Joint Distribution And Coincidence Probability Of The Number Of Dry Days And The Total Amount Of Precipitation In Southern Sumatra Fire-Prone Area

Sri NURDIATI [1], ***Mohamad Khoirun NAJIB*** [1], ***and Achmad Syarief THALIB*** [1]

DOI: 10.21163/GT_2022.172.10

ABSTRACT:
El Niño Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) can affect the increase in rainfall intensity and the number of dry days, also known as dry spells that can cause drought and increase the potential for forest fires. This study examines the effect of ENSO and IOD conditions on the joint distribution of the number of dry days and total precipitation in a fire-prone area in southern Sumatra, Indonesia. The joint distribution is constructed using rotated copulas from several families, including Gaussian, student's t, Clayton, Gumbel, Frank, Joe, Galambos, BB1, BB6, BB7, and BB8. Fire-prone areas are defined using k-mean clustering, while the copula parameters are estimated using the inference of function for margins (IFM) method. Based on the peak of joint probability density functions (PDFs), ENSO and IOD conditions had a significant effect in the dry season but had no significant effect in the rainy season. The peak of joint PDFs is getting to the dry-dry conditions when the ENSO and IOD indexes increase in the dry season. However, based on coincidence probability, ENSO conditions still influence the joint distribution between the number of dry days and total precipitation during the rainy season but not with IOD conditions. The lower the ENSO index, the higher the probability of wet conditions co-occurring in the number of dry days and total precipitation. Meanwhile, ENSO and IOD conditions significantly affect the coincidence probability between the number of dry days and total precipitation. Moderate-Strong El Niño has the most considerable coincidence probability of 68.5%, followed by Positive IOD with 62.6%. The two conditions had similar effects on the joint distribution of the number of dry days and total precipitation. Moreover, the association between the number of dry days and the total precipitation was stronger in the dry season than in the rainy season.

Key-words: *Bivariate copula, Exceedance probability, Rainfall, Risk assessment, Wildfire.*

1. INTRODUCTION

Indonesia - an archipelagic country on the equator flanked by two continents (Asia and Australia) and two oceans (Indian and Pacific Oceans)- has varied climatic conditions. These climatic conditions are influenced by the surrounding oceans, locally and globally. The high precipitation amount in almost all provinces is made possible by an ample supply of water, which comes from the surrounding sea (Nuryanto & Badriyah, 2014). Therefore, local sea conditions will affect the climatic conditions of the surrounding islands. Besides local sea conditions, climatic conditions are also influenced by global ocean conditions around Indonesia, such as the El Niño Southern Oscillation (ENSO) in the Pacific Ocean and the Indian Ocean Dipole (IOD) in the Indian Ocean (Kurniadi et al., 2021).

ENSO is an ocean-atmosphere interaction centered in the equatorial Pacific Ocean that has a global impact on Earth's system (McPhaden et al., 2006). ENSO causes Indonesian seas to be colder during El Niño events and warmer during La Niña events, resulting in increased precipitation during La Niña and decreased precipitation during El Niño (Nur'utami & Hidayat, 2016). Meanwhile, IOD

[1]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, IPB University, Bogor, Indonesia,*
**Corresponding author: nurdiati@apps.ipb.ac.id; mkhoirun_najib@apps.ipb.ac.id; achmad_st@apps.ipb.ac.id*

is a phenomenon indicated by differences in sea surface temperature anomalies in the western and eastern parts of the Indian Ocean.

The positive phase of IOD causes a decrease in sea surface temperature in Indonesia, the same as El Niño, accompanied by a reduction in precipitation (Khaldun et al., 2018). Contrary, the negative phase of IOD is like La Niña conditions in the Pacific Ocean. ENSO and IOD phases in each year are not always the same. However, the impact will be more significant if ENSO and IOD are simultaneously in a warm (or cold) phase. For example, the droughts in 1997 and 2015 coincided with ENSO being in an El Niño phase and IOD being in a positive phase (Avia & Sofiati, 2018).

El Niño and positive IOD, which reduce precipitation, indirectly contributed significantly to drought occurrence in several regions in Indonesia, although with different strengths in each area (Ummenhofer et al., 2013). Drought reduces the water content of vegetation, loss of water content in extensive woods, leading to plant death, and increased fire potential (Suharjo & Velicia, 2018). Therefore, precipitation is urgently needed so vegetation's water content can remain within normal limits. The prolonged drought in 1997 caused by the late rainy season due to El Niño and positive IOD climate variability of 1997/1998 has triggered widespread forest fires in the equatorial regions of Sumatra and Kalimantan (Nikonovas et al., 2022).

Previous research has found that forest fires in Kalimantan are more sensitive to ENSO than IOD. Meanwhile, the effect of IOD on forest fires was more pronounced in the southern part of Sumatra because it is located near the Indian Ocean (Nurdiati et al., 2021). Moreover, although it has been mentioned that the lack of precipitation can be a triggering factor for forest fires, the total precipitation is insufficient to characterize forest fires in Indonesia. Other indicators such as the number of dry days, i.e., the number of days with less than 1 mm of precipitation per day (Brown et al., 2010), can better characterize forest fires than total precipitation. Although many experts have investigated the effect of ENSO and IOD on precipitation in Indonesia (Lestari et al., 2018), there has been limited research examining the joint distribution of the number of dry days and total precipitation under different ENSO and IOD phase conditions.

The copula function is the most common method for constructing a joint distribution. The copula is a statistical method that can describe the relationship between variables that is not too tight on distribution assumptions and can clearly show dependency relationships at extreme points (Singh et al., 2020). This method can describe the dependency structure between variables with different marginals and model their tail dependencies (Caraka et al., 2016). Copula has been widely used to address uncertainties in coincidence probability (Yang et al., 2020), conditional probability (H. W. Li et al., 2021), and joint return period analysis (Naeini et al., 2021). However, according to a systematic literature review, the copula is still limited to being applied in forest fire analysis (Najib et al., 2021).

According to the background presented, this study analyzes the joint distribution of the number of dry days and total precipitation in fire-prone areas in southern Sumatra using the copula function. This study begins by identifying fire-prone areas in Sumatra using k-mean clustering based on the burned area dataset. In the specified fire-prone areas, the number of dry days and total precipitation were extracted from 1981 to 2020. After that, the copula function was employed to construct the joint distribution of the number of dry days and total precipitation. The joint distribution is constructed based on the seasonality and phases of ENSO and IOD to analyze the effect of these conditions on the dependencies and coincidence probability. The study results are expected to improve understanding of the dependencies between the variables studied and become an initial provision for further analysis of extreme events such as floods, droughts, and forest fires in the study area.

2. STUDY AREA AND DATASETS

This study focused on southern Sumatra, located at 98°E-107.5°E and 1.5°S-6°S (**Fig. 1**). This study will use four types of data: precipitation, Niño 3.4 index, dipole mode index, and burned area datasets. Precipitation data was derived from Climate Hazards Group Infra-Red Precipitation with Station data (CHIRPS) with a spatial resolution of 0.25°×0.25° in daily and monthly temporal resolutions from 1981 to 2020. Total precipitation data were taken from monthly data, while dry spell

data were extracted using daily data. Moreover, The Niño 3.4 and dipole mode index is time-series data obtained from the average of Hadley Center Sea Ice and Sea Surface Temperature (HadISST1) data. The data was downloaded from the Physical Sciences Laboratory (PSL) of the National Oceanic and Atmospheric Administration (NOAA) from 1981 to 2020. Both are using 1981-2010 as an anomaly in the index calculation. The burned area data uses the fourth generation of the Global Fire Emission Database (GFED4.1s) from 1997-2016 to define fire-prone areas. The GFED4.1s dataset provides monthly global burned area data with a spatial resolution of 0.25°×0.25° in Hierarchical Data Format version 5 (HDF5).

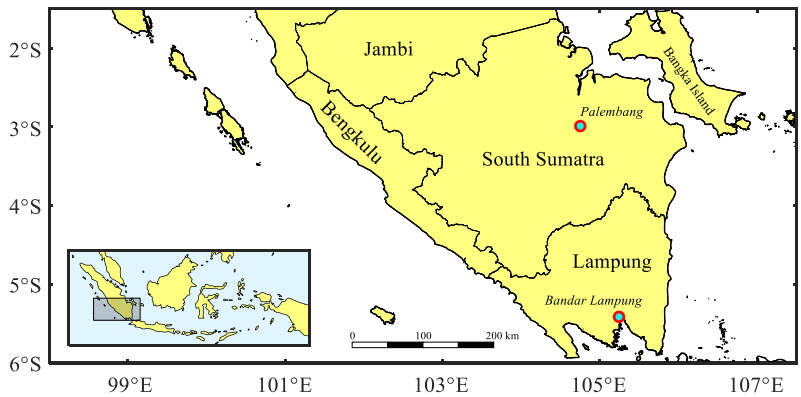

Fig. 1. Location of the southern part of Sumatra on a map of Indonesia

3. METHODS

3.1. k-Means Clustering

This study begins by collecting data from several sources mentioned (*data integration*), then defining fire-prone areas in southern Sumatra using k-mean clustering. After that, the number of dry days and total precipitation are extracted and aggregated (*data reduction*). Thus, the time series of the number of dry days and total precipitation are obtained, then partitioned based on seasonality, ENSO, and IOD conditions.

K-means clustering is the most common and popular unsupervised algorithm for clustering. The k-means objective can be written in a minimization problem given by

$$\underset{\mu_j}{\operatorname{argmin}} \sum_{j=1}^{k} \sum_{x_j \in D_j'} \|x_j - \mu_j\|^2 \tag{1}$$

where μ_j represent the mean of the j-th cluster and D_j' denotes the subdomain of the dataset associated with the j-th cluster (Brunton & Kutz, 2022). We use standardization before clustering to provide better, more efficient, and more accurate results (Mohamad & Usman, 2013).

3.2. Correlation Coefficient

Correlation analysis measures the strength of the relationship between two variables, called a correlation coefficient. Three popular correlation coefficients are Pearson (r), Spearman (ρ), and Kendall (τ) correlation coefficients (Hauke & Kossowski, 2011). The existence of a correlation between variables is a sufficient condition for copula-based analysis. Because the copula used mostly has a relationship with the Kendall correlation, this study uses the Kendall correlation coefficient to examine the relationship between variables.

3.3. Copula-Based Joint Distribution

The copula is a function that *couples* multivariate distribution functions with one-dimensional marginal distribution functions. In another sense, the copula is a multivariate distribution function whose one-dimensional margin is uniform at the interval (0,1). The copula function can describe dependencies between variables, which is an important step in connecting one climate information to another.

The advantage of the copula function is that each variable is joined by its marginal distribution instead of its original value. Unlike other multivariate distributions, the copula function is flexible in the type of marginal distribution used and does not require input variables to share a similar distribution. Let H be a joint distribution function with F and G as edges, then there is a copula C such that for every $x, y \in R$, so that

$$H(x,y) = C\big(F(x), G(y)\big) \tag{2}$$

If F and G are continuous cumulative distribution functions (CDFs), then C is unique (Nelsen, 2006). Assuming the variables are continuously distributed, then C is unique and the joint density function $h(x,y)$ can be written as

$$h(x,y) = c\big(F(x), G(y)\big) \cdot f(x) \cdot g(y) \tag{3}$$

where

$$c\big(F(x), G(y)\big) = \frac{\partial^2}{\partial F(x)\, \partial G(y)} C\big(F(x), G(y)\big) \tag{4}$$

where f and g are continuous probability density functions (PDFs) and c is a copula density.

3.3.1. Parameters Estimation

This study uses the inference of function for margin (IFM) method to estimate the copula parameters. The first step of this method is to estimate the distribution function used as edges for the copula. There are three continuous distribution functions used, namely generalized extreme value (GEV), normal (N), and lognormal (LN). The fittest distribution function was selected and tested using the Anderson-Darling hypothesis test. After obtaining the fittest distribution function, the second step of the IFM method is estimating the copula parameters. Different copulas have their own characteristics in describing the dependency structure of a pair of variables. There are 11 copulas considered to investigate the dependency structure between the two variables, including Gaussian, Student t, Clayton, Gumbel, Frank, Joe, Galambos, BB1, BB6, BB7, and BB8. The fittest copula function was selected and tested using the Cramer-von Mises hypothesis test (Najib et al., 2022b). Moreover, the first and second steps of the IFM method aim to find the parameter values by maximizing the log of the likelihood of each function. Code for fitting parameter values of univariate distribution and copula functions is available on the GitHub page: https://github.com/mkhoirun-najiboi/mycopula.

3.3.2. Coincidence Probability

By dividing the three parts of each edge, each variable's dry, normal, and wet conditions can be determined. Coincidence probability is the probability of dry, normal, and wet conditions of two variables simultaneously, whether the conditions are the same or not (Najib et al., 2022a). It is called synchronous probability if the two conditions are the same (wet-wet, normal-normal, and dry-dry). Conversely, if the conditions are different (wet-normal, dry-wet, etc.), then it is called asynchronous probability (**Fig. 2**).

ww	**nw**	**dw**
wn	**nn**	**dn**
wd	**nd**	**dd**

Fig. 2. Illustration of the coincidence probability of the number of dry days (x) and total precipitation (y): synchronous probability with blue shading and otherwise asynchronous probability. The subscripts w, n and d mean wet, normal and dry conditions.

4. RESULTS AND DISCUSSION

4.1. Southern Sumatra Fire-Prone Areas

The definition of fire-prone areas using k-mean clustering is done using several choices of the number of clusters. Based on several experiments, clustering with 8 clusters gives the lowest RMSE value compared to the others **Fig. 3a**. The RMSE in question is the difference between the time series on the data burned area in the fire-prone area and the data burned area in the study area. Fire-prone areas are defined by removing clusters that have low hotspot characteristics. Based on **Fig. 3b**, clusters A and C from the 8-mean clustering results have low hotspot characteristics.

The results of clustering on the burned area data using 8-means clustering are shown in **Fig. 3c**. Each color represents eight different clusters named with the letters A to H to avoid misunderstanding if given a name based on numbers. Fire-prone areas are defined by eliminating clusters with low burning area characteristics. **Fig. 3d** shows the maximum value of the burn fraction in a month that has occurred in an area of 0.25°×0.25°. The fire-prone areas are focused on three districts in South Sumatra province: Ogan Komering Ilir, Banyuasin, and Musi Banyuasin. The most considerable burn fraction reached 78.6% in the Ogan Komering Ilir district in October 1997. Based on **Fig. 3b**, the clusters assessed as not prone to fire are clustered C and A, so the other clusters are designated fire-prone areas.

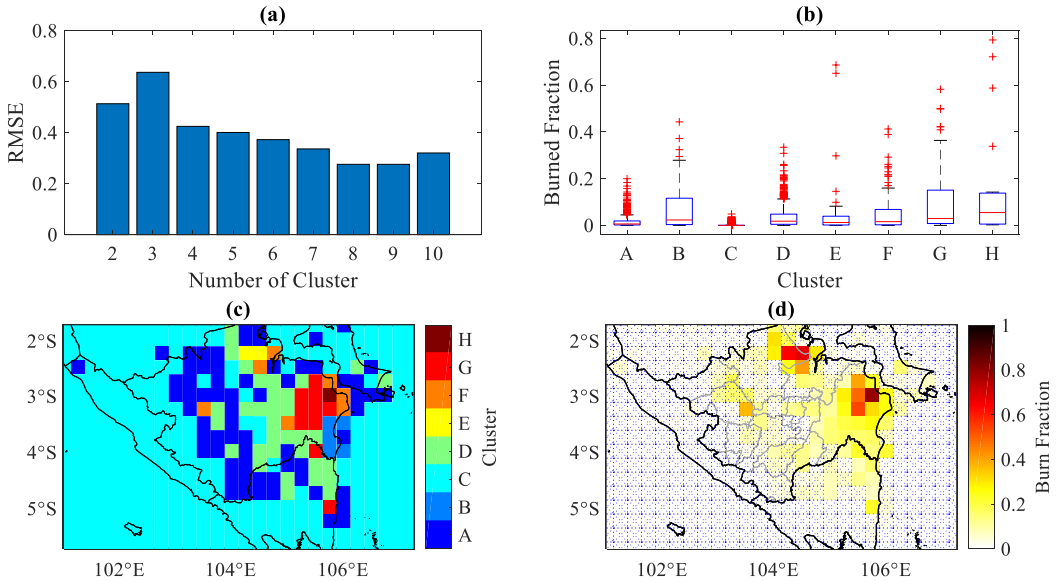

Fig. 3. Definition of fire-prone areas in southern Sumatra: (a) RMSE value from the results of clustering from several choices of the number of clusters, (b) 8-means clustering results (cluster characteristics), (c) 8-means clustering results (cluster location) and (d) maximum burn fraction value for each area.

The area covered by the blue dotted box is the area removed from the fire-prone area. The burned area, precipitation, and the number of dry days data are aggregated to obtain the time series of each data using this fire-prone area.

4.2. Correlation Analysis

This section determines which of the cumulative monthly, 2-monthly, and 3-monthly number of dry days and total precipitation data provides the highest correlation to the burned area data. **Table 1** shows the Kendall correlation value for each condition. In the total rainfall data, the Kendall correlation coefficient shows that the 2-month cumulative variable provides the highest correlation. Meanwhile, the 3-month cumulative variable gives the highest correlation on the number of dry days. Therefore, 2-month and 3-month cumulative variables were selected for total precipitation and the number of dry days data, respectively.

Table 1.

Pearson, Spearman, and Kendall correlation values between climate indicators (the number of dry days and total precipitation) and burned area.

Pair variables	**Cumulative Variable**		
	1-month	*2-month*	*3-month*
Number of dry days and burned area	58.8%	67.0%	**67.3**%
Total precipitation and burned area	54.6%	**63.6**%	57.2%

The existence of sufficient correlation is a requirement for using copula-based analysis. The copula parameter has a close relationship with the Kendall-τ correlation coefficient. Therefore, the Kendall-τ correlation coefficient between the variables that have been partitioned by season and phase of ENSO and IOD needs to be considered (**Table 2**).

There are two seasons to consider, i.e., the rainy (Nov-Apr) and dry (May-Oct) seasons. Meanwhile, four ENSO conditions were considered: La Niña (Nino Index ≤ -0.5), Normal ($-0.5 <$ Nino Index < 0.5), weak El Niño ($0.5 \leq$ Nino Index < 1), and moderate-strong El Niño (Nino Index ≥ 1, for simplicity, will be called strong El Niño). Furthermore, the three IOD conditions consisted of Negative IOD (DMI ≤ -0.4), Neutral IOD (to differentiate from "Normal" terms in ENSO conditions, $-0.4 <$ DMI < 0.4), and Positive IOD (DMI ≥ 0.4).

Table 2.

Kendall-τ correlation coefficient between the number of dry days and total precipitation under different ENSO and IOD conditions.

Season	**ENSO Conditions**				**IOD Conditions**		
	La Niña	***Normal***	***Weak El Niño***	***Strong El Niño***	***Negative***	***Neutral***	***Positive***
Rainy	-33.4%	-31.6%	-31.2%	-48.8%	-30.8%	-35.8%	-51.4%
Dry	-59.2%	-65.1%	-76.8%	-81.9%	-35.0%	-66.8%	-84.2%

The Kendall-τ correlation coefficient between the number of dry days and total precipitation has a negative relationship, with the lowest correlation being -30.8% and the highest being -84.2%. The strength of the correlation of each pair-variable ranged from moderate to very strong. Thus, all pair variables can be analyzed using the copula function.

4.3. Parameters Estimation

4.3.1. Distribution Functions

The first step in copula-based joint distribution analysis using the IFM method is to estimate the univariate distribution function of each edge. **Tables 3** and **4** show the fittest distribution functions of the number of dry days and total precipitation in the rainy and dry seasons, respectively. The Anderson-Darling hypothesis test was employed to test the suitability of the distribution. The null hypothesis of this test is that the data comes from a population spread on the selected distribution function. With a *p*-value of more than the 5% significance level, each of the selected distributions was

in accordance with the data because the hypothesis test failed to reject the null hypothesis. Thus, each selected distribution function can be used as an edge for constructing the copula function.

Table 3.
Fittest distribution functions (i.e., Generalized Extreme Value/GEV, Normal, or Log Normal/LN) with the *p*-value of the Anderson-Darling hypothesis test (rainy season).

	ENSO Conditions				**IOD Conditions**		
	La Niña	***Normal***	***Weak El Niño***	***Strong El Niño***	***Negative***	***Neutral***	***Positive***
Number of dry days	GEV (81.7%)	GEV (94.8%)	GEV (85.0%)	GEV (59.7%)	Normal (96.1%)	GEV (73.2%)	GEV (53.4%)
Total Precipitation	LN (74.8%)	LN (84.7%)	GEV (86.7%)	GEV (78.9%)	Normal (88.5%)	Normal (47.2%)	GEV (60.0%)

Table 4.
Fittest distribution functions (i.e., Generalized Extreme Value/GEV, Normal, or Log Normal/LN) with the *p*-value of the Anderson-Darling hypothesis test (dry season).

	ENSO Conditions				**IOD Conditions**		
	La Niña	***Normal***	***Weak El Niño***	***Strong El Niño***	***Negative***	***Neutral***	***Positive***
Number of dry days	GEV (86.3%)	GEV (64.2%)	GEV (60.7%)	Normal (27.9%)	LN (71.2%)	GEV (52.1%)	Normal (8.7%)
Total Precipitation	Normal (99.4%)	Normal (99.1%)	GEV (68.9%)	GEV (96.6%)	GEV (99.9%)	GEV (95.3%)	LN (77.2%)

4.3.2. Copula Functions

Some copula functions, such as Clayton, Gumbel, and Joe, can only be constructed on variables that have a positive relationship. By reconsidering **Table 2** that the relationship between the two variables is negative, a rotated copula is needed to obtain a copula function with a negative relationship. Three forms of rotation will be applied, namely 90, 180, and 270 degrees:

$$C_{90^\circ}(u, v|\theta) = v - C(1-u, v| - \theta) \tag{5}$$

$$C_{180^\circ}(u, v|\theta) = u + v - 1 + C(1-u, 1-v|\theta) \tag{6}$$

$$C_{270^\circ}(u, v|\theta) = u - C(u, 1-v| - \theta) \tag{7}$$

Table 5 shows the fittest copula function with the p-value of the Cramer-von Mises hypothesis test. With a p-value of more than 5% significance level, each of the selected copula functions was appropriate because the hypothesis test failed to reject the null hypothesis. Thus, each chosen copula function can be used to construct the joint distribution between the number of dry days and total precipitation in each condition.

Table 5.
Fittest copula functions with the p-value of the Cramer-von Mises hypothesis test.

Season	**ENSO Conditions**				**IOD Conditions**		
	La Niña	***Normal***	***Weak El Niño***	***Strong El Niño***	***Negative***	***Neutral***	***Positive***
Rainy	Joe-90° (53.6%)	BB7-270° (63.6%)	Gumbel-270° (60.3%)	BB8-270° (59.2%)	Clayton° (76.6%)	Galambos-90° (37.3%)	BB8-270° (43.4%)
Dry	Joe-270° (70.6%)	BB1-90° (59.9%)	Gumbel-90° (72.4%)	Frank-270° (64.3%)	Joe-270° (77.5%)	BB8-270° (60.8%)	Frank° (35.4%)

With a variety of copula functions that have been developed and widely used, various characteristics of the relationship between variables can be handled by the copula (Z. Li et al., 2020). As in **Table 5**, each copula function formed for each condition varies from Clayton to BB7 copulas. The various functions of the copula can cope with different relationship behaviors, such as joint behavior in extreme or tail regions.

4.4. Joint Density Functions

After selecting the distribution and copula functions that can represent the data, the joint density function is calculated using Eq. 4, then visualized using a contour plot for each ENSO (**Fig. 4**) and IOD (**Fig. 5**) conditions. Contour levels have been equalized for each condition so that each can be compared with the others. In each plot, nine square areas show each combination of coincidence probability areas.

Based on **Fig. 4,** each ENSO condition in the rainy season has no significant effect on the joint distribution of the number of dry days and total precipitation. This can be seen from the dashed line, which is the peak of the joint probability density functions (PDFs). The peak of joint PDFs in the rainy season does not experience a significant shift and is around normal to wet, with dry days of about 20 days per 3-month and total precipitation of around 600 mm per 2-month. However, some values become outliers and lie in dry conditions during Weak El Niño and Strong El Niño. This outlier occurred in November and December, indicating that there was a delay in the arrival of the rainy season or the prolongation of the dry season during El Niño.

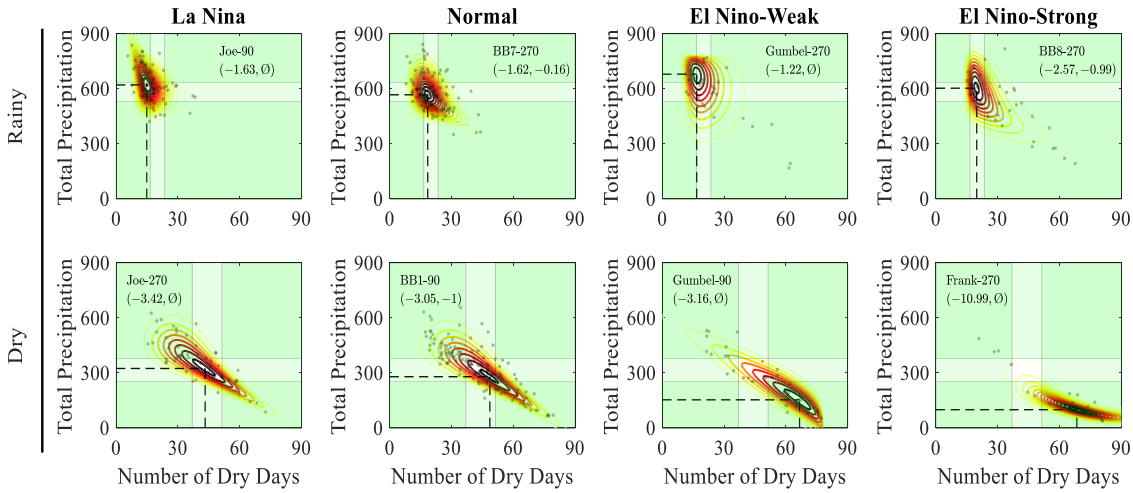

Fig. 4. Contour plot of joint density function for each ENSO condition in the rainy and dry seasons.

Meanwhile, the opposite happened for each ENSO condition in the dry season. The effect of the ENSO condition is discernible from the peak of joint PDFs movement in each condition. The peak of joint PDFs is getting to the dry-dry area when the ENSO index increases. The strong El Niño condition with the highest ENSO index causes the peak of joint PDF to be at the driest state compared to other ENSO conditions in dry seasons, with rainfall intensity of 99 mm in 2-month and the number of dry days of as much as 68.4 days in 3-month.

As before, based on **Fig. 5**, each IOD condition in the rainy season does not significantly affect the joint distribution of the number of dry days and precipitation. In the dry season, it is seen that there is an influence of IOD conditions which makes the peak of joint PDFs of each condition increasingly enters the dry-dry area as the IOD index increases. Positive IOD conditions with the highest IOD index cause the peak of joint PDF to be in the driest state. The peak of joint PDF has a rainfall intensity of 121.5 mm in 2-month and the number of dry days of as much as 65.3 days in three months.

Based on **Figs. 4** and **5** in the dry season, Positive IOD conditions and Strong El Niño have similar effects on the joint PDFs of the number of dry days and total precipitation. Both PDF joints are relatively in dry-dry condition with little probability in normal-dry condition. In contrast, the joint PDFs during Weak El Niño conditions still touch normal-normal conditions, although the peak of the joint PDF is in dry-dry conditions.

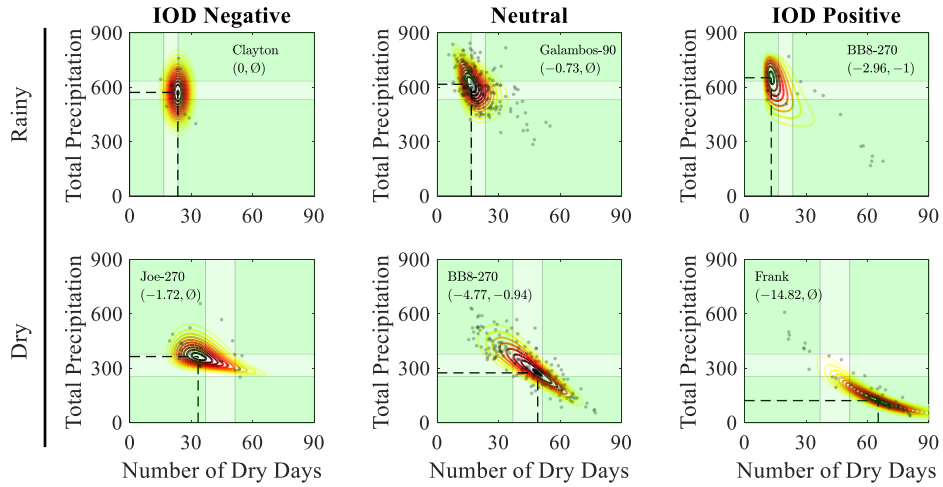

Fig. 5. Contour plot of joint density function for each IOD condition in the rainy and dry seasons.

4.5. Coincidence Probability

Using a suitable copula function, a joint distribution function between climate indicators can be constructed and the joint probability between indicators in dry and wet conditions can be estimated. This probability is often referred to as coincidence or exceedance probability (Wei et al., 2020). Coincidence probability is used to observe the frequency of co-occurrence between the number of dry days and total precipitation under dry, normal, and wet conditions. The values of the synchronous probability are presented in **Tables 6** and **7** for the rainy and dry seasons, respectively. Both also display the total synchronous and asynchronous probability.

Table 6 shows the effect of ENSO on increasing the probability of wet-wet conditions and decreasing the probability of dry-dry conditions as the ENSO index decreases in the rainy season. The lower the ENSO index, the higher the probability of wet conditions co-occurring in the number of dry days and total precipitation. During La Niña, the probability of wet conditions simultaneously is 30.3%. These results indicate that although the ENSO phenomenon does not show a significant effect from the peak of joint PDFs, the effect is clearly seen in the coincidence probability of the number of dry days and total precipitation.

Table 6.

Coincidence probability of the joint PDFs of the number of dry days and total precipitation (rainy season).

Coincidence	ENSO Conditions				IOD Conditions		
	La Niña	***Normal***	***El Niño-Weak***	***El Niño-Strong***	***Negative***	***Neutral***	***Positive***
dry-dry	3.4%	17.8%	19.7%	39.9%	18.5%	13.5%	38.3%
normal-normal	15.3%	17.5%	9.9%	12.9%	15.0%	14.9%	7.6%
wet-wet	30.3%	12.7%	11.4%	2.9%	1.8%	19.8%	13.7%
Total Sync.	*48.9%*	*48.0%*	*41.0%*	*55.6%*	*35.2%*	*48.3%*	*59.5%*
Total Async.	*51.1%*	*52.0%*	*59.0%*	*44.4%*	*64.8%*	*51.7%*	*40.5%*

Different things are shown in the IOD conditions, where there is no evident influence from each IOD condition. There is no pattern of decreasing or increasing the probability of co-occurrence when the IOD index increases or decreases. The highest probability for co-occurrence in wet conditions is under Neutral IOD conditions. These results reaffirm that the IOD phenomenon does not significantly affect the joint PDF of the number of dry days and total precipitation in the rainy season.

Table 7.

Coincidence probability of the joint PDFs between the number of dry days and total precipitation (dry season).

Coincidence	ENSO Conditions				IOD Conditions		
	La Niña	***Normal***	***El Niño-Weak***	***El Niño-Strong***	***Negative***	***Neutral***	***Positive***
dry-dry	14.1%	21.7%	48.8%	68.5%	1.0%	22.3%	62.6%
normal-normal	20.9%	21.4%	13.7%	5.1%	19.7%	21.9%	8.9%
wet-wet	35.9%	25.7%	15.0%	4.4%	34.8%	26.7%	8.4%
Total Sync.	*70.8%*	*68.7%*	*77.5%*	*78.0%*	*55.5%*	*70.8%*	*80.0%*
Total Async.	*29.2%*	*31.3%*	*22.5%*	*22.0%*	*44.5%*	*29.2%*	*20.0%*

Table 7 shows that ENSO and IOD conditions significantly affect the coincidence probability between the number of dry days and total precipitation, confirming the statement in **Fig. 5**. The higher the ENSO and IOD indexes, the higher the dry condition probability for the number of dry days and total precipitation simultaneously. Strong El Niño and Positive IOD have a coincidence probability of 68.5% and 62.6%, respectively. The two values show almost the same probability, indicating that Strong El Niño and Positive IOD have similar effects on the coincidence probability of the number of dry days and total precipitation. Although the Weak El Niño gives a reasonably high value of 48.8%, this value is still relatively less than the Strong El Niño and Positive IOD conditions. During Strong El Niño, the probability of wet-wet conditions is 4.4%, while the probability of wet-wet conditions in Positive IOD is 8.4%.

Based on the value of total synchronous and asynchronous probability, only Strong El Niño and Positive IOD conditions give a total synchronous probability of more than 50% in the rainy season. In other ENSO and IOD conditions, the total asynchronous probability is higher than the total synchronous probability. This shows that the relationship between the number of dry days and total precipitation is not too high in the rainy season. Thus, the probability of different conditions occurring in the number of dry days and total precipitation is still relatively high, such as the probability of the number of dry days in wet conditions but total precipitation in normal conditions, etc.

Meanwhile, in the dry season, the total synchronic probability is very high, around 70%, which indicates the relationship between the number of dry days and total precipitation is very strong. Only when the IOD condition is negative the total synchronous probability value is 55%, but still higher than the total asynchronous probability. Coincidence probability between the number of dry days and the total amount of precipitation under Moderate-Strong El Niño and Positive IOD is more than 50% in the dry season, which implies that drought and fire potential preparedness should receive greater attention in these conditions.

5. CONCLUSIONS

South Sumatra is an area in Indonesia that is very vulnerable to forest fires with respect to ENSO and IOD events (Putra et al., 2019). Understanding the shared behavior between local climatic conditions such as the number of dry days and the total amount of precipitation is considered important as an early stage to form a probabilistic-based model in the future. The copula model is a very effective way to understand the shared behavior of climate indicators (Dixit & Jayakumar, 2022; Hussain et al., 2022).

This study has defined fire-prone areas in southern Sumatra using k-means clustering. Based on the fire-prone areas, 3-month the number of dry days and 2-month total precipitation gave the highest correlation to the burned area. The joint distribution is constructed using a rotated copula based on ENSO and IOD conditions from these two variables. Based on the peak of joint PDFs, ENSO and IOD conditions did not significantly affect the rainy season but substantially in the dry season. However, based on coincidence probability, ENSO conditions still influence the joint distribution of the number of dry days and total precipitation during the rainy season but not with IOD conditions.

Moreover, the association between the number of dry days and the total precipitation was stronger in the dry season than in the rainy season.

Although the copula is a powerful technique for accessing the joint distribution between variables that are not normally distributed, this method relies heavily on selecting the marginal distribution as the edges of the copula. This study only focuses on time-independent distribution functions, so climate change is not considered in the analysis here. The use of time-varying copula can improve further understanding and make the joint distribution model more precise (Ausin & Lopes, 2010; Xu et al., 2021). In addition, high-dimensional copulas can be used in the same analysis by making the ENSO or IOD index as one of the edges of the copula. This means that the joint distribution model used is no longer divided based on ENSO or IOD categories, but the joint distribution model that is formed can assess the behavior of climate indicators based on the movement of the ENSO or IOD index using tri-variate copulas or more, such as nested Archimedean copula (Zhao et al., 2022).

REFERENCES

Ausin, M. C., & Lopes, H. F. (2010). Time-varying joint distribution through copulas. *Computational Statistics and Data Analysis*, *54*(11), 2383–2399. https://doi.org/10.1016/j.csda.2009.03.008

Avia, L. Q., & Sofiati, I. (2018). Analysis of El Niño and IOD Phenomenon 2015/2016 and Their Impact on Rainfall Variability in Indonesia. *IOP Conference Series: Earth and Environmental Science*, *166*(1). https://doi.org/10.1088/1755-1315/166/1/012034

Brown, P. J., Bradley, R. S., & Keimig, F. T. (2010). Changes in extreme climate indices for the Northeastern United States, 1870-2005. *Journal of Climate*, *23*(24), 6555–6572. https://doi.org/10.1175/2010JCLI3363.1

Brunton, S. L., & Kutz, J. N. (2022). *Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press.

Caraka, R. E., Yasin, H., Sugiarto, W., & Ismail, K. M. (2016). Time Series Analysis Using Copula Gauss and Ar(1)-N.Garch(1,1). *Media Statistika*, *9*(1). https://doi.org/10.14710/medstat.9.1.1-13

Dixit, S., & Jayakumar, K. V. (2022). Spatio-temporal analysis of copula-based probabilistic multivariate drought index using CMIP6 model. *International Journal of Climatology*, *42*(8), 4333–4350. https://doi.org/10.1002/joc.7469

Hauke, J., & Kossowski, T. (2011). Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, *30*(2), 87–93. https://doi.org/10.2478/v10117-011-0021-1

Hussain, B., Qureshi, N. A., Buriro, R. A., Qureshi, S. S., Pirzado, A. A., & Saleh, T. A. (2022). Interdependence between temperature and precipitation: modeling using copula method toward climate protection. *Modeling Earth Systems and Environment*, *8*(2), 2753–2766.https://doi.org/10.1007/s40808-021-01256-8

Khaldun, M. H. I., Wirasatriya, A., Dwi Suryo, A. A., & Kunarso. (2018). The Influence of Indian Ocean Dipole (IOD) on the Variability of Sea Surface Temperature and Precipitation in Sumatera Island. *IOP Conference Series: Earth and Environmental Science*, *165*(1). https://doi.org/10.1088/1755-1315/165/1/012008

Kurniadi, A., Weller, E., Min, S. K., & Seong, M. G. (2021). Independent ENSO and IOD impacts on rainfall extremes over Indonesia. *International Journal of Climatology*, *41*(6), 3640–3656. https://doi.org/10.1002/joc.7040

Lestari, D. O., Sutriyono, E., Sabaruddin, S., & Iskandar, I. (2018). Respective Influences of Indian Ocean Dipole and El Niño-Southern Oscillation on Indonesian Precipitation. *Journal of Mathematical and Fundamental Sciences*, *50*(3), 257–272. https://doi.org/10.5614/j.math.fund.sci.2018.50.3.3

Li, H. W., Li, Y. P., Huang, G. H., & Sun, J. (2021). Quantifying effects of compound dry-hot extremes on vegetation in Xinjiang (China) using a vine-copula conditional probability model. *Agricultural and Forest Meteorology*, *311*. https://doi.org/10.1016/j.agrformet.2021.108658

Li, Z., Shao, Q., Tian, Q., & Zhang, L. (2020). Copula-based drought severity-area-frequency curve and its uncertainty, a case study of Heihe River basin, China. *Hydrology Research*, *51*(5), 867–881. https://doi.org/10.2166/nh.2020.173

McPhaden, M. J., Zebiak, S. E., & Glantz, M. H. (2006). ENSO as an integrating concept in earth science.

Science, *314*(5806), 1740–1745. https://doi.org/10.1126/science.1132588

Mohamad, I. Bin, & Usman, D. (2013). Standardization and its effects on K-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, *6*(17), 3299–3303. https://doi.org/10.19026/rjaset.6.3638

Naeini, E. M. B., Akhoond-Ali, A. M., Radmanesh, F., Koupai, J. A., & Soltaninia, S. (2021). Comparison of the Calculated Drought Return Periods Using Tri-variate and Bivariate Copula Functions Under Climate Change Condition. *Water Resources Management*, *35*(14), 4855–4875. https://doi.org/10.1007/s11269-021-02965-6

Najib, M. K., Nurdiati, S., & Sopaheluwakan, A. (2021). Copula in Wildfire Analysis: A Systematic Literature Review. *InPrime: Indonesian Journal of Pure and Applied Mathematics*, *3*(2), 101–111. https://doi.org/10.15408/inprime.v3i2.22131

Najib, M. K., Nurdiati, S., & Sopaheluwakan, A. (2022a). Copula-based joint distribution analysis of the ENSO effect on the drought indicators over Borneo fire-prone areas. *Modeling Earth Systems and Environment*, *8*(2), 2817–2826. https://doi.org/10.1007/s40808-021-01267-5

Najib, M. K., Nurdiati, S., & Sopaheluwakan, A. (2022b). Multivariate fire risk models using copula regression in Kalimantan, Indonesia. *Natural Hazards*. https://doi.org/10.1007/s11069-022-05346-3

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Springer Science & Business Media.

Nikonovas, T., Spessa, A., Doerr, S. H., Clay, G. D., & Mezbahuddin, S. (2022). ProbFire: A probabilistic fire early warning system for Indonesia. *Natural Hazards and Earth System Sciences*, *22*(2), 303–322. https://doi.org/10.5194/nhess-22-303-2022

Nur'utami, M. N., & Hidayat, R. (2016). Influences of IOD and ENSO to Indonesian Rainfall Variability: Role of Atmosphere-ocean Interaction in the Indo-pacific Sector. *Procedia Environmental Sciences*, *33*, 196–203. https://doi.org/10.1016/j.proenv.2016.03.070

Nurdiati, S., Bukhari, F., Julianto, M. T., Najib, M. K., & Nazria, N. (2021). Heterogeneous Correlation Map Between Estimated ENSO And IOD From ERA5 And Hotspot In Indonesia. *Jambura Geoscience Review*, *3*(2), 65–72. https://doi.org/10.34312/jgeosrev.v3i2.10443

Nuryanto, D. E., & Badriyah, I. U. (2014). Sea Surface Temperature Effect to Indonesian Maritime Continent Rainfall on September 2006 (Pengaruh Perubahan Suhu Permukaan Laut Terhadap Curah Hujan Benua Maritim Indonesia Pada September 2006). *Jurnal Meteorologi Dan Geofisika*, *15*(3), 147–155.

Putra, R., Sutriyono, E., Kadir, S., Iskandar, I., & Lestari, D. O. (2019). Dynamical link of peat fires in South Sumatra and the climate modes in the Indo-Pacific region. *Indonesian Journal of Geography*, *51*(1), 18–22. https://doi.org/10.22146/ijg.35667

Singh, H., Pirani, F. J., & Najafi, M. R. (2020). Characterizing the temperature and precipitation covariability over Canada. *Theoretical and Applied Climatology*, *139*(3–4), 1543–1558. https://doi.org/10.1007/s00704-019-03062-w

Suharjo, B. H., & Velicia, W. A. (2018). The Role of Rainfall Towards Forest and Land Fires Hotspot Reduction in Four Districs in Indonesia on 2015-2016. *Jurnal Silvikultur Tropika*, *09*(1), 24–30. https://doi.org/10.29244/j-siltrop.9.1.24-30

Ummenhofer, C. C., D'Arrigo, R. D., Anchukaitis, K. J., Buckley, B. M., & Cook, E. R. (2013). Links between Indo-Pacific climate variability and drought in the Monsoon Asia Drought Atlas. *Climate Dynamics*, *40*(5–6), 1319–1334. https://doi.org/10.1007/s00382-012-1458-1

Wei, X., Zhang, H., Singh, V. P., Dang, C., Shao, S., & Wu, Y. (2020). 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. *Hydrology Research*, *51*(5), 1120–1135. https://doi.org/10.2166/nh.2020.106

Xu, P., Wang, D., Wang, Y., Qiu, J., Singh, V. P., Ju, X., Zhang, A., Wu, J., & Zhang, C. (2021). Time-varying copula and average annual reliability-based nonstationary hazard assessment of extreme rainfall events. *Journal of Hydrology*, *603*. https://doi.org/10.1016/j.jhydrol.2021.126792

Yang, J., Wang, Y., Yao, J., Chang, J., Xu, G., Wang, X., & Hu, H. (2020). Coincidence probability analysis of hydrologic low-flow under the changing environment in the Wei River Basin. *Natural Hazards*, *103*(2), 1711–1726. https://doi.org/10.1007/s11069-020-04051-3

Zhao, Z., Wang, H., Shi, Q., & Wang, C. (2022). Study on drought events in China based on time-varying nested Archimedean-copula function. *Water Supply*, *22*(1), 795–811. https://doi.org/10.2166/ws.2021.233