CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (1) (2026), Pages 1155-1170
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Comparison of Extreme Value Logistic and Copula Approaches for Bivariate Extreme Value in Pekanbaru

Fadilla Afsari[1] and A'yunin Sofro[2*]

[1] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Indonesia*
[2] *Department of Actuarial Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Indonesia*

Abstract

Global climate change has intensified extreme weather events in tropical regions, increasing heat-related risks in Pekanbaru City. This study aims to analyze the joint behavior of extreme maximum temperature and humidity using a multivariate extreme value framework. Daily data from 2014–2024 were processed using a seasonal block maxima approach, resulting in 24 seasonal extreme observations. The marginal distributions follow a Weibull-type Generalized Extreme Value (GEV) distribution, indicating bounded extreme values. Dependence analysis indicates a weak association, where the Joe copula converges to the independence case, while the Frank copula yields the lowest AIC value among the fitted copula models. However, the difference in AIC values is small and should be interpreted cautiously. Joint return period analysis indicates that joint exceedance of seasonal block maxima is relatively infrequent (approximately 100 years), whereas single-variable exceedances occur more frequently (approximately 5.2 years). These return period estimates are exploratory and subject to uncertainty due to the limited sample size. Overall, the results suggest a weak dependence structure between the variables, and the findings should be interpreted cautiously in the context of climate-related risk assessment.

Keywords: Extreme value theory; copula; climate risk; temperature; humidity

1. Introduction

Global climate change has increased the frequency and intensity of extreme weather phenomena, particularly in tropical regions such as Indonesia. The Intergovernmental Panel on Climate Change (IPCC) reports that rising extreme temperatures accompanied by high atmospheric humidity increase the risk of extreme events and their impacts on natural systems and human health [1]. Extreme heat events are therefore not isolated phenomena but are influenced by the interaction of multiple climatic variables. Consequently, statistical approaches based solely on average values are insufficient to adequately represent the risks associated with extreme climate conditions.

Pekanbaru City, as a major urban and economic center in Riau Province, is exposed to such extreme conditions. Observations from the Sultan Syarif Kasim II Meteorological Station (BMKG) indicate that daily maximum temperatures frequently exceed 35 °C in recent years

*Corresponding author. E-mail: ayuninsofro@unesa.ac.id

Submitted: February 10, 2026 Reviewed: March 06, 2026 Accepted: May 09, 2026
DOI: https://doi.org/10.18860/cauchy.v11i1.41081

[2]. Regional climate variability is strongly influenced by monsoonal circulation and sea surface temperature dynamics, which shape temperature and rainfall patterns [3]. These conditions may lead to joint occurrence of high temperature and high humidity, increasing the potential for compound extreme events.

Extreme Value Theory (EVT) provides a statistical framework for modeling rare but high-impact events by focusing on the tail behavior of distributions [4, 5]. The Generalized Extreme Value (GEV) distribution has been widely applied in studies of extreme temperature and precipitation under climate variability [6–8]. However, most existing studies in Indonesia still adopt univariate approaches, in which each climatic variable is modeled independently. This limitation may lead to an underestimation of risk, particularly when extremes occur jointly.

To address this limitation, Bivariate Extreme Value Theory (BEVT) and copula-based models allow separate modeling of marginal distributions and dependence structures [9, 10]. Previous studies show that BEVL models and copula approaches can represent dependence structures in climatic variables [11, 12]. Nevertheless, studies that explicitly compare BEVL with multiple copula models in the context of extreme temperature and humidity in tropical regions remain limited.

Therefore, this study aims to model the joint behavior of extreme maximum temperature and humidity in Pekanbaru City using a seasonal block maxima approach. The dependence structure is analyzed using the Bivariate Extreme Value Logistic (BEVL) model and copula-based approaches, including the Joe copula and the Frank copula. The Joe copula is included to examine potential upper-tail dependence, while the Frank copula is considered to represent symmetric dependence without tail dependence. This study contributes by providing a comparative evaluation of these models based on joint return period analysis and the Akaike Information Criterion (AIC), thereby offering a statistical assessment of compound extreme climate behavior and its potential implications for climate-related hazards.

2. Methods

This study employs a quantitative approach based on Bivariate Extreme Value Theory (EVT) and copula models to analyze extreme air temperature and humidity in Pekanbaru City. The dataset consists of daily observations of air temperature and relative humidity recorded at the Sultan Syarif Kasim II Meteorological Station (BMKG) over the period 2014–2024. The analyzed variables are air temperature x and relative humidity y. The EVT framework is adopted due to its effectiveness in modeling rare but high-impact climate extremes [4, 7, 8].

Prior to modeling, data preprocessing was conducted by removing non-physical observations, particularly zero values, to ensure data quality. Exploratory analysis indicates strong seasonal variability consistent with the Asian–Australian monsoon system. Therefore, a seasonal block maxima approach was applied using two periods, ONDJFM and AMJJAS, to obtain relatively homogeneous extreme observations and reduce temporal dependence.

The seasonal maxima of temperature and humidity were modeled using the Generalized Extreme Value (GEV) distribution. Parameter estimation for the GEV model was performed using Maximum Likelihood Estimation (MLE). The adequacy of the fitted marginal models was evaluated using the Kolmogorov–Smirnov (KS) test and quantile–quantile (QQ) plots [13, 14].

To model the dependence structure, the Bivariate Extreme Value Logistic (BEVL) model and copula-based approaches were employed [6, 12, 15, 16]. The Joe copula was used to capture upper-tail dependence, while the Frank copula was considered as an alternative model for symmetric dependence without tail dependence [9, 17, 18]. Parameter estimation for the BEVL model was performed using Maximum Likelihood Estimation (MLE).

For the copula models, initial parameter values were obtained via inversion of Kendall's tau and subsequently refined using MLE to ensure stable estimation. Finally, the joint behavior of extreme temperature and humidity was evaluated using joint return periods under AND and OR scenarios. Model performance was assessed using the Akaike Information Criterion (AIC), which balances goodness of fit and model complexity [19, 20].

2.1. Extreme Value Theory (EVT)

Extreme Value Theory (EVT) is used to analyze extreme events arising from stochastic processes, particularly rare observations with potentially significant impacts, such as extreme temperature, rainfall, floods, and wind events [21]. In contrast to conventional statistical approaches that focus on central tendencies, EVT emphasizes the behavior of observations in the tail of the distribution, where extreme values are located [22]. Therefore, EVT provides an appropriate framework for modeling climate extremes characterized by low occurrence frequency but potentially substantial impacts.

In this study, the Block Maxima approach is employed by partitioning the data into seasonal blocks and extracting the maximum value from each block. This approach is adopted to reflect the seasonal variability of the climate system in Indonesia, particularly in the Sumatra region, which is influenced by the Asian–Australian monsoon system. The resulting seasonal maxima are subsequently used as the basis for modeling the marginal extreme distributions.

The resulting block maxima are then modeled using the Generalized Extreme Value (GEV) distribution, which provides a unified framework for representing different types of extreme value behavior. The Block Maxima approach is selected due to its suitability for regularly recorded time-series data and its direct interpretation in terms of return levels. In addition, this approach facilitates the interpretation of seasonal extreme climate behavior in the study area.

Although the Peaks Over Threshold (POT) approach is also widely used in EVT, it requires the selection of an appropriate threshold, which may introduce additional uncertainty. Therefore, the Block Maxima approach is considered more appropriate for the present study, given the objective of capturing seasonal extreme behavior in temperature and humidity. This consideration is particularly relevant because the study focuses on seasonal climate variability rather than individual daily exceedances.

2.2. Block Maxima

The block maxima method is applied to extract extreme values from a time series by dividing the data into equal time intervals and selecting the maximum value within each block [23, 24]. In this study, seasonal blocks are used, consisting of two main periods, namely ONDJFM (October–March) and AMJJAS (April–September). This seasonal division is intended to represent the dominant climatic variability in the study region.

This seasonal classification is intended to ensure that each block represents relatively homogeneous climatic conditions, which is important for satisfying the assumptions of Extreme Value Theory. It also reflects the influence of monsoonal patterns that govern climate variability in the study region. Therefore, the seasonal block maxima approach is considered appropriate for representing climate extremes in tropical environments.

The resulting series of seasonal block maxima is subsequently modeled using the Generalized Extreme Value (GEV) distribution, which is widely used to describe the probabilistic behavior of extreme observations. This framework enables the estimation of return levels associated with rare events. In addition, the GEV framework provides a unified representation of different forms of extreme value behavior.

While the use of seasonal blocks reduces the effective sample size, it provides a more realistic representation of extreme climate conditions and helps mitigate temporal dependence among observations. This trade-off is commonly accepted in EVT applications involving climatological data. Consequently, the approach allows a more stable interpretation of seasonal extreme climate variability in the study area.

2.3. Generalized Extreme Value Distribution

The Generalized Extreme Value (GEV) distribution is used to model the maximum values of a time series based on the block maxima principle. This distribution arises from the Fisher–Tippett theorem, which states that, under appropriate normalization, the distribution of block maxima converges to a limiting distribution as

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to F(x), \quad n \to \infty, \tag{1}$$

where $F(x)$ belongs to one of three families, namely the Gumbel, Fréchet, or Weibull distributions [25, 26]. These three types are unified within the GEV distribution framework.

The GEV distribution is characterized by three parameters: the location parameter μ, the scale parameter $\sigma > 0$, and the shape parameter ξ. The shape parameter determines the tail behavior of the distribution, where $\xi = 0$ corresponds to the Gumbel type, $\xi > 0$ to the Fréchet type, and $\xi < 0$ to the Weibull type [4]. The cumulative distribution function (CDF) of the GEV distribution is given by

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

In this study, the GEV distribution is fitted to the seasonal block maxima of temperature and humidity. Parameter estimation is performed using the Maximum Likelihood Estimation (MLE) method, which is widely used in extreme value analysis due to its statistical efficiency [4].

To ensure the adequacy of the fitted marginal models, goodness-of-fit diagnostics are conducted using the Kolmogorov–Smirnov (KS) test and quantile–quantile (QQ) plots. These procedures are used to verify that the GEV distribution appropriately represents the observed extreme data before proceeding to dependence modeling.

2.4. Dependence

Dependence is used to describe the interrelationship between two random variables, such as extreme temperature and extreme humidity. Two variables are considered independent if the occurrence of an extreme event in one variable does not affect the probability of an extreme event in the other, while the presence of dependence may influence the likelihood of joint extreme events. Therefore, dependence analysis is important for understanding the joint behavior of climate extremes.

In this study, dependence is analyzed to characterize the joint behavior of extreme temperature and humidity. A commonly used measure of nonlinear dependence is Kendall's tau, which quantifies the strength of monotonic association between variables [9, 10]. Kendall's tau is used to provide an initial assessment of dependence and to obtain starting values for copula parameter estimation.

To capture the dependence structure more flexibly, copula-based models are employed, including the Joe copula and the Frank copula, alongside the Bivariate Extreme Value Logistic (BEVL) model. The Joe copula is particularly useful for modeling upper-tail dependence, while the Frank copula is suitable for representing symmetric dependence without tail dependence. These models are considered appropriate for evaluating different forms of dependence in multivariate extreme events.

The final parameter estimation for the dependence models is carried out using maximum likelihood estimation, ensuring consistency and stability of the fitted models. This estimation procedure is widely applied in dependence modeling due to its statistical efficiency. In addition, maximum likelihood estimation allows the fitted models to be compared consistently within the proposed modeling framework.

2.5. Bivariate Extreme Value Logistic (BEVL)

The Bivariate Extreme Value Logistic (BEVL) model is used to describe the dependence structure between two extreme variables, such as maximum air temperature and maximum relative humidity. Let (X, Y) denote a vector of block maxima, where each marginal variable follows a Generalized Extreme Value (GEV) distribution with parameters (μ_X, σ_X, ξ_X) and (μ_Y, σ_Y, ξ_Y), respectively.

To construct the joint model, the marginal variables are transformed into a standard Fréchet scale. The joint cumulative distribution function (CDF) of the BEVL model is expressed as

$$F_{X,Y}(x, y) = \exp\{-V(\tilde{x}, \tilde{y})\}, \tag{3}$$

where $V(\tilde{x}, \tilde{y})$ is the logistic exponent measure defined as

$$V(\tilde{x}, \tilde{y}) = \left(\tilde{x}^{-1/\alpha} + \tilde{y}^{-1/\alpha}\right)^{\alpha}, \qquad 0 < \alpha \leq 1. \tag{4}$$

The parameter α controls the strength of dependence between the two variables. A value of $\alpha = 1$ corresponds to independence, while smaller values indicate stronger dependence between extremes. In practical applications, values of α close to 1 represent weak or near-independent dependence structures.

In this study, the BEVL model is fitted to the seasonal block maxima of temperature and humidity to characterize their joint extreme behavior. Parameter estimation is carried out using the Maximum Likelihood Estimation (MLE) method based on paired observations (x_i, y_i).

It should be noted that when the estimated value of α is close to unity, the dependence structure approaches independence. Therefore, careful interpretation is required, particularly when the sample size is limited, as weak dependence may reflect both the underlying data structure and estimation uncertainty.

2.6. Joe Copula

Based on Sklar's Theorem, the joint distribution of two continuous random variables can be expressed through a copula function that links the marginal distributions with their dependence structure [9]. In general, the joint cumulative distribution function of random variables X and Y is written as

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)), \tag{5}$$

where $C : [0, 1]^2 \to [0, 1]$ denotes a copula function, and F_X and F_Y represent the marginal distribution functions of X and Y, respectively.

The Joe copula belongs to the Archimedean copula family and is commonly used to model upper-tail dependence, which is relevant for joint extreme events [9, 10]. The Joe copula for the variable pair $(u, v) = (F_X(x), F_Y(y))$ is defined as

$$C_{\text{Joe}}(u, v; \theta) = 1 - \left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right]^{1/\theta}, \quad \theta \geq 1. \tag{6}$$

In this study, the Joe copula is applied to model the dependence between extreme temperature and humidity using pseudo-observations obtained from the fitted GEV marginal distributions. The pseudo-observations are defined as

$$u_i = \widehat{F}_X(x_i), \qquad v_i = \widehat{F}_Y(y_i), \quad i = 1, \ldots, n.$$

Parameter estimation is performed using a two-step procedure. First, an initial estimate is obtained through inversion of Kendall's tau to provide a stable starting value. This is followed by maximum likelihood estimation (MLE) to obtain the final parameter estimate.

The dependence parameter θ controls the strength of upper-tail dependence, where larger values indicate stronger dependence. It should be noted that $\theta = 1$ corresponds to the independence copula. Therefore, when the estimated parameter is close to this boundary, the dependence structure may be weak or effectively absent, and the interpretation should be made with caution.

2.7. Frank Copula

Based on Sklar's Theorem, copula functions provide a flexible framework to model the dependence structure between random variables independently from their marginal distributions [9]. In addition to the Joe copula, this study also employs the Frank copula as an alternative model to capture the dependence between extreme temperature and humidity.

The Frank copula is an Archimedean copula introduced by Frank (1979), which is capable of modeling symmetric dependence structures without tail dependence [9, 17, 18]. The copula function for the variable pair $(u, v) = (F_X(x), F_Y(y))$ is defined as

$$C_{\text{Frank}}(u, v; \theta) = -\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right], \quad \theta \in \mathbb{R} \setminus \{0\}. \tag{7}$$

The parameter θ controls the strength and direction of dependence. Positive values of θ indicate positive dependence, while negative values indicate negative dependence. When $\theta \to 0$, the Frank copula converges to the independence copula, indicating no dependence between variables.

In this study, the Frank copula is applied to pseudo-observations obtained from the fitted GEV marginal distributions, defined as

$$u_i = \widehat{F}_X(x_i), \qquad v_i = \widehat{F}_Y(y_i), \quad i = 1, \ldots, n.$$

Parameter estimation is performed using a two-step procedure. An initial estimate is obtained using inversion of Kendall's tau, followed by maximum likelihood estimation (MLE) to obtain the final parameter value. This approach improves numerical stability and ensures consistent parameter estimation.

It should be noted that the Frank copula does not exhibit tail dependence. Therefore, it is particularly useful for modeling weak or moderate dependence structures, especially when extreme events do not tend to occur simultaneously.

2.8. Return Level

The return level is defined as a threshold value that is expected to be exceeded, on average, once within a specified return period of T years. In the context of the Generalized Extreme Value (GEV) distribution, the return level is obtained by inverting the cumulative distribution function [5, 27]. For a univariate GEV model, the T-year return level x_T is defined as the solution to

$$\Pr(X \leq x_T) = F_X(x_T) = 1 - \frac{1}{T}. \tag{8}$$

If $X \sim \text{GEV}(\mu, \sigma, \xi)$ with $\xi \neq 0$, the return level is given by

$$x_T = \mu + \frac{\sigma}{\xi} \left[\left\{ -\ln \left(1 - \frac{1}{T} \right) \right\}^{-\xi} - 1 \right]. \tag{9}$$

For $\xi = 0$, corresponding to the Gumbel case, the return level is expressed as

$$x_T = \mu - \sigma \ln \left[-\ln \left(1 - \frac{1}{T} \right) \right]. \tag{10}$$

In this study, return levels are estimated for the seasonal block maxima of temperature and humidity in Pekanbaru City. These estimates represent the magnitude of extreme events expected to be exceeded, on average, once within a given return period (e.g., 10 or 20 years). The obtained return levels are subsequently used as thresholds in the joint analysis to evaluate the frequency of compound extreme events under the AND and OR scenarios.

2.9. Joint Return Period

In addition to univariate return levels, this study considers the joint return period to characterize the recurrence of extreme events involving temperature and humidity. Joint return periods are defined under two exceedance conditions, namely the OR case and the AND case [12, 27].

The *OR* case, denoted by $\{X > a \vee Y > b\}$, represents the condition in which at least one of the variables exceeds its threshold. Using the complement rule, this probability is expressed as

$$\Pr(X > a \vee Y > b) = 1 - \Pr(X \leq a, Y \leq b) = 1 - F_{X,Y}(a, b).$$

Thus, the joint return period for the OR case is defined as

$$T_{X>a \vee Y>b} = \frac{1}{1 - F_{X,Y}(a, b)}. \tag{11}$$

The *AND* case, denoted by $\{X > a \wedge Y > b\}$, corresponds to the simultaneous exceedance of both variables. Using the inclusion–exclusion principle, the joint exceedance probability is given by

$$\Pr(X > a \wedge Y > b) = 1 - F_X(a) - F_Y(b) + F_{X,Y}(a, b).$$

Accordingly, the joint return period for the AND case is expressed as

$$T_{X>a \wedge Y>b} = \frac{1}{1 - F_X(a) - F_Y(b) + F_{X,Y}(a, b)}. \tag{12}$$

In this study, the joint distribution function $F_{X,Y}(x, y)$ is obtained from the selected dependence model, including the BEVL model and copula-based approaches. The marginal distributions F_X and F_Y are given by the fitted GEV models for temperature and humidity, respectively. These formulations are used to evaluate the frequency of compound extreme events under different dependence structures.

2.10. Akaike Information Criterion (AIC)

The Akaike Information Criterion (AIC) is used to evaluate and compare the performance of competing statistical models by balancing goodness of fit and model complexity. For a model with k estimated parameters and a maximum log-likelihood value $\hat{\ell}$, the AIC is defined as

$$\text{AIC} = -2\hat{\ell} + 2k. \tag{13}$$

A model with a smaller AIC value is generally considered to provide a better trade-off between model fit and parsimony. In this study, the AIC is used to compare several dependence models, including the Bivariate Extreme Value Logistic (BEVL) model and copula-based approaches, namely the Joe copula and the Frank copula.

The comparison of AIC values is intended to assess the relative performance of each model in representing the dependence structure between extreme temperature and humidity. Rather than identifying a single definitive model, the AIC is used as a complementary criterion to support model selection and interpretation in conjunction with the dependence characteristics observed in the data.

3. Results and Discussion

This section presents the results of the extreme value analysis for temperature and humidity in Pekanbaru City. The discussion begins with descriptive statistics and seasonal block maxima extraction, followed by the estimation of marginal GEV models and dependence structures using the BEVL and copula approaches. The results are subsequently interpreted through joint return period analysis and model comparison using the Akaike Information Criterion (AIC).

3.1. Descriptive Statistics

Descriptive statistics are used to provide an initial overview of the observed data. Prior to the analysis, a data preprocessing step was conducted to ensure data quality and reliability. Observations with zero values in temperature and relative humidity were excluded, as such values are not physically meaningful under normal atmospheric conditions in the study area.

Table 1 presents the descriptive statistics of temperature and relative humidity in Pekanbaru City after the data cleaning process. The table summarizes the mean, minimum, maximum, and standard deviation of each variable. These statistics provide an initial description of the variability and range of the observed climate data.

Table 1: Descriptive Statistics of Temperature and Humidity in Pekanbaru (Cleaned Data)

Description	Variable	
	Temperature (x)	Humidity (y)
Mean	32.58	81.97
Min.	23.14	51.00
Max.	36.80	100.00
Std. Dev.	1.63	5.84

As shown in Table 1, the observed temperature ranges from 23.14 °C to 36.80 °C, with an average value of 32.58 °C. Meanwhile, relative humidity varies between 51% and 100%, with a mean of 81.97%. The larger standard deviation of humidity compared to temperature indicates that humidity exhibits greater variability over time, reflecting the dynamic nature of humidity conditions in the study area.

Overall, the relatively high maximum values and observed variability in both variables indicate the presence of extreme observations. These characteristics are consistent with the application of Extreme Value Theory (EVT) in the subsequent analysis. In addition, the descriptive results provide an empirical basis for further modeling of extreme climate behavior.

3.2. Block Maxima

In this study, extreme values of temperature and humidity were obtained using a seasonal block maxima approach. The daily observations were divided into two main seasonal periods, namely ONDJFM (October–March) and AMJJAS (April–September), following the Asian–Australian monsoon system that influences climatic variability in Indonesia. This seasonal classification is intended to represent relatively homogeneous climatic conditions within each block.

The dataset covers daily observations from January 2014 to December 2024. For the ONDJFM season, months from October to December were assigned to the following season year, while months from January to March remained in the corresponding year. Based on this seasonal-year definition, a total of 24 seasonal block maxima were obtained.

The first and last seasonal blocks are partial blocks because the observation period starts in January 2014 and ends in December 2024. These partial blocks were retained to preserve temporal continuity and maintain the number of extreme observations. However, because partial blocks may not be fully comparable with complete seasonal blocks, this issue is acknowledged as a limitation and the resulting estimates are interpreted cautiously. Although a formal sensitivity analysis was not conducted, the inclusion of partial seasonal blocks is not expected to substantially affect the overall results, as the analysis primarily indicates a weak dependence structure.

The maximum temperature and humidity values were extracted within each seasonal block. The resulting seasonal extreme values are presented in Table 2. Seasonal extreme temperatures range from 32.6 °C to 36.8 °C, while humidity ranges from 90.1% to 100%. These values represent the observed seasonal extremes during the study period. The extracted block maxima are subsequently used for GEV modeling and dependence analysis.

Table 2: Seasonal Extreme Values of Temperature and Humidity Based on the Block Maxima Method

Season	Block	Temperature (x)	Humidity (y)
ONDJFM	1	35.2	94
AMJJAS	2	35.9	93
ONDJFM	3	35.0	92
AMJJAS	4	35.2	94
ONDJFM	5	35.8	95
AMJJAS	6	36.6	93
ONDJFM	7	35.3	97
AMJJAS	8	35.4	91
ONDJFM	9	35.4	95
AMJJAS	10	36.4	96
ONDJFM	11	36.0	96
AMJJAS	12	36.8	95
ONDJFM	13	35.5	98
AMJJAS	14	36.0	95
ONDJFM	15	35.6	96
AMJJAS	16	36.0	94
ONDJFM	17	35.7	98
AMJJAS	18	35.9	97
ONDJFM	19	34.5	99
AMJJAS	20	35.7	100
ONDJFM	21	35.0	98
AMJJAS	22	35.7	96
ONDJFM	23	36.1	94
AMJJAS	24	32.6	90.1

3.3. Estimation of Generalized Extreme Value (GEV) Parameters

The Generalized Extreme Value (GEV) distribution was employed to model the seasonal maxima of temperature and humidity in Pekanbaru City. Parameter estimation was carried out using the Maximum Likelihood Estimation (MLE) method. The estimated GEV parameters for both variables are summarized in Table 3.

Table 3: Estimated Parameters of the GEV Distribution

Parameter	Variable	
	Temperature (x)	Humidity (y)
Location (μ)	34.5470	91.3289
Scale (σ)	0.7708	3.6162
Shape (ξ)	−0.2966	−0.3910

Based on these estimates, the cumulative distribution functions (CDFs) of the fitted GEV distributions for temperature and humidity are expressed as

$$F_X(x) = \exp\left\{-\left[1 - 0.2966\left(\frac{x - 34.5470}{0.7708}\right)\right]^{\frac{1}{0.2966}}\right\}, \tag{14}$$

$$F_Y(y) = \exp\left\{-\left[1 - 0.3910\left(\frac{y - 91.3289}{3.6162}\right)\right]^{\frac{1}{0.3910}}\right\}. \tag{15}$$

The estimated location parameters indicate the central tendency of extreme temperature and humidity at approximately 34–35 °C and 91%, respectively. The larger scale parameter for humidity indicates greater variability in extreme humidity compared to temperature. The negative shape parameters ($\xi < 0$) for both variables indicate that the extremes follow a Weibull-type GEV distribution, implying the presence of an upper bound in the distribution.

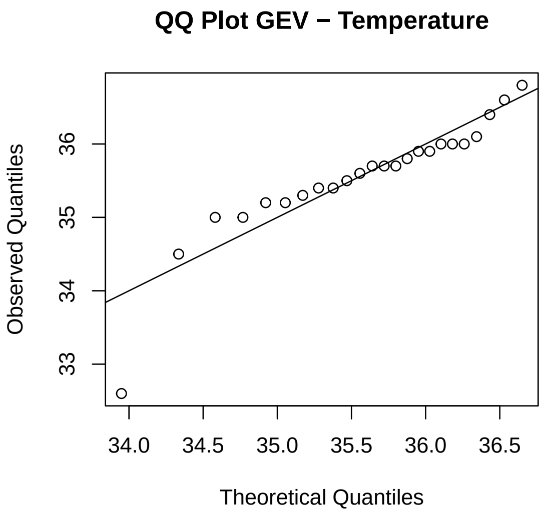
The fitted GEV distributions are subsequently used to transform the extreme observations into the standard $[0, 1]$ interval through their marginal cumulative distribution functions. This transformation is used for the subsequent dependence modeling using the Bivariate Extreme Value Logistic (BEVL) model and copula-based approaches. The transformed pseudo-observations provide a standardized basis for evaluating the dependence structure between extreme temperature and humidity.

3.4. Goodness-of-Fit of the GEV Model

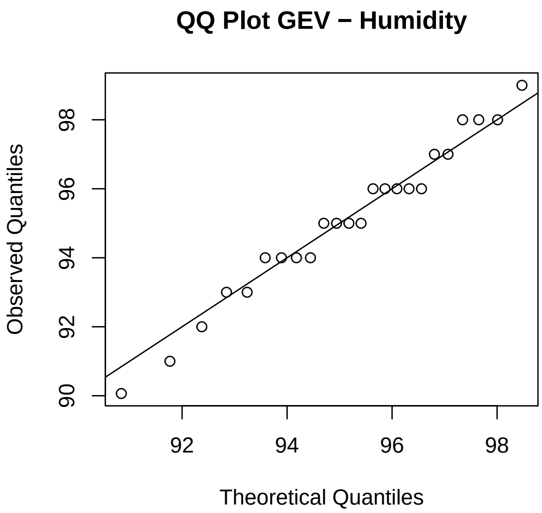
To evaluate the adequacy of the fitted Generalized Extreme Value (GEV) distributions for seasonal maximum temperature and humidity, both statistical and graphical goodness-of-fit (GOF) diagnostics were performed, including the Kolmogorov–Smirnov (KS) test and quantile–quantile (QQ) plots. These diagnostics were conducted to assess the consistency between the fitted GEV distributions and the observed seasonal extremes. The evaluation provides an important basis for subsequent dependence modeling.

The KS test results indicate that the fitted GEV distributions are consistent with the observed data. For temperature, the test yields a statistic of $D = 0.1679$ with a p-value of 0.4587. For humidity, the test statistic is $D = 0.1262$ with a p-value of 0.7945. Since both p-values are greater than the commonly used significance level of 0.05, there is no statistical evidence to reject the null hypothesis that the data follow the fitted GEV distributions.

In addition to the statistical test, QQ plots were used to visually assess the agreement between empirical and theoretical quantiles. Fig. 1 shows that the observed data points generally lie close to the reference line, suggesting a reasonable agreement between the fitted model and the observed extremes. Some deviations from the reference line are observed at the lower and upper tails, particularly for temperature. These deviations are relatively limited and are likely influenced by the small number of seasonal block maxima used in the analysis.



(a) Temperature



(b) Humidity

Fig. 1: QQ plots of the fitted GEV distributions for seasonal maximum temperature and humidity.

Overall, the graphical diagnostics are broadly consistent with the statistical test results. Therefore, the fitted GEV distributions are considered adequate for representing the marginal behavior of extreme temperature and humidity, while the results should be interpreted with caution given the limited sample size. The fitted GEV distributions are subsequently used to transform the extreme observations into the standard $[0, 1]$ interval through their marginal cumulative distribution functions for the subsequent dependence modeling.

3.5. Estimation of Bivariate Extreme Value Logistic (BEVL) Model Parameters

The dependence between seasonal maximum temperature and humidity was modeled using the Bivariate Extreme Value Logistic (BEVL) model. Parameter estimation was performed using the Maximum Likelihood Estimation (MLE) method. The estimated parameters are presented in Table 4.

Table 4: Estimated Parameter Values of the BEVL Model

Parameter	**Variable**	
	Temperature (x)	**Humidity (y)**
Location (μ)	34.5723	91.3346
Scale (σ)	0.8612	3.6321
Shape (ξ)	−0.2191	−0.2992
Dependence (α)	0.9998	

The negative shape parameters ($\xi < 0$) indicate that the marginal extremes of temperature and humidity follow Weibull-type GEV distributions, implying the presence of an upper bound. The estimated dependence parameter $\alpha = 0.9998$ is very close to unity. Under the BEVL model, $\alpha = 1$ corresponds to independence, whereas smaller values indicate stronger dependence. Therefore, this result suggests that the dependence between extreme temperature and humidity is weak and close to independence.

This result is consistent with the findings from the Joe copula model, where the estimated parameter also converges to the independence case. The consistency between the BEVL and copula-based approaches indicates that the dependence structure is relatively weak. These results suggest that joint exceedance of seasonal block maxima of temperature and humidity is relatively limited. However, this interpretation should be made with caution, as the relatively small number of seasonal block maxima may affect the stability of dependence estimation.

3.6. Joe Copula Approach

The dependence between seasonal maximum temperature and humidity was further examined using the Joe copula based on pseudo-observations derived from the fitted Generalized Extreme Value (GEV) marginal distributions. Parameter estimation was performed using maximum likelihood estimation, yielding $\theta = 1$. In the Joe copula family, a parameter value of $\theta = 1$ corresponds to the independence copula. This result suggests that the dependence between extreme temperature and humidity is weak and effectively close to independence.

The upper-tail dependence coefficient is given by $\lambda_U = 2 - 2^{1/\theta}$. Substituting $\theta = 1$ yields $\lambda_U = 0$, indicating the absence of upper-tail dependence. This result indicates that joint exceedance of seasonal block maxima in the upper tail is limited.

The convergence of the Joe copula parameter to the independence boundary reflects the weak dependence structure present in the data. However, this result should be interpreted with caution, as the relatively small number of seasonal block maxima may affect the stability of dependence estimation.

3.7. Joint Return Period Using Joe Copula

The joint return period is used to quantify the frequency of extreme temperature and humidity events under different exceedance scenarios. In this study, joint probabilities were estimated using the Joe copula combined with the marginal Generalized Extreme Value (GEV) distributions. The extreme thresholds were defined based on the 10-year return levels of each variable.

Two exceedance scenarios were considered: the AND case, representing the joint exceedance of seasonal block maxima of temperature and humidity, and the OR case, representing the exceedance of at least one variable. The estimated joint return period for the AND case is on the order of 100 years, whereas the OR case is approximately 5.26 years. Since the analysis is based on two seasonal blocks per year, the return periods are initially obtained in units of seasonal blocks and subsequently converted into years by dividing by two.

These results indicate that joint exceedance of seasonal block maxima is less frequent than single-variable exceedances. This pattern is consistent with the weak dependence structure identified in the copula analysis. Fig. 2 illustrates that only a limited number of observations exceed both thresholds within the same seasonal block, indicating that joint exceedance events are relatively infrequent in the dataset.

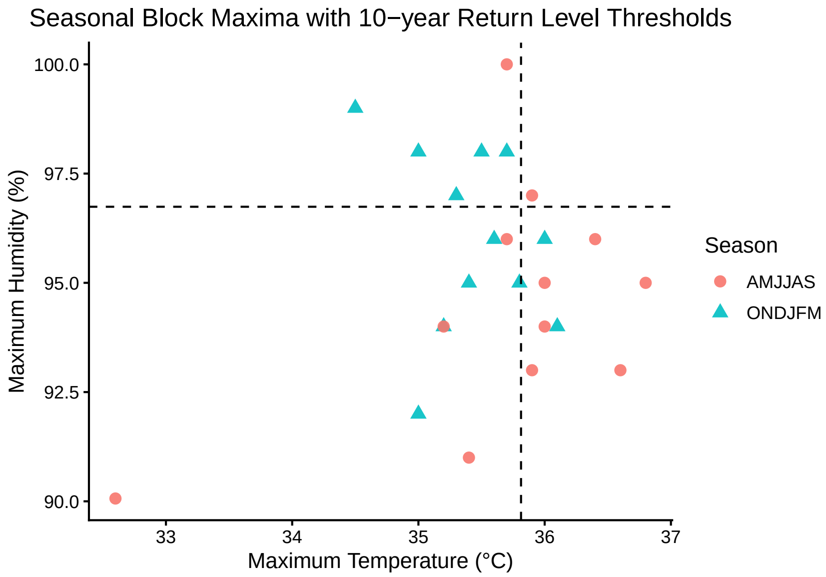


Fig. 2: Scatter plot of seasonal block maxima of maximum temperature and humidity with 10-year return level thresholds.

The Joe copula estimation yields a dependence parameter $\theta = 1$ and an upper-tail dependence coefficient $\lambda_U = 0$, corresponding to the independence case. This result suggests that the dependence between extreme temperature and humidity is weak and close to independence, particularly in the upper tail. It should be noted that the estimation of long return periods (e.g., on the order of 100 years) is subject to substantial uncertainty due to the limited sample size of 24 seasonal block maxima. Therefore, these values should be interpreted as exploratory rather than precise estimates.

For comparison, the joint return periods were also estimated using the Frank copula. The estimated return period for the AND case is on the order of 120.02 years, while the OR case is approximately 5.22 years. These values are broadly consistent with those obtained from the Joe copula and further support the indication of weak dependence.

Overall, the results from both copula models suggest a weak dependence structure between extreme temperature and humidity, although this conclusion should be interpreted with caution given the limited sample size.

3.8. Frank Copula Approach

To further examine the dependence structure between seasonal maximum temperature and humidity, the Frank copula was employed as an alternative Archimedean copula model. Parameter estimation was carried out in two stages, namely inversion of Kendall's tau to obtain an initial value, followed by maximum likelihood estimation for the final parameter estimate. The initial estimation based on Kendall's tau yielded a parameter value of $\theta = -0.5966$, while the final maximum likelihood estimation resulted in $\theta = -0.4326$. The optimization procedure converged successfully.

The negative parameter value indicates a weak negative dependence between extreme temperature and humidity. This suggests that the association between the two variables is limited and remains relatively weak. Unlike the Joe copula, which converges to the independence case, the Frank copula yields a non-zero parameter estimate, indicating the presence of a weak global dependence structure.

However, as the Frank copula does not exhibit tail dependence, this association should be interpreted as weak overall dependence rather than upper-tail dependence. Overall, the Frank copula results indicate that the dependence structure between extreme temperature and humidity in Pekanbaru is weak and slightly negative. This result is consistent with the findings from the BEVL and Joe copula models, which also suggest that strong dependence is not supported by the data.

3.9. Akaike Information Criterion (AIC)

The Akaike Information Criterion (AIC) is employed to evaluate the performance of likelihood-based models by balancing goodness of fit and model complexity, where a smaller AIC value indicates a more parsimonious model [19, 20]. In this study, AIC is used to provide an indicative comparison of the BEVL model, the Joe copula, and the Frank copula. The AIC values obtained from the fitted models are presented in Table 5.

Table 5: AIC values of the competing dependence models

Model	AIC Value
Bivariate Extreme Value Logistic (BEVL)	170.935
Joe Copula	2.000
Frank Copula	1.924

It is important to note that the AIC values reported in this study should be interpreted with caution. The BEVL model is fitted directly to the joint distribution of the data, whereas the copula models are estimated based on pseudo-observations obtained from the fitted marginal GEV distributions. As a result, the underlying likelihood structures are not strictly comparable. Therefore, the AIC values presented here are used only as an indicative measure rather than a definitive criterion for model selection.

Based on Table 5, the Frank copula yields the lowest AIC value among the fitted copula models. However, the difference between the AIC values of the Frank and Joe copulas is very small and does not provide sufficient evidence to support a clear distinction between the two models. The Joe copula produces an AIC value close to 2, which corresponds to a log-likelihood value near zero.

This situation arises when the model converges to the independence case ($\theta = 1$), indicating that the dependence between the variables is weak. Therefore, the AIC value of the Joe copula reflects the near-independence structure of the data rather than superior model performance. In contrast, the BEVL model yields a substantially larger AIC value. Given the differences in model formulation and likelihood structure, this result should be interpreted cautiously and only as an indicative comparison.

Overall, the AIC results support the finding that the dependence between extreme temperature and humidity is weak, and no strong conclusion regarding model superiority can be drawn from this comparison.

4. Conclusion

This study applies a bivariate extreme value framework to analyze the joint behavior of extreme temperature and humidity in Pekanbaru City. The marginal results indicate that both variables can be represented using a Weibull-type Generalized Extreme Value (GEV) distribution, implying the presence of finite upper bounds in the fitted models.

The dependence analysis indicates that the relationship between extreme temperature and humidity is weak. The Joe copula converges to the independence case, while the Frank copula yields the lowest AIC value among the fitted copula models. However, this difference is small and should be interpreted with caution, as the likelihood structures of the models are not strictly comparable. Therefore, the results do not support a strong model selection conclusion.

The joint return period analysis indicates that joint exceedance of seasonal block maxima is less frequent than single-variable exceedances. Nevertheless, long return period estimates are subject to substantial uncertainty due to the limited number of seasonal block maxima and should be regarded as exploratory rather than precise.

Overall, the findings suggest a weak dependence structure between the variables, and the comparison between the BEVL and copula approaches should be interpreted as exploratory. These results provide a preliminary statistical description of extreme climate behavior in the study area, while their interpretation should remain cautious given the data limitations.

CRediT Authorship Contribution Statement

Fadilla Afsari: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Writing–Original Draft, Visualization. **A'yunin Sofro:** Conceptualization, Methodology, Validation, Writing–Review & Editing, Supervision, Project Administration.

Declaration of Generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used DeepL and Grammarly as AI-assisted tools for language translation, grammar checking, and improving sentence clarity. These tools were employed solely to enhance linguistic quality and readability. The use of AI-assisted technologies did not influence the research design, data analysis, results, or scientific conclusions. All content was carefully reviewed and revised by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

Declaration of Competing Interest

The authors declare no competing interests

Funding and Acknowledgments

This research was supported by Non-APBN funding from Universitas Negeri Surabaya under the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia for the 2026–2027 funding period.

The authors would like to express their sincere gratitude to the supervisors for their valuable guidance and constructive suggestions during the research process. The authors also gratefully acknowledge the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) for providing the daily temperature and humidity data from the Sultan Syarif Kasim II Meteorological Station used in this study. In addition, appreciation is extended to all parties who contributed to the completion of this research.

Data and Code Availability

The data supporting the findings of this study were obtained from the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG), specifically from the Sultan Syarif Kasim II Meteorological Station in Pekanbaru. Access to the data is subject to BMKG data sharing policies.

The processed datasets and computational codes used in the analysis are not publicly available but can be obtained from the corresponding author upon reasonable request. Requests for data and code access will be considered for academic and research purposes, subject to confidentiality and data use agreements.

References

[1] IPCC. *Climate Change 2021: The Physical Science Basis*. Cambridge University Press Cambridge, UK and New York, NY, USA, 2021. DOI: 10.1017/9781009157896.

[2] Badan Meteorologi, Klimatologi, dan Geofisika. *Data Iklim Stasiun Meteorologi Sultan Syarif Kasim II Pekanbaru*. 2023. https://dataonline.bmkg.go.id.

[3] Edi Aldrian and R. Dwi Susanto. "Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature". In: *International Journal of Climatology* 23.12 (2003), pp. 1435–1452. DOI: 10.1002/joc.950.

[4] Stuart Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer London, 2001. https://link.springer.com/book/10.1007/978-1-4471-3675-0.

[5] Alexander J. McNeil. *Extreme Value Theory for Risk Managers*. 1999. https://www.sfu.ca/~rjones/econ811/readings/McNeil%201999.pdf.

[6] W. Habibulloh and A. Y. Sofro. "Prediksi suhu udara di Jawa Tengah menggunakan extreme value theory". In: *MATHunesa: Jurnal Ilmiah Matematika* 11.3 (2023), pp. 489–495. DOI: 10.26740/mathunesa.v11n3.p489-495.

[7] Hadi Tabari. "Extreme value analysis dilemma for climate change impact assessment on global flood and extreme precipitation". In: *Journal of Hydrology* 593 (2021), p. 125932. DOI: 10.1016/j.jhydrol.2020.125932.

[8] D. Rypkema and S. Tuljapurkar. "Modeling extreme climatic events using the generalized extreme value distribution". In: *Handbook of Statistics*. Vol. 44. Elsevier, 2021, pp. 39–71. DOI: 10.1016/bs.host.2020.12.002.

[9] Roger B. Nelsen. *An Introduction to Copulas*. 2nd ed. Springer New York, 2006. https://link.springer.com/book/10.1007/0-387-28678-0.

[10] Harry Joe. *Multivariate Models and Multivariate Dependence Concepts*. CRC Press Boca Raton, FL, 1997.

[11] A. Oktaviarina and A. Y. Sofro. "Analysis between temperature and wind speed in East Java using bivariate extreme value theory". In: *Journal of Physics: Conference Series* 1417.1 (2019), p. 012019. DOI: 10.1088/1742-6596/1417/1/012019.

[12] A. E. Kirana and A. Y. Sofro. "Prediksi kelembaban dan curah hujan maksimum di Kabupaten Malang menggunakan bivariate extreme value logistic". In: *MATHunesa: Jurnal Ilmiah Matematika* 12.3 (2024), pp. 465–474. DOI: 10.26740/mathunesa.v12n3.p465-474.

[13] N. Smirnov. "Table for estimating the goodness of fit of empirical distributions". In: *The Annals of Mathematical Statistics* 19.2 (1948), pp. 279–281. DOI: 10.1214/aoms/1177730256.

[14] M. B. Wilk and R. Gnanadesikan. "Probability plotting methods for the analysis of data". In: *Biometrika* 55.1 (1968), pp. 1–17. DOI: 10.1093/biomet/55.1.1.

[15] N. C. Dzupire, P. Ngare, and L. Odongo. "A copula-based bivariate model for temperature and rainfall processes". In: *Scientific African* 8 (2020), e00365. DOI: 10.1016/j.sciaf.2020.e00365.

[16] A. Y. Sofro, W. Habibulloh, and K. N. Khikmah. "Prediction of air temperature using spatial extreme value with copula approach". In: *Journal of Theory and Applications in Mathematics* 8.4 (2024), pp. 1217–1232. DOI: 10.31764/jtam.v8i4.25436.

[17] M. J. Frank. "On the simultaneous associativity of F(x, y) and x + y - F(x, y)". In: *Aequationes Mathematicae* 19.1 (1979), pp. 194–226. DOI: 10.1007/BF02189866.

[18] A. C. Favre et al. "Multivariate hydrological frequency analysis using copulas". In: *Water Resources Research* 40.1 (2004). DOI: 10.1029/2003WR002456.

[19] Hirotugu Akaike. "A new look at the statistical model identification". In: *IEEE Transactions on Automatic Control* 19.6 (1974), pp. 716–723. DOI: 10.1109/TAC.1974.1100705.

[20] Kenneth P. Burnham and David R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. New York, NY: Springer, 2002. https://link.springer.com/book/10.1007/b97636.

[21] Samuel Kotz and Saralees Nadarajah. "Extreme Value Distributions: Theory and Applications". In: *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer, 2011. https://link.springer.com/referencework/10.1007/978-3-642-04898-2.

[22] Edward Omey, Francisco Mallor, and Eulalia Nualart. *An Introduction to Statistical Modelling of Extreme Values: Application to Calculate Extreme Wind Speeds*. Tech. rep. KU Leuven, Nov. 2009. https://lirias.kuleuven.be/retrieve/231511.

[23] Ana Ferreira and Laurens de Haan. "On the block maxima method in extreme value theory: PWM estimators". In: *The Annals of Statistics* 43.1 (2015), pp. 276–298. DOI: 10.1214/14-AOS1280.

[24] Christopher P. Zwonik. "Assessing Trends in Future Precipitation Extremes in the Northeastern United States Using the Method of Block Maxima". Master's thesis. 2020.

[25] J. D. Prang. "Sebaran Nilai Ekstrim Terampat dalam Fenomena Curah Hujan". Undergraduate thesis. Institut Pertanian Bogor, 2006. http://repository.ipb.ac.id/handle/123456789/43133.

[26] Richard Minkah. "An application of extreme value theory to the management of a hydroelectric dam". In: *SpringerPlus* 5.1 (2016), pp. 1–12. DOI: 10.1186/s40064-016-1719-2.

[27] Ning Li, Xiaohui Liu, Wei Xie, Jie Wu, and Peng Zhang. "The return period analysis of natural disasters with statistical modeling of bivariate joint probability distribution". In: *Risk Analysis* 33.1 (2013), pp. 134–145. DOI: 10.1111/j.1539-6924.2012.01838.x.