptive statistics, parameter estimation for stationary and nonstationary distributions, and model evaluation using AIC. The findings help explain climate variability in Jakarta and assess the suitability of various distribution models over time.

3.1 Descriptive Statistics

Descriptive statistics are statistics that analyze data by describing or describing the data that has been collected without intending to make general conclusions or generalizations. The results of descriptive statistical analysis of rainfall and air temperature data in Jakarta in 1864-2020 can be seen in Table 1.

Table 1: Descriptive statistics

Variable	Minimum	Maximum	Average	Standard Deviation
Rainfall (mm)	173.0000	1068.9000	443.6062	166.5255
Air Temperature (°C)	26.1600	29.9000	27.6955	0.9289

Based on Table 1, the descriptive statistical analysis reveals patterns that reflect the climatic dynamics in Jakarta over the period 1864–2020. Rainfall exhibits significant variation, ranging from 173 mm to 1068.9 mm. This wide range of 895.9 mm indicates substantial annual rainfall fluctuations, which may be influenced by weather phenomena or long-term climate changes. The average rainfall of 443.61 mm with a standard deviation of 166.53 mm suggests uncertainty in the annual rainfall pattern. On the other hand, air temperature demonstrates relatively high stability, with much smaller variations compared to rainfall. The temperature range from the lowest to the highest is only 3.8°C, spanning from 26.1°C to 29.9°C, with an average of 27.70°C and a standard deviation of 0.93°C. This stability indicates that, while rainfall varies significantly, air temperature tends to remain relatively consistent on an annual basis. This initial analysis provides a comprehensive overview of the fundamental climatic patterns in Jakarta.

3.2 Estimation of Stationary Distribution Parameters

Stationary distribution estimation is performed using maximum likelihood estimation (MLE) and the fminsearch algorithm. The main point of estimating distribution parameters using maximum likelihood estimation (MLE) is to maximize the likelihood function, which is the joint probability function of each distribution of rainfall and air temperature data. The likelihood function represents the probability of observing the entire dataset given a set of parameters. It is constructed as the product of the probability density function of each individual observation, which is formulated as $\mathcal{L}(\theta|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i; \theta)$. The results of estimating the stationary distribution parameters using MLE can be seen in Table 2.

Table 2: Parameter estimation results of stationary distribution using MLE

Distribution	Rainfall (mm)			Air temperature (°C)		
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
Normal	443.6030	166.5250	-	27.6994	0.9279	-
GEV	366.2170	128.9680	0.0917	27.3010	0.8012	-0.1002
Lognormal	6.0274	0.3697	-	3.3209	0.0333	-

Table 2 shows rainfall has a more significant variation than air temperature in Jakarta. In the normal distribution, rainfall has a mean of 443.6030 mm with a standard deviation of 166.5250 mm, while air temperature is more stable with a mean of 27.6994°C and a standard deviation of 0.9279°C. In the GEV distribution, rainfall shows an average of 366.2170 mm with almost symmetrical distribution tails, while air temperature tends to decrease in extremes with a shape parameter of -0.1002. The lognormal distribution shows that rainfall is more variable than air temperature, which remains narrower and more consistent around its mean.

Estimating stationary distribution parameters using the fminsearch algorithm is done by finding the optimum value of parameters by minimizing the log-likelihood function of each distribution on rainfall and air temperature data. The results of stationary distribution parameter estimation using the fminsearch algorithm can be seen in Table 3.

Table 3: Stationary distribution parameter estimation results using the fminsearch algorithm

Distribution	Rainfall (mm)			Air temperature (°C)		
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
Normal	443.6026	165.9874	-	27.9664	0.9249	-
GEV	366.2174	128.9685	0.0191	27.3009	0.8012	-0.1002
Lognormal	6.0274	0.3685	-	3.3209	0.0332	-

Based on Table 3, parameter estimation using the fminsearch algorithm has results similar to parameter estimation using MLE. Thus, the results of stationary distribution parameter estimation using the fminsearch algorithm in Table 3 are compared with the results of stationary distribution parameter estimation using the MLE method in Table 2 to see the model fit results of the two methods. The results of the comparison of the stationary distribution parameter values of each distribution with MLE and the fminsearch algorithm are shown in Table 4.

Table 4: Results of comparison of distribution parameter values

Distribution	Difference in parameter values					
	Rainfall (mm)			Air Temperature (°C)		
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
Normal	0.0004	0.5376	-	0.2670	0.0003	-
GEV	0.0004	0.0005	0.0000	0.0001	0.0000	0.0000
Lognormal	0.0004	0.0012	-	0.0000	0.0001	-

Based on Table 4, the lowest difference in parameter results on rainfall data is 0.0000 mm, and the highest difference is 0.5376 mm. The lowest difference in parameter results in rainfall data is 0.0000°C, and the highest difference is 0.2670°C. These small differences support the convergence properties of the Nelder-Mead algorithm. As implemented in the fminsearch function, it has been shown to converge reliably to a local minimum in low-dimensional smooth optimization problems. This indicates that the fminsearch algorithm has effectively reached parameter estimates close to the maximum likelihood estimates.

3.3 Estimation of Nonstationary Distribution Parameters

The fminsearch algorithm is used to estimate the parameters of non-stationary distributions of normal, GEV, and lognormal distributions by gradually involving time variables in the estimation process. This approach aims to understand how time variables affect each distribution parameter. In the first step, time variables are only included in the location parameters, while the scale and shape parameters remain constant (unaffected by time). This step helps identify if there are any trends in the data that can be explained by changes in the location of the distribution over time. In the next step, the time variable is also incorporated into the scale and the location parameter, while the shape parameter remains unchanged. This stepwise approach allows for a more in-depth look at how the time variable affects the variability of the data (through the scale parameter) in addition to shifting the center of the distribution (through the location parameter). By performing this stepwise estimation, the impact of adding time variables on the distribution parameters can be seen more clearly. The general form of the time-varying location and scale parameters is expressed as linear functions of time: $\mu_t = \mu_0 t + \mu_1$, $\sigma_t = \sigma_0 t + \sigma_1$, where t represents time, and $\mu_0, \mu_1, \sigma_0, \sigma_1$ are parameters to be estimated. The nonstationary parameter estimation results for the rainfall (mm) and air temperature (°C) data are presented in Table 5.

Table 5: Estimated values of distribution parameters of rainfall data (mm) and air temperature (°C)

Data	Distribution	Parameter values of the `fminsearch` algorithm		
		$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
Rainfall (mm)	Normal	$0.6800t + 390.5624$	163.1750	-
		$0.5567t + 399.9393$	$0.3156t + 137.6182$	-
	GEV	$0.4572t + 331.8594$	127.9178	0.0129
		$0.4893t + 329.7612$	$0.1569t + 115.8374$	0.0067
	Lognormal	$0.0014t + 5.9183$	0.3631	-
		$0.0014t + 5.9196$	$0.0001t + 0.3596$	-
Air temperature (°C)	Normal	$0.0187t + 26.2382$	0.3910	-
		$0.0186t + 26.2473$	$0.0002t + 0.3787$	-
	GEV	$0.0184t + 26.0970$	0.3459	-0.1182
		$0.0181t + 26.0987$	$0.0005t + 0.2958$	0.0032
	Lognormal	$0.0007t + 3.2683$	0.0139	-
		$0.0007t + 3.2683$	$0.0000t + 0.0139$	-

The stationary distribution parameter estimation results obtained using the fminsearch algorithm in Table 5 will be compared with the stationary distribution parameter estimation results also calculated using the same algorithm in Table 3. This comparison assesses and determines the best distribution model between the two results. Once the best model is determined, the results will be visualized graphically to represent how the best distribution model reflects the data pattern.

3.4 Comparison of Parameter Results

The estimated distribution parameters calculated using the fminsearch algorithm are evaluated using the AIC to determine the best model or distribution parameters. The main principle in this evaluation is that the model with the lowest AIC value is considered the best, as it indicates that the model has an optimal balance between complexity and fit to the data. The results of the AIC values of the normal distribution, GEV distribution, and lognormal distribution with stationary and nonstationary distribution parameters estimated by the fminsearch algorithm on rainfall (mm) and air temperature (°C) data are shown in Table 6.

Table 6: AIC value on rainfall (mm) and air temperature (°C) data

Parameter type	Distribution	Number of parameters	Rainfall (mm)	Air temperature (°C)
Stationary	Normal	2	2028.6000	419.6600
	GEV	3	2006.2000	411.3600
	Lognormal	2	2002.8800	417.5000
Nonstationary	Normal	3	2025.2000	154.7600
	Normal	4	2024.2000	156.6540
	GEV	4	2004.7400	148.1880*
	GEV	5	2005.8800	152.1760
	Lognormal	3	2000.3400*	149.9900
	Lognormal	4	2002.3200	151.9900

*Lowest AIC value

Based on Table 6, the lognormal distribution model with three parameters has the smallest AIC value (2000.3400) for rainfall data. The GEV distribution model with four parameters has the smallest AIC value (148.1880) for air temperature data. These results show that the 3-parameter lognormal model is most suitable for rainfall data, and the 4-parameter GEV model is most appropriate for air temperature data in Jakarta in the 1864-2020 period compared to other models.

3.5 Visualization of the Best Distribution

The best distribution model based on the AIC value is visualized through the probability density function (pdf) graph for rainfall and air temperature data. The pdf graph for rainfall is depicted with a 3-parameter lognormal model, while for air temperature, it is depicted with a 4-parameter GEV model. These graphs illustrate the distribution of historical rainfall and air temperature data in Jakarta in various years, as shown in Figure 1.

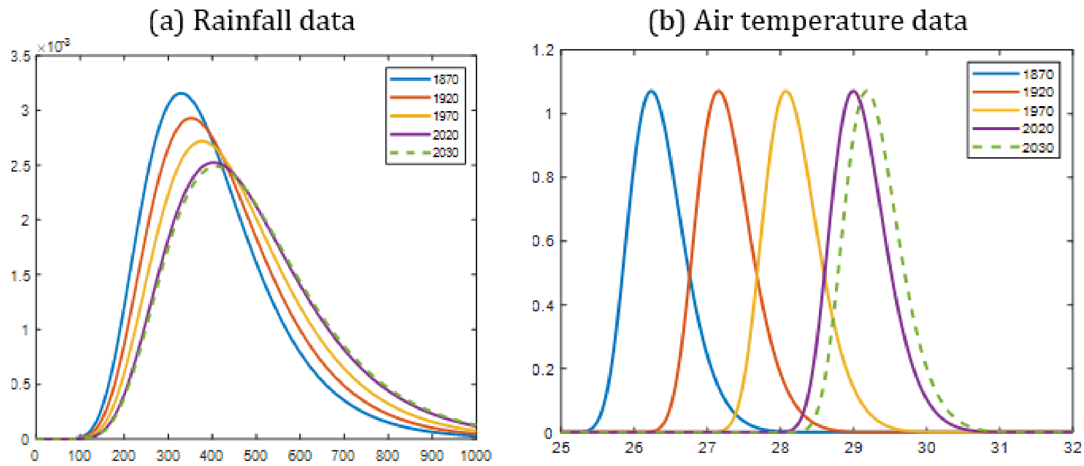

Figure 1: Pdf distribution graph of (a) rainfall data and (b) air temperature data

Based on Figure 1, the pdf curves show that the lognormal and GEV models can model rainfall variations from historical years (1870, 1920, 1970) to future predictions (2030). The rainfall curve shifts to the right, indicating increased rainfall over time. The peak of the rainfall distribution in 1870 was about 400 mm per year, but the peak became flatter and shifted to the right in subsequent years. This indicates that extreme rainfall events are becoming more frequent. The distribution curve on the air temperature graph shows a consistent rightward shift, with the shape of the graph remaining the same from year to year. This indicates that the variation is constant, with the annual average temperature continuing to increase.

In addition to the pdf graph, the best distribution model based on the AIC value is also visualized through quartile and average graphs for rainfall and air temperature data. This graph displays three trend lines: the average, Q1 and Q3 (lower quartile 25% and upper quartile 25%), and Q0 and Q4 (lower limit 9% and upper limit 91% of the data). The quartile and mean graphs for rainfall use a 3-parameter lognormal model, while that for air temperature uses a 4-parameter GEV model. This visualization provides a more comprehensive picture of the distribution of rainfall data in Jakarta from 1864–2030, as shown in Figure 2.

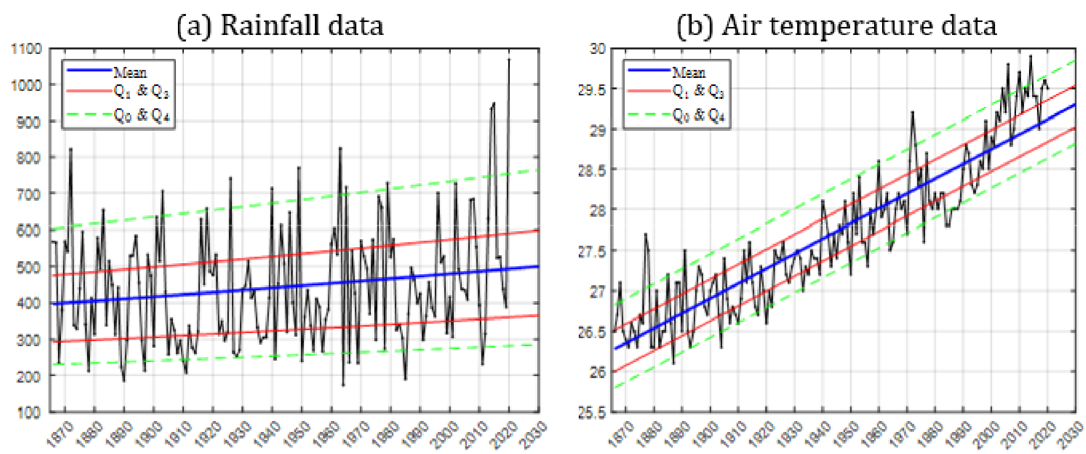

Figure 2: Quartile and mean distribution graphs of (a) rainfall data and (b) air temperature data

Based on Figure 2, the graph shows three trend lines: blue for the average, red for Q1 and Q3, and green for Q0 and Q4. The rainfall graph shows a slight increase in the average trend, with the distance between the red and green lines getting wider. This indicates an increase in rainfall variability over time. The increasing Q4 line indicates an increase in extreme rainfall, while the Q0 line remains stable, indicating low rain is still occurring. The average trend in the air temperature graph is increasing, with an average temperature in 2020 of more than 29°C. It can be predicted that the average air temperature in the next ten years will be greater than the average in 2020. The distance between the red and green lines is constant, indicating stable air temperature variability. Some outlier points on both graphs indicate extreme rainfall and temperature events in some years.

Thus, the distribution estimation results using the fminsearch algorithm are suitable for describing the distribution of rainfall and air temperature data in Jakarta from 1864 to 2020 and can be used to predict future rainfall and air temperature in Jakarta. Rainfall and air temperature in Jakarta in the future, namely until 2030, are expected to experience increased variations in precipitation and average air temperature. By conducting appropriate adaptation and mitigation planning in Jakarta, the impacts of extreme climate change from rainfall and air temperature that occur in the future can be minimized.

These findings align with previous studies such as [13] and [15], which emphasize the importance of nonstationary modeling for climate extremes. The ability of the lognormal and GEV distributions to capture the increasing variability and extremity of rainfall and temperature, respectively, suggests the necessity of adaptive planning. These results can support the formulation of climate adaptation policies and disaster risk management strategies in Jakarta.

4 Conclusion

The parameter estimation conducted using the fminsearch algorithm yields results for the stationary distribution parameters that closely align with those derived from the maximum likelihood estimation (MLE) method. The minimal deviation between the two methods is lower than 10^{-5}, while the most significant observed difference reaches 5.3760. These slight variations indicate that both fminsearch and MLE approaches offer similar levels of accuracy in estimating stationary distribution parameters.

The estimates generated using the fminsearch algorithm were leveraged for the nonstationary distribution parameters to identify the best-fitting distribution based on the Akaike Information Criterion (AIC). The selection criterion, which favors models with the lowest AIC, revealed that the optimal distribution for modeling rainfall data is the 3-parameter lognormal distribution, with an AIC value of 2000.3400. For air temperature data, the best distribution is a 4-parameter model, resulting in an AIC value of 148.1880.

These results highlight the effectiveness of the fminsearch algorithm in accurately estimating time-varying distribution parameters and identifying optimal models for climate data. These results may support the development of adaptive strategies for climate risk management in Jakarta. Future research may explore multivariate time-varying modeling or the inclusion of large-scale climate indices to enhance predictive capability.

CRediT Authorship Contribution Statement

Suci Nur Setyawati: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualization, Project administration. **Sri Nurdiati:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing – original draft, Writing – review and editing, Visualization, Supervision, Project administration, Funding acquisition. **I Wayan Mangku:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing –

original draft, Supervision, Project administration. **Mohamad Khoirun Najib:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Project administration.

Declaration of Generative AI and AI-assisted technologies

In the development of this work, Generative AI tools—namely ChatGPT—were utilized exclusively to improve language clarity and correct grammar. The analysis, interpretation, and main content were entirely created by the authors without AI involvement. All key findings and discussions were produced manually by the research team.

Declaration of Competing Interest

The authors affirm that we have no recognized financial conflicts of interest or personal affiliations that might have affected the objectivity of the work presented in this paper.

Funding and Acknowledgments

We gratefully acknowledge the BIMA Kemendikbudristek program for its financial support of this research through the "Penelitian Tesis Magister (PTM)" scheme, under contract number 006/C3/DT.05.00/PL/2025 and subcontract number 23342/IT3.D10/PT.01.03/P/B/2025. We also thank the Department of Mathematics, IPB University, for its valuable support throughout the course of this study.

Data Availability

The rainfall and air temperature data for Jakarta utilized in this study are publicly accessible from two sources: the KMNI website [https://dataplatform.knmi.nl/] for the period 1864–1990, and the BMKG website [https://dataonline.bmkg.go.id/dataonline-home] for the period 1991–2020.

References

[1] T. A. Carleton and S. M. Hsiang, "Social and economic impacts of climate," *Science (80-.).*, vol. 353, no. 6304, 2016. DOI: 10.1126/science.aad9837.

[2] M. Ariska, H. Akhsan, M. Muslim, M. Romadoni, and F. S. Putriyani, "Prediksi perubahan iklim ekstrem di kota palembang dan kaitannya dengan fenomena el nino–southern oscillation (enso) berbasis machine learning," *JIPFRI (Jurnal Inovasi Pendidikan Fisika dan Riset Ilmiah)*, vol. 6, no. 2, pp. 79–86, 2022. Available online.

[3] R. O. et al, "Climate change driving salinity: An overview of vulnerabilities, adaptations, and challenges for indonesian agriculture," *Weather. Clim. Soc.*, vol. 16, no. 1, pp. 29–49, 2024. DOI: 10.1175/WCAS-D-23-0025.1.

[4] C. Parmesan, Y. T. Morecroft, and H.-O. Pörtner, *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the IPCC Sixth Assessment Report*. Cambridge University Press, 2022. DOI: 10.1017/9781009325844.

[5] H. P. Rahayu, R. Haigh, D. Amaratunga, B. Kombaitan, D. Khoirunnisa, and V. Pradana, "A micro scale study of climate change adaptation and disaster risk reduction in coastal urban strategic planning for jakarta," *International Journal of Disaster Resilience in the Built Environment*, vol. 11, no. 1, pp. 119–133, 2020. DOI: 10.1108/IJDRBE-10-2019-0073.

[6] P. X. et al, "Time-varying copula and average annual reliability-based nonstationary hazard assessment of extreme rainfall events," *J. Hydrol.*, vol. 603, no. PA, p. 126 792, 2021. DOI: 10.1016/j.jhydrol.2021.126792.

[7] H. Chen and J. Sun, "Contribution of human influence to increased daily precipitation extremes over china," *Geophys. Res. Lett.*, vol. 44, no. 5, pp. 2436–2444, 2017. DOI: 10.1002/2016GL072439.

[8] B. C. Call, P. Belmont, J. C. Schmidt, and P. R. Wilcock, "Changes in floodplain inundation under nonstationary hydrology for an adjustable, alluvial river channel," *Water Resources Research*, vol. 53, no. 5, pp. 3811–3834, 2017. DOI: 10.1002/2016WR020277.

[9] M. Ghanbari, M. Arabi, J. Obeysekera, and W. Sweet, "A coherent statistical model for coastal flood frequency analysis under nonstationary sea level conditions," *Earth's Future*, vol. 7, no. 2, pp. 162–177, 2019. DOI: 10.1029/2018EF001089.

[10] Q. Zhang, X. Gu, V. P. Singh, and X. Chen, "Evaluation of ecological instream flow using multiple ecological indicators with consideration of hydrological alterations," *Journal of Hydrology*, vol. 529, pp. 711–722, 2015. DOI: 10.1016/j.jhydrol.2015.08.066.

[11] P. Ganguli and P. Coulibaly, "Does nonstationarity in rainfall require nonstationary intensity-duration-frequency curves?" *Hydrol. Earth Syst. Sci.*, vol. 21, no. 12, pp. 6461–6483, 2017. DOI: 10.5194/hess-21-6461-2017.

[12] V. Agilan and N. V. Umamahesh, "Covariate and parameter uncertainty in non-stationary rainfall idf curve," *Int. J. Climatol.*, vol. 38, no. 1, pp. 365–383, 2018. DOI: 10.1002/joc.5181.

[13] L. Yan, L. Li, P. Yan, H. He, J. Li, and D. Lu, "Nonstationary flood hazard analysis in response to climate change and population growth," *Water (Switzerland)*, vol. 11, no. 9, pp. 1–20, 2019. DOI: 10.3390/w11091811.

[14] Q. Tian, Z. Li, and X. Sun, "Frequency analysis of precipitation extremes under a changing climate: A case study in heihe river basin, china," *J. Water Clim. Chang.*, vol. 12, no. 3, pp. 772–786, 2021. DOI: 10.2166/wcc.2020.170.

[15] M. Yılmaz and F. Tosunoğlu, "Non-stationary low flow frequency analysis under climate change," *Theor. Appl. Climatol.*, vol. 155, no. 8, pp. 7479–7497, 2024. DOI: 10.1007/s00704-024-05081-8.

[16] D. Rahmayani and S. Sutikno, "Analisis curah hujan ekstrem non-stasioner dengan pendekatan block maxima di surabaya dan mojokerto," *J. Sains dan Seni ITS*, vol. 8, no. 2, 2020. DOI: 10.12962/j23373520.v8i2.44133.

[17] M. K. Najib, S. Nurdiati, and A. Sopaheluwakan, "Multivariate fire risk models using copula regression in kalimantan, indonesia," *Nat. Hazards*, vol. 113, no. 2, pp. 1263–1283, 2022. DOI: 10.1007/s11069-022-05346-3.

[18] S. Nurdiati, M. K. Najib, and M. Z. Bayu, "Analisis pembentukan sebaran bivariat berbasis copula antara luas area terbakar dan curah hujan di sumatra bagian selatan," *J. Mat. Integr.*, vol. 18, no. 2, p. 217, 2022. DOI: 10.24198/jmi.v18.n2.42024.217-227.

[19] R. A. Fisher, "On an absolute criterion for fitting frequency curves," *Stat. Sci.*, vol. 12, no. 1, pp. 39–41, 1997.

[20] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder-mead simplex method in low dimensions," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998. DOI: 10.1137/S1052623496303470.

[21] A. A. Grasa, *Econometric Model Selection: A New Approach*. Kluwer Academic Publishers, 1989.