BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1203-1216
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1203-1216

JOINT DISTRIBUTION AND PROBABILITY DENSITY OF CLIMATE FACTORS IN KALIMANTAN USING NESTED COPULAS

Sri Nurdiati[1*], Teduh W. Mas'oed[2], Mohamad K. Najib[3], Dewi Rahmawati[4]

[1,2,3,4]Department of Mathematics, Faculty of Mathematics and Natural Sciences, IPB University Jln. Meranti, Babakan, Dramaga, Bogor, 18660, Indonesia

*Corresponding author's e-mail: *nurdiati@apps.ipb.ac.id*

Article History:

Received: 30th September 2024
Revised: 31st January 2025
Accepted: 28th February 2025
Published: 1st April 2025

Keywords:

Climate Factors;
Copula;
ENSO;
IOD;
Kalimantan;
Multivariate Distribution.

ABSTRACT

In this study, we investigate the joint distribution of local and global climate factors in Kalimantan, Indonesia, using fully and partially nested copula models. The analysis focuses on capturing the dependencies between local factors (precipitation and the number of dry days) and global indices (ENSO and IOD). The methodology involves estimating the marginal distributions of each variable using goodness-of-fit tests, and then modeling the dependence structure between variables with a range of copulas. We used both one-parameter copulas, including Gaussian, Clayton, Gumbel, Joe, and Frank, as well as two-parameter copulas, such as BB1, BB7, and BB8, with rotations of 90°, 180°, and 270° applied to account for negative dependencies. Nested copula structures were employed to model multivariate dependencies, with fully nested and partially nested approaches used to capture interactions between all four variables. The results show that global climate indices, particularly ENSO and IOD, have a more substantial influence during the dry season, impacting drought conditions in Kalimantan. The copula method offers a flexible and efficient way to construct multivariate joint distributions, better representing complex climate relationships than traditional models. Future work could extend this approach to include additional climate variables and use real-time data for forest fire risk prediction.

How to cite this article:

S. Nurdiati, T. W. Mas'oed, M. K. Najib and D. Rahmawati.," JOINT DISTRIBUTION AND PROBABILITY DENSITY OF CLIMATE FACTORS IN KALIMANTAN USING NESTED COPULAS" *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1203-1216, June, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Forest and land fires are significant environmental challenges in Indonesia, particularly affecting regions like Central Sumatra, South Sumatra, South Kalimantan, and Merauke [1]. These areas frequently experience extensive burned land due to various factors, including climatic conditions and human activities. The El Niño-Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) are two key climatic phenomena influencing the occurrence and severity of these fires [2]. The positive phase of the IOD, characterized by warmer sea surface temperatures in the western Indian Ocean, can lead to drier conditions in Indonesia. Similarly, during the warm phase of ENSO, prolonged dry spells and elevated temperatures can exacerbate the likelihood of forest and land fires.

Indonesia is a country located at the equator, so many factors influence the season system in Indonesia, including global phenomena such as ENSO and IOD [3]. The Indian Ocean Dipole (IOD) is a phenomenon resulting from the interaction between the sea and the atmosphere in the Indian Ocean. IOD consists of positive and negative IOD. Meanwhile, El Nino-Southern Oscillation (ENSO) is a phenomenon of deviation from sea surface temperatures in the Pacific Ocean near the central and eastern equator. The ENSO phenomenon is divided into two phases: El Nino, the warm ENSO phase, and La Nina, the cold ENSO phase [4].

The phenomenon of positive IOD and El Nino co-occurring will cause a decrease in sea surface temperatures in Indonesian waters. This will impact reducing the intensity of rainfall, which will indirectly create a drought risk in several regions of Indonesia [5]. Drought causes litter and trees to dry out or even die, and groundwater reserves will decrease, thus triggering forest and land fires.

Several studies have previously been carried out to examine the influence of rainfall and global climate indicators, namely ENSO and IOD, on the incidence of forest and land fires in Indonesia. Nuralita et al. [5] studied the relationship between rainfall and the El Nino climate phenomenon and positive IOD, which was analyzed using statistical time functions. Ryadi et al. [6] discussed the influence of El Nino and La Nina on the level of land drought, distribution of rainfall and sea surface temperature using linear regression statistical analysis. Najib et al. [7] used the copula method to study the influence of ENSO on the joint distribution of drought indicators, such as total precipitation, the number of dry days, and precipitation anomalies in relation to increased risk of forest fires.

Based on the research that has been carried out, the multidimensional copula method is not widely used in research on the influence of drought and global climate indicators on forest fire incidents in Indonesia. Copula is a function that allows users to separate marginal distributions from dependency structures in certain multivariate distributions [8]. In multivariate statistical analysis, copula is a valuable approach to studying non-linear dependencies between events in multivariate cases. Thus, the copula can model the dependency structure between indicators and is flexible enough to model data that does not meet the characteristics of a normal distribution. As an annual disaster in Indonesia, forest and land fires cause various losses. It would be better if this phenomenon could be detected early as an initial step in preventing it.

Therefore, this study analyzes the joint distribution between total precipitation and the number of dry days in Kalimantan, ENSO, and IOD. The total precipitation and the number of dry days data used in this research have been processed in previous research by Najib et al. [7]. Meanwhile, data on global climate indicators, i.e., ENSO and IOD, were obtained from the Oceanic Nino index, SST 3.4 Index (ONI), and Dipole Mode Index (DMI). The analysis was carried out using the fully and partially nested copulas method based on the Inference of Function for Margins (IFM) to estimate copula parameters.

2. RESEARCH METHODS

A copula is a statistical concept that represents the dependence pattern between random variables in probability theory and multivariate statistics. It is beneficial when working with complicated, high-dimensional data where the relationships between variables are difficult to define using conventional approaches. Copulas are used in various disciplines, including economics, finance, hydrology, and earth science [9]–[11]. Copula theory's main premise is to distinguish between modeling marginal distributions and modeling the joint distribution, enabling more flexible and thorough representations of dependence patterns.

2.1. Definition of Copulas

Sub-copulas are first defined as a specific class of grounded 2-increasing functions with margins, and then copulas are defined as sub-copulas with domain $\boldsymbol{I}$ where $\boldsymbol{I} = [\mathbf{0}, \mathbf{1}]$.

Definition 1. *A 2-dimensional sub-copula (or 2-subcopula, or just a sub-copula) is a function C' that has the following properties:*

1. *Domain $C' = S_1 \times S_2$ where S_1 and S_2 are subsets of I containing 0 and 1;*
2. *C' is grounded and 2-increasing;*
3. *For every $u \in S_1$ and $v \in S_2$, $C'(u,1) = u$ and $C'(1,v) = v$*

Definition 2. *A 2-dimensional copula (or 2-copula, or just a copula) is a 2-subcopula C whose domain is I^2.*

Equivalently, a copula is a function C from I^2 to I with the following properties:

1. For every $u, v \in I$, $C(u,0) = C(0,v) = 0$, $C(u,1) = u$, and $C(1,v) = v$
2. For every $u_1, u_2, v_1, v_2 \in I$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$$

It should be noted that for every (u, v) in C', $0 \leq C' \leq 1$, implying that Range C' is likewise a subset of I. In summary, copulas model the dependence structure among several random variables. In contrast, sub-copulas are a subset of copulas that explicitly explain the dependence structure among a specified subset of those variables.

2.1.1. Sklar's Theorem

The foundation of many, if not most, of the practical applications of the theory of copulas to statistics is Sklar's theorem, which is essential to understanding that theory. Sklar's theorem explains the importance of copulas in the link between multivariate distribution functions and their univariate margins. As a result, this section opens with a quick overview of distribution functions.

Definition 3. *A cumulative distribution function (or distribution function, or CDF) is a function F_X with domain $\mathbb{R}$ that has the property*

1. *F_X is non-decreasing,*
2. *$F_X(-\infty) = 0$ and $F_X(\infty) = 1$.*

Definition 4. *A joint cumulative distribution function (or joint distribution function, or joint CDF) is a function F_{XY} with domain $\mathbb{R}^2$ that has the property*

1. *F_{XY} is 2-increasing,*
2. *$F_{XY}(x, -\infty) = F_{XY}(-\infty, x) = 0$ and $F_{XY}(\infty, \infty) = 1$*

Based on **Definition 4**, F_{XY} is grounded, and because Domain $F_{XY} = \mathbb{R}^2$. F_{XY} has margins F_X and F_Y given by $F_X(x) = F_{XY}(x, \infty)$ and $F_Y = F_{XY}(\infty, y)$. Since F_{XY} is a grounded 2-increasing function, F_{XY} is non-decreasing in each argument, meaning that the horizontal, vertical, and diagonal sections of a joint distribution function F_{XY} are all non-decreasing. Therefore, F_X and F_Y are distribution functions. Sklar's Theorem is a fundamental result in the field of copula theory, which is used to study the dependence structure between random variables. Sklar's Theorem provides a way to construct multivariate joint distribution functions from marginal distribution function of individual variables and a copula function that describes their dependence. Copulas allow us to separate the effect of dependence from effects of the marginal distributions [12].

Theorem 1. **(Sklar's Theorem).** *Let X and Y be random variables with marginal cumulative distribution functions (CDFs) $F_X(x)$ and $F_Y(y)$, respectively. Then there exists a copula function $C(u, v)$ such that for any values x and y in the support of the respective marginal distributions:*

$$F_{XY}(x, y) = C\big(F_X(x), F_Y(y)\big) \tag{1}$$

where F_{XY} is the joint distribution function (CDF) of X and Y. If F_X and F_Y are continuous, then C is unique; otherwise, C is uniquely determined on Range $F_X \times$ Range F_Y.

In other words, the joint distribution functions can be expressed as a function of the marginal distribution functions through a copula function, which characterizes the dependence structure between the variables. This theorem is particularly useful in cases where you want to separate the modeling of marginal distributions from the modeling of dependence. Copula methods have become ubiquitous when analyzing, modeling, and quantifying the dependence between variables [13].

2.2. Copula-based Joint Density Function

Given the copula and the marginal distributions of the variables, the joint density function can be derived. It describes the likelihood of observing specific values for all the variables simultaneously.

Definition 5. *A probability density function (or density function, or PDF) is a function f_X with domain $\mathbb{R}$ that has the property*

1. *f_X is non-negative, $f_X(x) \geq 0$,*
2. $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
3. $P(a \leq x \leq b) = \int_a^b f_X(x)\, dx$

Definition 6. *A joint probability density function (or joint density function, or joint PDF) is a function f_{XY} with domain $\mathbb{R}^2$ that has the property*

1. *f_{XY} is non-negative, $f_{XY}(x,y) \geq 0$,*
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\, dx\, dy = 1$
3. $P(a \leq x \leq b, c \leq y \leq d) = \int_c^d \int_a^b f_{XY}(x,y)\, dx\, dy$

The distribution function (CDF) and density function (PDF) are complementary concepts for describing random variables. The CDF gives a cumulative view of probabilities, while the PDF provides a detailed view of probabilities at specific values. They are mathematically related through differentiation (for continuous random variables) and thoroughly describe a random variable's probability distribution. Therefore, several references provide other definitions of CDF, as follows.

Definition 7. *A cumulative distribution function (or distribution function, or CDF) is a function F_X with domain $\mathbb{R}$ defined for each number x, such that:*

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt \tag{2}$$

where f_X is a probability density function.

2.2.1. Consequence of Sklar's Theorem

Just as a distribution function has a density function, a copula also has a density. Therefore, a copula can also be written as the integral of its density function.

Definition 8. *Let C is a unique copula, then C can be expressed as:*

$$C(u,v) = \int_0^v \int_0^u c(s,t)\, ds\, dt \tag{3}$$

where $u = F_X(x)$ and $v = F_Y(y)$ and c is the corresponding copula density function.

The important consequence of Sklar's theorem then stated that every joint probability density f_{XY} is also writable by the product of its marginal probability densities f_X and f_Y and the copula density c.

Corollary 1. *Let f_{XY} is a joint density with marginal densities f_X and f_Y, then there exists a copula density c such that*

$$f_{XY}(x,y) = c\big(F_X(x), F_Y(y)\big) \cdot f_X(x) \cdot f_Y(y) \tag{4}$$

Proof. By deriving the right and left sides of Sklar's Theorem (**Equation (1)**) with respect to *x* and *y*, then **Equation (4)** proven. ■

2.3. High-Dimension: Nested Copula

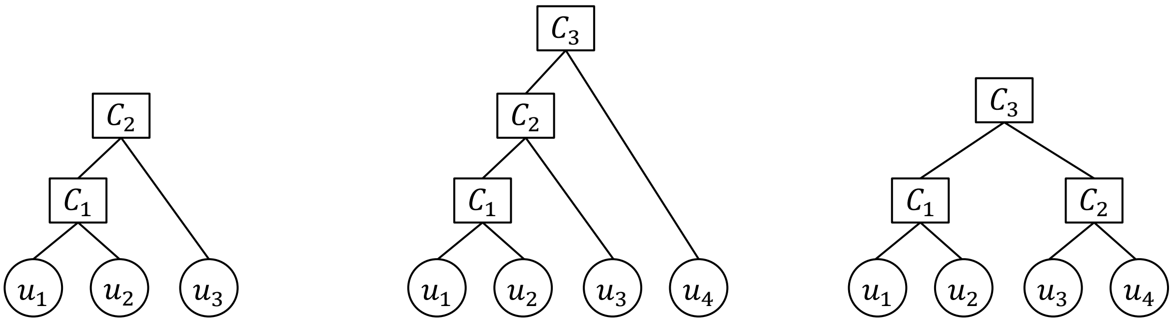
For high-dimensional cases, Sklar's theorem can be extended as follows.

Theorem 2. **(Sklar's Theorem)** *Let $X_1, X_2, \dots, X_d$ be random variables with marginal cumulative distribution functions (CDFs) $F_1, F_2, \dots, F_d$, respectively. Then there exists a copula function C such that for any values $x_1, x_2, \dots, x_d$ in the support of the respective marginal distributions:*

$$F_{X_1,X_2,\dots,X_d}(x_1, x_2, \dots, x_d) = C\big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big) \tag{5}$$

where $F_{X_1,X_2,\dots,X_d}$ is the joint distribution function (CDF) of $X_1, X_2, \dots, X_d$. If $F_1, F_2, \dots, F_d$ are continuous, then C is unique; otherwise, C is uniquely determined on Range $F_1 \times$ Range $F_2 \times \cdots \times$ Range F_d.

Copulas with higher dimensions, $d \geq 3$, can be constructed by exploiting the arrangement of several bivariate copulas. One approach is a nested copula, which constructs a multi-dimensional copula as a nested structure of a bivariate copula. Nested copulas offer several conveniences in constructing high-dimensional multivariate distributions, such as structural simplicity and computational conveniences. **Figure 1** provides the structural construction of nested copulas in the 3-variate and 4-variate cases. Mathematically, the formulation for nested 3-copula and (fully and partially) nested 4-copula can be written as follows.



(a) **(b)** **(c)**

Figure 1. Nested Structure of High-Dimensional Copulas, (a) Nested 3-Copula, (b) Fully Nested 4-Copula, and (c) Partially Nested 4-Copula

Definition 9. **(Nested 3-Copula)** *Let X_1, X_2, X_3 be random variables with marginal cumulative distribution functions (CDFs) F_1, F_2, F_3, respectively. Then, 3-copula C can be constructed from a nested structure of two bivariate copula (2-copulas) as follows:*

$$C(u_1, u_2, u_3) = C_2(C_1(u_1, u_2), u_3) \tag{6}$$

where $u_i = F_i(x_i)$ for $i = 1,2,3$.

Definition 10. **(Nested 4-Copula).** *Let X_1, X_2, X_3, X_4 be random variables with marginal cumulative distribution functions (CDFs) F_1, F_2, F_3, F_4, respectively. Then, 4-copula C can be constructed from a nested structure of three bivariate copula (2-copulas) as follows:*

1. *Fully nested 4-copula*

$$C(u_1, u_2, u_3, u_4) = C_3(C_2(C_1(u_1, u_2), u_3), u_4) \tag{7}$$

2. *Partially nested 4-copula*

$$C(u_1, u_2, u_3, u_4) = C_3\big(C_1(u_1, u_2), C_2(u_3, u_4)\big) \tag{8}$$

where $u_i = F_i(x_i)$ *for* $i = 1,2,3,4$.

According to **Definition 9**, **Definition 10**, and **Theorem 2**, the joint distribution and probability density functions for the 3-variate and 4-variate cases are given by the following proposition.

Proposition 1. *Let* X_1, X_2, X_3 *be random variables with marginal cumulative distribution functions (CDFs)* F_1, F_2, F_3, *respectively. Then, the joint distribution* $F_{1,2,3}$ *can be constructed from a nested structure of two bivariate copula (2-copulas) as follow:*

$$F_{1,2,3}(x_1, x_2, x_3) = C_2\left(C_1\big(F_1(x_1), F_2(x_2)\big), F_3(x_3)\right) \tag{9}$$

As a result, the joint probability density function is written by:

$$f_{1,2,3}(x_1, x_2, x_3) = c_2\left(C_1\big(F_1(x_1), F_2(x_2)\big), F_3(x_3)\right) \cdot c_1\big(F_1(x_1), F_2(x_2)\big) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \tag{10}$$

Proof. By deriving the right and left sides of **Equation (9)** with respect to x_1, x_2 and x_3, then **Equation (10)** proven. ■

Proposition 2. *Let* X_1, X_2, X_3, X_4 *be random variables with marginal cumulative distribution functions (CDFs)* F_1, F_2, F_3, F_4, *respectively. Then, the joint distribution* $F_{1,2,3,4}$ *can be constructed from a* fully nested *structure of three bivariate copula (2-copulas) as follow:*

$$F_{1,2,3,4}(x_1, x_2, x_3, x_4) = C_3\left(C_2\left(C_1\big(F_1(x_1), F_2(x_2)\big), F_3(x_3)\right), F_4(x_4)\right) \tag{11}$$

As a result, the joint probability density function is written by:

$$\begin{aligned} f_{1,2,3,4}(x_1, x_2, x_3, x_4) &= c_3\left(C_2\left(C_1\big(F_1(x_1), F_2(x_2)\big), F_3(x_3)\right), F_4(x_4)\right) \cdot c_2\left(C_1\big(F_1(x_1), F_2(x_2)\big), F_3(x_3)\right) \\ &\quad \cdot c_1\big(F_1(x_1), F_2(x_2)\big) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \end{aligned} \tag{12}$$

Proof. By deriving the right and left sides of **Equation (11)** with respect to x_1, x_2, x_3 and x_4, then **Equation (12)** proven. ■

Proposition 3. *Let* X_1, X_2, X_3, X_4 *be random variables with marginal cumulative distribution functions (CDFs)* F_1, F_2, F_3, F_4, *respectively. Then, the joint distribution* $F_{1,2,3,4}$ *can be constructed from a* partially *nested structure of three bivariate copula (2-copulas) as follow:*

$$F_{1,2,3,4}(x_1, x_2, x_3, x_4) = C_3\left(C_1\big(F_1(x_1), F_2(x_2)\big), C_2\big(F_3(x_3), F_4(x_4)\big)\right) \tag{13}$$

As a result, the joint probability density function is written by:

$$\begin{aligned} f_{1,2,3,4}(x_1, x_2, x_3, x_4) \\ &= c_3\Big(C_1\big(F_1(x_1), F_2(x_2)\big), C_2\big(F_3(x_3), F_4(x_4)\big)\Big) \cdot c_1\big(F_1(x_1), F_2(x_2)\big) \\ &\cdot c_2\big(F_3(x_3), F_4(x_4)\big) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \end{aligned} \quad (14)$$

Proof. By deriving the right and left sides of **Equation (13)** with respect to x_1, x_2, x_3 and x_4, then **Equation (14)** proven. ■

2.4. Estimation Method

The Inference of Function for Margin (IFM) method is a parametric method consisting of two steps, with the basis of each step containing the log likelihood approach **[14]**. This method is usually used to estimate the parameters of a multidimensional copula. The first step in this method is to construct a log likelihood function to estimate the marginal parameter vector $\hat{\alpha}_i$, i.e.

$$\hat{\alpha}_i = \arg\max \ln L_i = \arg\max \ln \prod_{t=1}^{N} f_i\big(x_i^t; \alpha_i\big) \quad (15)$$

where f_i is the probability density function of the random variable X_i. The second step of the IFM method is to estimate the copula parameters by maximizing the log value of the copula likelihood function L. For bivariate cases, it is written as follow.

$$\hat{\theta} = \arg\max \ln L = \arg\max \ln \prod_{t=1}^{N} c(F_1(x_1^t; \hat{\alpha}_1), F_2(x_2^t; \hat{\alpha}_2); \theta) \quad (16)$$

where $\hat{\theta}$ is the estimate of the copula parameter θ and c is the copula probability density function **[15]**, **[16]**. For 3- and 4-variate cases, copulas are formed through a nested structure of bivariate copulas as in **Figure 1**.

The IFM method assumes independent working between marginal parameters and their log likelihood. The rules for estimating parameters with IFM are almost identical to the maximum likelihood method (MLE). We selected the most suitable copula function after the parameters were obtained from the estimation process. Selection is carried out using the goodness-of-fit test on each copula function based on several test statistics such as Root Mean Squared Error (RMSE), Akaike's Information Criterion (AIC), and Cramer-von Mises (CvM).

The fully nested copula structure is constructed with the following procedures (based on the degree of dependence between the pair variables):

a. As the first two variables (1 and 2), select the variables with the highest degree of dependence (rank-based), such as the Kendall-tau correlation.

b. Using variables 1 and 2, estimate the copula.

c. Evaluate the degree of dependence (rank-based) between empirical copula from step 2 with the remaining variables.

d. Select variable 3, which has the maximum degree of dependence (rank-based) with the copula constructed using variables 1 and 2.

e. Continue the process until the last variable is considered.

Meanwhile, the partially nested copula structure is constructed by combining two or more fully nested ones.

3. RESULTS AND DISCUSSION

3.1. Datasets

We use four types of data in this research: total precipitation, the number of dry days, the Oceanic Nino Index (ONI), and the Dipole Mode Index (DMI). Total precipitation and dry days represent Kalimantan's local climate conditions. At the same time, ONI and DMI are global climate indices representing conditions in the Pacific and Indian Oceans. These four data sources come from different sources. Below is a brief explanation of each data source used.

Total precipitation and the number of dry days are datasets processed in previous research by Najib et al. [7] with a spatial resolution of 0.25° × 0.25°. The temporal resolution used is monthly. ONI is a time series data downloaded from https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/, calculated using the average Hadley Center Sea Ice and Sea Surface Temperature (HadISST) data with monthly time intervals derived from observations in the Pacific Ocean (5°S-5°N and 170°-120°W). Meanwhile, DMI is a time series data downloaded from https://psl.noaa.gov/gcos_wgsp/Timeseries/DMI/ with monthly time intervals. DMI is data calculated using the average of HadISST data based on the sea surface temperature gradient anomaly between the western Indian Ocean (50°E-70°E and 10°S-10°N) and the southeastern Indian Ocean (90°-110°E and 10°S-0°). We downloaded all data from 2001-2020. Then, we divided the data set into the dry season (May to October) and the rainy season (November to April). The determination of the dry season is based on the six months with the lowest average precipitation, as seen in Figure 2.

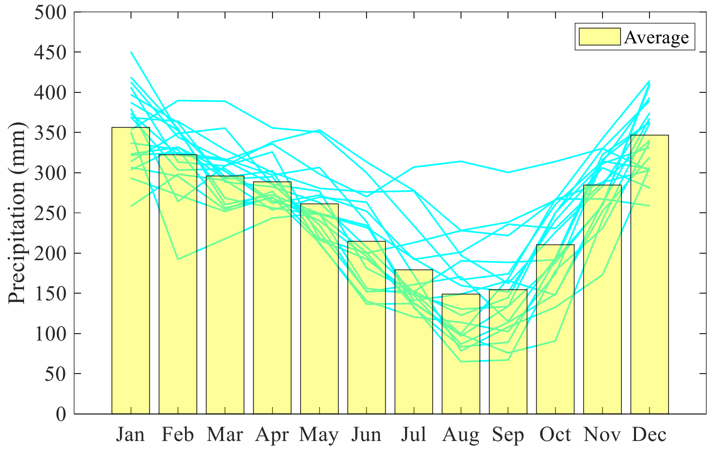


Figure 2. Average Precipitation for Each Month in the Study Area in 2001-2020

3.2. Marginal Distributions

Data analysis begins by identifying the marginal distribution for each data group. The ten types of distribution used are normal, lognormal, extreme value, generalized extreme value, exponential, gamma, Weibull, logistic, loglogistic, and inverse Gaussian distributions. The initial step is to estimate each distribution's parameters by maximizing the likelihood function's log value. After that, the process continues by testing the fit distribution using the Anderson-Darling test. The distribution is then selected based on the most significant -value of Anderson-Darling statistics. Table 1 shows the chosen distribution and its parameters for each variable with the Anderson-Darling test.

Table 1. Selected Marginal Distribution and Its Parameters for Each Variable with Anderson-Darling Test

Variables	Seasons	
	Dry	Rainy
Total precipitation (X_1)	Weibull: A = 217.6901, B = 3.2481 fails to reject h_0 ($p = 0.8233$)	Gamma: $a = 39.4287$, $b = 8.007$ fails to reject h_0 ($p = 0.9677$)
Number of dry days (X_2)	Lognormal: $\mu = 2.7107$, $\sigma = 0.25076$ fails to reject h_0 ($p = 0.5164$)	Loglogistic: $\mu = 2.333$, $\sigma = 0.099$ fails to reject h_0 ($p = 0.6720$)
ONI (X_3)	Logistic: $\mu = 0.054669$, $\sigma = 0.34694$ fails to reject h_0 ($p = 0.9563$)	Generalized Extreme Value: $k = -0.15$, $\sigma = 0.839$, $\mu = -0.356$ fails to reject h_0 ($p = 0.9662$)
DMI (X_4)	Generalized Extreme Value: $k = -0.19$, $\sigma = 0.3173$, $\mu = 0.0206$ fails to reject h_0 ($p = 0.9803$)	Logistic: $\mu = 0.10255$, $\sigma = 0.131$ fails to reject h_0 ($p = 0.9825$)

Table 1 shows that the families of marginal distributions selected for each variable are diverse. No one distribution family dominates. However, all selected distributions passed the Anderson-Darling statistical test, as indicated by their *p*-values of more than 5%. **Figure 3** compares each variable's empirical and theoretical distribution in the dry and rainy seasons.

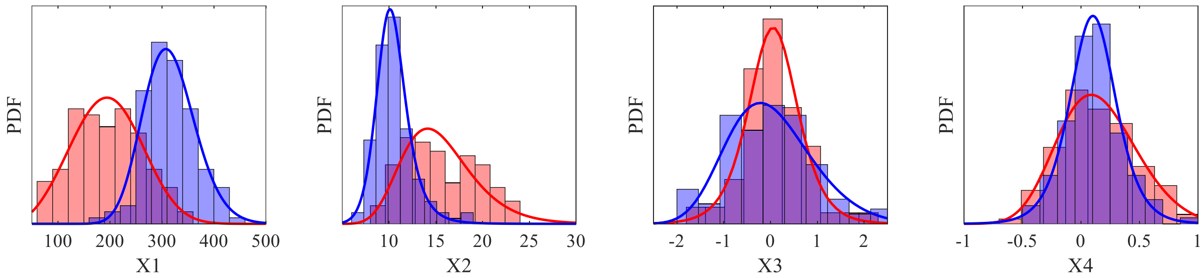


Figure 3. Empirical (*Bar Plot*) and Theoretical (*Line Plot*) Distribution of Each Variable in Dry (*Red*) and Rainy (*Blue*) Seasons.

Figure 3 shows that the distributions in the dry and rainy seasons are significantly different for variables X_1 (total precipitation) and X_2 (number of dry days), which shows that the season influences the conditions of these two variables. Meanwhile, the distribution of variables X_3 and X_4 is not significantly different in the dry and rainy seasons. This is because ENSO and IOD are annual phenomena, so they do not depend on the season in Kalimantan.

3.3. Bivariate Copulas

In our study, we employed a variety of copula functions to characterize the dependency structures between variables better. Copulas are powerful statistical tools that allow for the modeling of complex dependencies beyond linear relationships, particularly in multivariate distributions. Different copulas capture distinct aspects of the dependency structure, so we considered both one-parameter and two-parameter copulas. One-parameter copulas, such as the Clayton, Frank, Gumbel, Joe, and Galambos copulas, are more straightforward in their design, relying on a single parameter to capture the dependency between variables. Each of these copula's models dependency in unique ways. For example, the Clayton copula is particularly suited for capturing strong lower-tail dependence, while the Gumbel copula effectively describes strong upper-tail dependence [17]. These copulas provide a flexible framework for understanding different types of asymmetric dependencies that can arise in data.

On the other hand, two-parameter copulas, including the BB1, BB6, BB7, and BB8 copulas, offer an additional layer of complexity by introducing a second parameter. This allows them to capture more intricate dependency structures, including varying levels of both tail dependencies. For instance, the BB1 copula can simultaneously model both upper and lower tail dependencies with different intensities, providing a more nuanced representation of the underlying relationships between variables.

Utilizing one-parameter and two-parameter copulas ensured that a wide range of dependency patterns could be accurately captured, allowing for a more comprehensive and flexible data analysis. However, some copulas can only capture the positive dependencies between variables, while the dependencies between precipitation and hotspots used in this study are negative. Therefore, a rotated copula, which is 90°, 180°, and 270°, is necessary [18[, [19].

When copula is rotated 180°, the resulting copula is often called a survival copula. If Clayton copula can capture lower tail dependencies, then survival Clayton copula can capture upper tail dependencies. Moreover, negative dependencies can be captured using 90° and 270° rotated copulas [20]. The parameter for 90° and 270° rotated copula is negative of the corresponding copula parameter. **Figure 4** shows the dependency structure constructed from pairs of variables using bivariate copula. Some information is displayed such as the Kendall-tau correlation coefficient, the selected bivariate copula function and its parameters, and the contour plot of the joint probability density functions (joint PDFs) between pairs of variables.

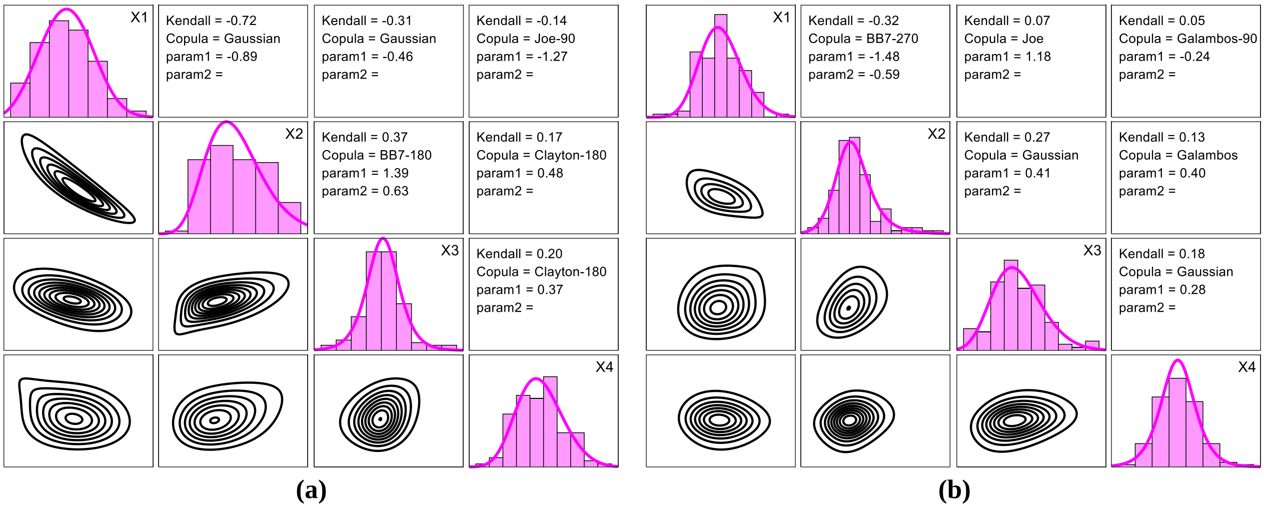


Figure 4. Dependency Structure Between Two Variables Using Bivariate Copula: (a) Dry and (b) Rainy Seasons

Figure 4 the highest dependency is in the relationship between variables X_1 and X_2 in the dry season. The Kendall-Tau correlation coefficient for this pair-variable is -72% which indicates a negative relationship between the two variables. Meanwhile, the most suitable copula function is the Gaussian copula with parameters: $\rho = -0.89$. Gaussian copulas are a family of elliptical copulas that able to access both positive and negative relationships between two or more variables. As a result, rotation on the Gaussian copula is not necessary. The variable relationship that applies rotation is between X_1 and X_4, where the selected copula function is the Joe copula with a rotation of 90 degrees (written Joe-90). Furthermore, some variable relationships are more suitable to the copula function rotated 180 degrees or the survival copula function, such as the relationship between variables X_2–X_3 and X_2–X_4. Based on **Figure 4**(a) and **Figure 4**(b), the dependencies between variables in the rainy season are weaker than in the dry season.

3.4. Nested Copula Structures

In this study, we simulate constructing 3- and 4-variate nested copulas to model the dependencies in climate data from Kalimantan. Using nested copulas allows us to capture complex multivariate dependencies, which is particularly valuable when analyzing environmental data, where interactions between variables are often nonlinear and exhibit varying degrees of correlation.

For the 3-variate case, we construct a fully nested 3-copula, where each pair of variables is linked through a hierarchy of copulas. This structure effectively models how the dependencies between the three climate variables evolve and interact across multiple layers. By nesting the copulas, we can account for dependencies that may not be captured by a single copula, enhancing the precision of our model. In the 4-variate case, we explore two different configurations: fully nested 4-copulas and partially nested 4-copulas.

The fully nested 4-copula follows a similar structure to the 3-variate case but with an additional variable, leading to a more intricate hierarchical structure. In this setup, all four climate variables are interconnected through layers of nested copulas, allowing us to model how their interdependencies propagate across different levels. In contrast, the partially nested 4-copula provides more flexibility by allowing only a subset of the variables to follow a nested structure. This means that while some variables may have a hierarchical relationship, others are modeled using simpler or direct copula connections. This approach is useful when not all variables exhibit the same level of dependence, enabling the model to accommodate varying strengths and forms of relationships between variables.

3.4.1. Nested 3-Copula

We construct fully nested 3-copulas for all combinations of the four variables used in the dry season. The construction steps for this fully nested 3-copula are as follows.

a. Estimate F_1, F_2, and F_3 (marginal distribution).

b. Select the two variables with the highest degree of dependence, for example X_1 and X_2, then transform the two variables using their respective marginal distributions (notate the results as U_1 and U_2).

c. Estimate C_1 using U_1 and U_2 (first bivariate copula).

d. Transform the remaining variable X_3 using its marginal distribution (notate the result as U_3).

e. Estimate C_2 using U_3 and $C_2(U_1, U_2)$ (second bivariate copula).

Using the selected marginal distributions and copula functions, the joint distribution and joint probability density functions can be calculated using **Equation (9)** and **Equation (10)**. **Table 2** shows the results of the nested 3-copula structures and selected copula functions for all combinations of the four variables used in the dry season.

Table 2. Nested 3-Copula Structures and Selected Copula Functions for All Combinations of Four Variables Used in the Dry Season

Variables	Structure	C1	C2	Goodness-of-fits
1,2,3	(1-2)-3	Gaussian Param1 = -0.89198	Galambos-180 Param1 = 0.16967	RMSE = 0.017 CvM = 0.036 pVal = 0.668
1,2,4	(1-2)-4	Gaussian Param1 = -0.89198	BB1-90 Param1 = 0 Param2 = -1	RMSE = 0.015 CvM = 0.026 pVal = 0.710
1,3,4	(1-3)-4	Gaussian Param1 = -0.46039	Frank Param1 = 0.19357	RMSE = 0.022 CvM = 0.059 pVal = 0.583
2,3,4	(2-3)-4	BB7-180 Param1 = 1.3906 Param2 = 0.62636	Clayton-180 Param1 = 0.37842	RMSE = 0.024 CvM = 0.069 pVal = 0.549

The brackets in the structure notation indicate the order in which pairs of variables are selected first. For example, the structure (1-2)-3 shows that the fitting process chooses the pair of variables X_1 and X_2 because this pair of variables has the highest level of dependence, as shown in **Figure 4**. After that, the copula results are based on X_1 and X_2, denoted C_1 (called inner copula), then paired with variable X_3 to produce copula C_2 (called outer copula). This also applies to other cases. Based on **Table 2**, the copula functions chosen to construct the trivariate joint distribution are various. This is obtained based on different patterns for each pair of variables. Moreover, goodness-of-fit tests show that the model (joint distribution) obtained fits the original data well. The joint distribution formed with a CvM p-value of more than 5% is acceptable for further analysis.

3.4.2. Nested 4-Copula

As in the previous subsection, we simulate the construction of a 4-dimensional copula (4-copula) with two approaches: fully and partially nested 4-copula. The general formulation can be seen in Figures 1.b and 1.c. Construction begins by selecting two variables with the highest degree of dependence (say x_1 and x_2), then transforming them into u_1 and u_2 and then the first bivariate copula C_1 can be obtained from them. After that, for fully nested, C_1 is sequentially paired with u_3 then u_4, i.e., the transformation results of x_3 and x_4 where x_3 has a dependency with C_1 that is greater than x_4. Meanwhile, for partially nested, C_1 is paired with the copula produced by u_3 and u_4. **Table 3** shows the results of the fully and partially nested 4-copula structures and selected copula functions of four climate variables in the dry season.

Table 3. Fully and Partially Nested 4-Copula Structures and Selected Copula Functions for Climate Data in Kalimantan in the Dry Season

Structure	C1	C2	C3	Goodness-of-fits
Fully ((1-2)-3)-4	Gaussian Param1 = -0.89198	Galambos-180 Param1 = 0.16967	Galambos Param1 = 0.17405	RMSE = 0.016 CvM = 0.029 pVal = 0.696
Partially (1-2)-(3-4)	Gaussian param1 = -0.8920	Clayton-180 param1 = 0.3732	Gumbel param1 = 1.5073	RMSE = 0.016 CvM = 0.030 pVal = 0.692

Table 4. Fully and Partially Nested 4-Copula Structures and Selected Copula Functions for Climate Data in Kalimantan in the Rainy Season

Structure	C1	C2	C3	Goodness-of-fits
Fully ((1-2)-3)-4	BB7-270 param1 = -1.4827 param2 = -0.58574	Clayton param1 = 0.093375	Clayton param1 = 0.038512	RMSE = 0.038 CvM = 0.175 pVal = 0.289
Partially (1-2)-(3-4)	BB7-270 param1 = -1.4827 param2 = -0.5857	Gaussian param1 = 0.2769	Gumbel param1 = 1.7449	RMSE = 0.022 CvM = 0.058 pVal = 0.583

The bracket notation in the structure indicates the sequence in which pairs of variables are selected. For instance, the structure ((1−2)−3)−4 signifies that the variables X_1 and X_2 are chosen first, as this pair exhibits the strongest dependence, as illustrated in **Figure 4**. Following this, the copula formed between X_1 and X_2 , denoted as C_1, is then paired with the variable X_3, which has the highest dependence with C_1, resulting in the second copula C_2. Finally, C_2 is combined with the remaining variable X_4, producing the third copula, C_3. This process is similarly applied to other cases. As presented in **Table 3** and **Table 4**, it is evident that different copula functions are selected to construct the 4-variate joint distribution, reflecting the varying dependence structures among the pairs of variables. The selection of these copula functions is based on distinct dependence patterns observed for each pair of variables.

Moreover, goodness-of-fit tests reveal that the resulting joint distribution model competes strongly with the original data. The joint distribution is deemed acceptable for further analysis with a Cramér-von Mises (CvM) p-value well above the 5% threshold. This indicates that the constructed copula model accurately represents the multivariate dependence structure, allowing for reliable inferences in subsequent stages of the analysis. Copulas provide an effective and efficient way to build multivariate joint distributions.

3.5. Discussion

In this study, we have employed the copula methodology to analyze the joint distribution of local climate factors (total precipitation and number of dry days) and global climate indices (ENSO and IOD) in Kalimantan. We captured the complex dependence structure between these variables in dry and rainy seasons using fully and partially nested copula models. The results from bivariate copula analysis reveal significant negative dependencies between the local climate factors, particularly in the dry season. For instance, the strong negative relationship between total precipitation and the number of dry days indicates that as dry days increase, precipitation decreases. This is consistent with the overall dry conditions experienced in Kalimantan during the dry season. Interestingly, the dependencies involving global climate indices (ENSO and IOD) were weaker in the rainy season compared to the dry season. This suggests that the influence of global phenomena like ENSO and IOD on local climate conditions in Kalimantan is more pronounced during periods of drought.

The nested copula structures provided further insights into the multivariate dependencies. The fully nested 3- and 4-copula models were effective in capturing the joint behavior of the local and global climate factors. The selection of various copula functions (e.g., Gaussian, Clayton, Frank) for different pairs of variables reflects the diverse dependence structures present in the data. Additionally, the goodness-of-fit tests confirmed that the copula models provide an acceptable fit to the data, making them suitable for further analysis, such as predicting drought events or forest fires.

4. CONCLUSION

This study demonstrates the effectiveness of copula methods to model the complex dependencies between local and global climate factors in Kalimantan. By employing both fully and partially nested copula structures, we successfully captured the intricate relationships between precipitation, the number of dry days, and global climate indices such as ENSO and IOD. The results highlight the global phenomena, particularly ENSO and IOD, have a more substantial impact during the dry season, influencing drought conditions in Kalimantan. Additionally, the copula method offers simplicity and flexibility in constructing multivariate joint distributions, allowing for a more detailed and accurate representation of dependencies than traditional linear models. This provides a significant advantage in modeling complex climate interactions and predicting extreme events, such as droughts or forest fires, based on the interplay between local and global climate factors.

REFERENCES

[1] S. Nurdiati *et al.*, “THE IMPACT OF EL NIÑO SOUTHERN OSCILLATION AND INDIAN OCEAN DIPOLE ON THE BURNED AREA IN INDONESIA,” *Terr. Atmos. Ocean. Sci.*, vol. 33, no. 15, 2022, doi: 10.1007/S44195-022-00016-0.

[2] X. Pan, M. Chin, C. M. Ichoku, and R. D. Field, “CONNECTING INDONESIAN FIRES AND DROUGHT WITH THE TYPE OF EL NIÑO AND PHASE OF THE INDIAN OCEAN DIPOLE DURING 1979–2016,” *J. Geophys. Res. Atmos.*, vol. 123, no. 15, pp. 7974–7988, 2018, doi: 10.1029/2018JD028402.

[3] S. Nurdiati, F. Bukhari, M. T. Julianto, M. K. Najib, and N. Nazria, “HETEROGENEOUS CORRELATION MAP BETWEEN ESTIMATED ENSO AND IOD FROM ERA5 AND HOTSPOT IN INDONESIA,” *Jambura Geosci. Rev.*, vol. 3, no. 2, pp. 65–72, 2021, doi: 10.34312/jgeosrev.v3i2.10443.

[4] A. M. Hidayat, U. Efendi, L. Agustina, and P. A. Winarso, “KORELASI INDEKS NINO 3.4 DAN SOUTHERN OSCILLATION INDEX (SOI) DENGAN VARIASI CURAH HUJAN DI SEMARANG,” *J. Sains Teknol. Modif. Cuaca*, vol. 19, no. 2, p. 75, 2018, doi: 10.29122/jstmc.v19i2.3143.

[5] I. Narulita, R. Rahayu, E. Kusratmoko, S. Supriatna, and M. Djuwansah, “ANCAMAN KEKERINGAN METEOROLOGIS DI PULAU KECIL TROPIS AKIBAT PENGARUH EL-NINO DAN INDIAN OCEAN DIPOLE (IOD) POSITIF, STUDI KASUS: PULAU BINTAN,” *J. Lingkung. dan Bencana Geol.*, vol. 10, no. 3, p. 127, 2020, doi: 10.34126/jlbg.v10i3.252.

[6] G. Y. I. Ryadi, A. Sukmono, and B. Sasmito, “PENGARUH FENOMENA EL-NINO DAN LA-NINA PADA PERSEBARAN CURAH HUJAN DAN TINGKAT KEKERINGAN LAHAN DI PULAU BALI,” *J. Geod. Undip*, vol. 8, no. 4, pp. 41–49, 2019.

[7] M. K. Najib, S. Nurdiati, and A. Sopaheluwakan, “COPULA-BASED JOINT DISTRIBUTION ANALYSIS OF THE ENSO EFFECT ON THE DROUGHT INDICATORS OVER BORNEO FIRE-PRONE AREAS,” *Model. Earth Syst. Environ.*, vol. 8, no. 2, pp. 2817–2826, 2022, doi: 10.1007/s40808-021-01267-5.

[8] P. R. Dewick and S. Liu, “COPULA MODELLING TO ANALYSE FINANCIAL DATA,” *J. Risk Financ. Manag.*, vol. 15, no. 3, 2022, doi: 10.3390/jrfm15030104.

[9] I. Ismail, F. N. N. Abd Mutalip, and K. Jacob, “A COMPREHENSIVE REVIEW ON THE DEVELOPMENT OF COPULAS IN FINANCIAL FIELD,” *J. Intell. Fuzzy Syst.*, vol. 45, no. 4, pp. 6047–6062, 2023, doi: 10.3233/JIFS-223481.

[10] B. Rašiová and P. Árendáš, “COPULA APPROACH TO MARKET VOLATILITY AND TECHNOLOGY STOCKS DEPENDENCE,” *Financ. Res. Lett.*, vol. 52, 2023, doi: 10.1016/j.frl.2022.103553.

[11] V. Gumus, Y. Avsaroglu, O. Simsek, and A. Basak, “EVALUATING THE DURATION, SEVERITY, AND PEAK OF HYDROLOGICAL DROUGHT USING COPULA,” *Theor. Appl. Climatol.*, vol. 152, no. 3–4, pp. 1159–1174, 2023, doi: 10.1007/s00704-023-04445-w.

[12] G. Geenens, “COPULA MODELING FOR DISCRETE RANDOM VECTORS,” *Depend. Model.*, vol. 8, no. 1, pp. 417–440, 2020, doi: 10.1515/demo-2020-0022.

[13] G. Geenens, “(Re-)Reading Sklar (1959)—A PERSONAL VIEW ON SKLAR’S THEOREM,” *Mathematics*, vol. 12, no. 3, 2024, doi: 10.3390/math12030380.

[14] M. K. Najib, S. Nurdiati, and A. Sopaheluwakan, “QUANTIFYING THE JOINT DISTRIBUTION OF DROUGHT INDICATORS IN BORNEO FIRE-PRONE AREA,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 880, no. 1, 2021, doi: 10.1088/1755-1315/880/1/012002.

[15] X. Wei, H. Zhang, V. P. Singh, C. Dang, S. Shao, and Y. Wu, “COINCIDENCE PROBABILITY OF STREAMFLOW IN WATER RESOURCES AREA, WATER RECEIVING AREA AND IMPACTED AREA: IMPLICATIONS FOR WATER SUPPLY RISK AND POTENTIAL IMPACT OF WATER TRANSFER,” *Hydrol. Res.*, vol. 51, no. 5, pp. 1120–1135, 2020, doi: 10.2166/nh.2020.106.

[16] 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.

[17] Z. Li, Q. Shao, Q. Tian, and L. Zhang, “COPULA-BASED DROUGHT SEVERITY-AREA-FREQUENCY CURVE AND ITS UNCERTAINTY, A CASE STUDY OF HEIHE RIVER BASIN, CHINA,” *Hydrol. Res.*, vol. 51, no. 5, pp. 867–881, 2020, doi: 10.2166/nh.2020.173.

[18] A. Wiboonpongse, J. Liu, S. Sriboonchitta, and T. Denoeux, “MODELING DEPENDENCE BETWEEN ERROR COMPONENTS OF THE STOCHASTIC FRONTIER MODEL USING COPULA: APPLICATION TO INTERCROP COFFEE PRODUCTION IN NORTHERN THAILAND,” *Int. J. Approx. Reason.*, vol. 65, pp. 34–44, 2015, doi: 10.1016/j.ijar.2015.04.001.

[19] S. Sriboonchitta, J. Liu, A. Wiboonpongse, and T. Denoeux, “A DOUBLE-COPULA STOCHASTIC FRONTIER MODEL WITH DEPENDENT ERROR COMPONENTS AND CORRECTION FOR SAMPLE SELECTION,” *Int. J. Approx. Reason.*, vol. 80, pp. 174–184, 2017, doi: 10.1016/j.ijar.2016.08.006.

[20] J. Liu, D. Sirikanchanarak, S. Sriboonchitta, and J. Xie, “ANALYSIS OF HOUSEHOLD CONSUMPTION BEHAVIOR AND INDEBTED SELF-SELECTION EFFECTS: CASE STUDY OF THAILAND,” *Math. Probl. Eng.*, 2018, doi: 10.1155/2018/5486185.