JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 8, No. 4, October 2024, pp. 1217-1232

Prediction of Air Temperature in East Java using Spatial Extreme Value with Copula Approach

A'yunin Sofro[1]*, Wildan Habibulloh[1], Khusnia Nurul Khikmah[1]
[1]Department of Mathematics, Universitas Negeri Surabaya, Indonesia
ayuninsofro@unesa.ac.id

Article History:
Received : 20-07-2024
Revised : 06-10-2024
Accepted : 13-10-2024
Online : 15-10-2024

Keywords:
Air temperature;
Copula;
Extreme values;
Spatial.

ABSTRACT

The increase in world temperature or global warming is a form of imbalance in the average temperature on Earth. The increase in air temperature will increase the risk of disasters, which will occur more frequently in the future. Rising global temperatures are expected to cause changes that can have fatal consequences. To anticipate the dangers are predicted by predicting the future air temperature increase. One of the methods that can be used is spatial extreme value theory, which uses the Gaussian copula model approach and Student's t copula, where the choice of these two methods was based on the flexibility they offer in capturing tail dependencies due to their capacity to describe the dependence structure between many variables simultaneously. This makes it possible to get a return level or predicted value of air temperature by considering the elements of location in it. This research discusses both approaches and uses the maximum likelihood estimation (MLE) and pseudo maximum likelihood estimation (PMLE) methods to estimate the parameters. In addition, since spatial elements need to be considered, the trend surface model is also used. Akaike information criterion (AIC) is used to determine the best model for predicting air temperature based on extreme air temperature data in East Java Province from nine observation stations. The results show that the highest air temperature value is around the Banyuwangi temperature observation station located in Banyuwangi Regency in the next two-year return period. The AIC results show that the best model produced is the Gaussian copula approach with a smaller AIC value than the student's t-copula approach, which is 8.0174. This value with a lower AIC value generally indicates a better-fitting model.

https://doi.org/10.31764/jtam.v8i4.25436

A. INTRODUCTION

The increase in world temperature, or what is referred to by most people in the world, is certainly known as global warming. Global warming, a form of ecosystem imbalance due to an increase in the average temperature on Earth, is an interesting and serious subject of discussion because it is directly related to the fate of mankind (Upadhyay, 2020). The Intergovernmental Panel on Climate Change in its 5th assessment report shows an increase in global average air temperature estimated to reach 0.85 during the 1880 to 2012 (Ahmed et al., 2023). The increase in global temperature is expected to cause changes that can have fatal consequences; the melting of the ice sheets at the north and south poles, the extinction of various types of fauna, and the increasing intensity of extreme weather phenomena. One natural phenomenon that is rare but difficult to avoid is the increasing intensity of extreme weather phenomena such as high or extreme temperature changes. Head of the Meteorology, Climatology and Geophysics

Agency, said that even a one-degree Celsius increase in temperature can lead to extreme weather such as tropical cyclones, extreme rains, strong winds or tornadoes, high waves, which can trigger floods, flash floods, landslides and other hydrometeorological disasters. Therefore, extreme changes in air temperature are a big problem for human life in the world. Natural events; extreme air temperature changes, sea tides, strong winds and so on can be influenced by the conditions of a spatial (Zscheischler et al., 2020). Based on the negative impacts that are likely to occur due to extreme temperature changes, it is necessary to study extreme events, especially spatially. The role of spatial analysis is crucial, as it allows us to measure and describe the characteristics and behaviour of natural phenomena such as air temperature, wind speed, and others.

In spatial data, the main thing considered is the dependence between locations, where events in a nearby location tend to have similarities with events in more distant locations (Cressie & Moores, 2023). Extreme value theory (EVT) is one of the statistical methods to identify extreme events. EVT is developed from the univariate case with extreme events in one variable and is often applied to stock data. Temperature, precipitation, snow, and river discharge data are spatial data, which are multivariate because they are observed in several locations, so the spatial extreme value method was developed (Melina et al., 2023). In the case of multivariate data, copula approaches and max-stable processes. Several methods exist to analyze spatial extreme value events, including the copula approach conducted (Boulaguiem et al., 2022; Carreau & Toulemonde, 2020). Another study used a hierarchical Bayesian approach conducted (F. Hussain et al., 2022). In addition, some studies use the Max Stable approach, conducted (Engelke & Hitz, 2020).

Copula is divided into two types of elliptical and Archi-median. For the spatial extreme model, the copula that can be used is an elliptical copula. Copulas included in the elliptical copula are the Gaussian copula and the student's t-copula (Gimeno-Sotelo & Gimeno, 2022). This study chose the two approaches of Gaussian copula and Student's t-copula with the consideration that the choice between Gaussian copula and Student's t-copula depends on the specific characteristics of the data and which is the purpose of the study. Another consideration is that the research data are finite extreme values. Therefore, the Gaussian copula and Student's t-copula approaches proposed in this study are more appropriate. Comparison of the performance of the two approaches using fit tests and other evaluation metrics. This is intended to help identify the model that best captures the underlying dependency structure of the data. Some research on copula has been done, including research by Renard & Lang (2007) with the Gaussian copula approach in hydrology (Nazeri Tahroudi et al., 2021). Another study was conducted using the student's t-copula approach to rainfall. In addition, Zhong et al. (2021) took a Gaussian copula approach in modelling flood disaster losses due to rainfall. The results showed that the copula approach provided appropriate results for extreme observation data. In this research, the case study was conducted in East Java province. One of the leading causes of the increase in world temperature is the burning of fossil fuels such as oil, natural gas, and coal, which releases CO2 and other greenhouse gases into the atmosphere. As the Central Bureau of Statistics reported in 2022, East Java Province ranks first in the number of motorized vehicles in Indonesia. The number of motorized vehicles in East Java is around 23591769 units.

The number of motorized vehicles in East Java can indirectly affect changes in extreme temperatures made possible through several mechanisms: the operation of motorized vehicles releases large amounts of greenhouse gases, especially carbon dioxide (CO_2), contributing to global warming. In addition, increasing numbers of vehicles often necessitate the construction of new roads, parking lots, and other infrastructure, leading to changes in land use. These changes increase the absorption of solar radiation, contributing to higher temperatures. However, the research challenges that data limitations pose are what this study considers by integrating local data on temperature extremes to gain a more comprehensive understanding of the factors driving climate change in East Java and inform effective mitigation strategies. The study of climate change has become increasingly urgent due to its far-reaching implications for human society and the environment. Understanding the future temperature rise is crucial for developing effective mitigation and adaptation strategies.

B. METHODS

1. Extreme Value Theory

Extreme value theory (EVT) is one of the statistical methods developed to identify extreme events by looking at the patterns and characteristics of extreme events. These rare extreme events tend to have a significant impact, even briefly (Naveau et al., 2020; Towler et al., 2020). This method aims to determine the estimated probability of extreme events by considering the tail of the distribution function based on the extreme values obtained. Studies show that climate data have stochastic behaviour with heavy tails. The first step based on EVT for further analysis is determining extreme values by looking at the pattern and characteristics of extreme events. There are two approaches (block maxima (BM) and peaks over threshold (POT)) to identifying extreme values based on EVT (Orsini et al., 2020). This study uses temperature data, a seasonal data type, where the data pattern is influenced by the season, so the BM method is used.

2. Block Maxima (BM)

BM is the methods used to identify extreme values based on the highest value of observation data grouped in a certain period. The BM method divides the observation data into blocks at a certain period (Ramadhani et al., 2017). Through splitting a dataset into equal-sized, non-overlapping blocks, the block maxima technique determines the greatest value from each block. In order to locate and examine extreme values within a dataset, this technique is frequently employed. Each block is determined by the highest value, which is called the extreme value of each block. According to Guermah & Rassoul (2020); Szigeti et al. (2020), the BM method applies the Fisher-Tippet theorem that the extreme value sample data taken from the BM method will follow the generalized extreme value (GEV) distribution. The cumulative distribution function of the GEV distribution is:

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

Where x is the extreme value obtained from the BM method, μ is the location parameter, σ is scale parameter with $\sigma > 0$, and ξ is shape parameter and has a probability distribution function as follows (Tian et al., 2023):

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

GEV has three distribution parameters, namely location parameter (μ), scale parameter (σ), and shape parameter (ξ). The shape parameter follows three distributions, namely Weibull ($\xi < 0$), Gumbel ($\xi = 0$), and Frechet ($\xi > 0$). The CDF forms of the three distributions are presented:

a. Gumbel distribution (type I extreme value distribution) for $\xi = 0$

$$f(x;\mu,\sigma,\xi) = \exp\left\{-exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, -\infty < x < \infty \quad (3)$$

b. Frechet distribution (extreme value distribution type II) for $\xi > 0$

$$f(x;\mu,\sigma,\xi) = \exp\left\{-exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, -\infty < x < \infty \quad (4)$$

c. Weibull distribution (extreme value distribution type III) for $\xi < 0$

$$f(x;\mu,\sigma,\xi) = \begin{cases} exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, x < \mu \\ 1, x \geq \mu \end{cases} \quad (5)$$

The greater the value of (ξ), the heavier the tail, the greater the chance of extreme values. Therefore, based on the three types of GEV distributions, the distribution with the heavy tail is the Frechet distribution. The difference in these distributions can be seen at the ends of the three-parameter shapes, which makes the distribution challenging to observe. So, the GEV distribution combines the three by following the shape of the GEV distribution. The shape parameter with $\xi = 0$ is to be the medium tail; for $\xi > 0$, it is to be a long tail, and for $\xi < 0$, it is to be the short tail.

Parameter estimation of the GEV distribution can use the maximum likelihood estimation (MLE). The main thing in estimating parameters with MLE is maximizing a distribution's PDF likelihood function. Estimating the parameters μ, σ, and ξ of the GEV distribution using the MLE, where the likelihood function is a joint probability function $x_1, x_2, \ldots, x_n$. Based on the MLE method, the likelihood function of the GEV. For $\xi \neq 0$, likelihood function is as follow.

$$L(f|x_i;\mu,\sigma,\xi) = \left(\frac{1}{\sigma}\right)^n \prod_{i=1}^{n} \left\{\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}-1}\right\} exp\left\{-\prod_{i=1}^{n}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \quad (6)$$

then log likelihood is as follow.

$$\ln L(f|x_i;\mu,\sigma,\xi) = -\,n\ln(\sigma) + \sum_{i=1}^{n}\left[\left(\frac{1}{\xi}-1\right) ln\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right] - \left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right] \quad (7)$$

Furthermore, the approach chosen in this study is copula which is explained in the next section.

3. Copula Approach

Copula was introduced by Abe Sklar in 1959 (Sklar's theorem). A copula is a function that relates a multivariate distribution function to a distribution. Copula can be explored and characterized as the dependency structure between random variables through its marginal distribution function (Geenens, 2020). In addition, copulas are flexible tools that can capture various dependency structures, including linear, non-linear, and asymmetric relationships, often used in climate modeling to assess the joint probability of extreme events. The choice of copula function depends on the specific characteristics of the data and the research objectives, which in this study can effectively analyze the extreme value phenomenon and model the dependence structure between multiple variables. Copulas are considered capable of dealing with dependencies of non-normally distributed variables. Copulas are divided into two families; elliptical and Archi-median copulas. In the case of the spatial extreme copula model, an elliptical copula can be used. The elliptical copula and the student's included, along with the Gaussian copula t-copula.

The difference between copula and extreme value copula lies in the transformation process, where the copula transformation process uses a uniform standard distribution. At the same time, extreme value copula also uses the CDF of GEV. In this study, the spatial approach used is extreme value copula. Extreme value copula is more appropriate for heavy tail data when the parameter $\xi > 0$. Of the three GEV distributions, Frechet distribution is the most heavy-tail. So, the GEV parameter estimation results will follow the Frechet distribution. In the Gaussian Copula for extreme spatial cases, the transformation process uses the GEV CDF with the transformation equation defined in the following equation:

$$U_j = F_j(x_{ij}) \quad (8)$$

F_j is the GEV CDF, and x_{ij} is the j-th and i-th observational data. So, the copula CDF follows the following equation:

$$C(u_1,\dots,u_j) = \Phi\left(\Phi'(\mathrm{u}_1),\dots,\Phi'(\mathrm{u}_\mathrm{j})\right) \quad (9)$$

Where Φ is the CDF of the normal distribution because the copula used is the Gaussian copula, the PDF of the Gaussian copula follows the form (S. I. Hussain & Li, 2022).

$$C(u_1, \dots, u_j) = \frac{\partial C(u_1, \dots, u_j)}{\partial u_1, \partial u_2, \partial u_3, \dots, \partial u_j} \tag{10}$$

According to Sklar's theorem, where f is the GEV PDF the copula has the following joint probability function.

$$f(x_{1,1}, x_{1,2}, \dots, x_{i,j}) = f(x_{1,1}) \cdot f(x_{1,2}) \cdot \dots \cdot f(x_{i,j}) \cdot c(u_1, \dots, u_j) \tag{11}$$

The other comparison approach chosen in this study besides copula is Student's t-copula.

4. Student's t-Copula

Student's t-copula is the copulas suitable for spatial extreme modelling. It is defined like a Gaussian copula but uses a multivariate extension of the t-distribution. According to (Zhang et al., 2022), t-copula shows flexibility in covariance structure and tail dependence. Tail dependencies can be considered the conditional probability of extreme observations in one component at an extreme state. The t-copula has the potential to generate extreme values because t is a skewed distribution. The CDF form of the student's t-copula can be defined as follows:

$$C(u_1, u_2, \dots, u_m) = F_{t(v,\Sigma)}\left(F_{t(v)}^{-1}(u_1), F_{t(v)}^{-1}(u_2), \dots, F_{t(v)}^{-1}(u_m)\right) \tag{12}$$

F_t defines the CDF of the multivariate distribution t. Then, $F_{t(v)}^{-1}$is the inverse CDF of the multivariate distribution of t. To obtain the PDF of the student's t-copula, the derivative of the student's t-copula CDF function is performed. So, that the PDF form is as follows:

$$c(u_1, \dots, u_m) = \frac{\partial}{u_1}, \dots, \frac{\partial}{u_m} \cdot C(u_1, u_2, \dots, u_m) \tag{13}$$

The form of the student's t copula distribution function can be written as follows:

$$c(u) = \frac{\Gamma^{\frac{v+d}{2}}}{\Gamma^{\left(\frac{v}{2}\right)} |\rho(h)|^{\frac{1}{2}}} \left\{1 + \frac{u^T P^{-1} u}{v}\right\}^{-\frac{v+d}{2}} \tag{14}$$

P is the correlation function ρ, written as $\rho(h)$, h is the distance between post locations, and u is the copula transformation (Czado & Nagler, 2022). So, the student's t- t-copula PDF is obtained in the following equation:

$$c(u_1, \dots, u_m) = \frac{\Gamma^{\frac{v+d}{2}}}{\Gamma\left(\frac{v}{2}\right)} \left\{1 + \frac{V^{\tau}\big(\rho(h)\big)V}{v}\right\}^{-\frac{v+d}{2}} \cdot |\rho(h)|^{\frac{1}{2}} \tag{15}$$

Where $v = \left(F_{t(v)}^{-1}(u_1), F_{t(v)}^{-1}(u_2), \dots, F_{t(v)}^{-1}(u_m)\right)$. According to Sklar's theorem, each probability with a copula can be written by multiplying the marginal distribution PDF with the copula function CDF so that the general formula can be written based on the equation.

$$(x_1, \dots, x_m) = f_{x_1}(x_1), \dots, f_{x_m}(x_m), c(u_1, \dots, u_m) \tag{16}$$

Furthermore, the model results generated from the two proposed approaches are predicted using the return level.

5. Return Level

The return level is the maximum value expected to occur in the future. The concept used in the return level is to predict rare events, such as extreme temperature changes, that have a negative impact (Min & Halim, 2020). Return level is also an application of the GEV distribution, from the CDF of GEV, the probability of extreme temperature is $1 - p$, where p is the return period with $p = 1$, where k is the number of blocks formed so that:

$$\hat{R}_j = \hat{\mu}_j \left[1 - \{-ln(1-p)\}^{-\hat{\xi}_j}\right] \tag{17}$$

Where μ_j is the j-th location parameter estimate, σ_j is the j-th scale para-meter estimate, ξ_j is the j-th shape parameter estimate, and p is the return period (Rypkema & Tuljapurkar, 2021). The prediction results generated by the two proposed approaches are then evaluated based on the comparison of the predicted results produced with the actual value with the AIC goodness measure.

6. Akaike Information Criterion (AIC)

The best model selection is done by comparing more appropriate models for the data. Akaike Information Criterion (AIC) is a method of determining the best model. So, this method is used to determine or select the best model to result from this research. Choosing a simple model is better in specific contexts than choosing a complex one. AIC is defined by the following equation.

$$AIC = -2 \, log \, L(\hat{\beta}) + 2q \tag{18}$$

Where log $L(\hat{\beta})$ is the log-likelihood function of each approach, and q is the number of parameters estimated (Cavanaugh & Neath, 2019; Portet, 2020).

7. Data

The data used in this study are secondary data obtained from the Meteorology and Geophysics Agency (BMKG) website, https://dataonline.bmkg.go.id/. The data used is temperature data in East Java, consisting of 9 observation stations from 2013 to 2022. List of 9 observation stations; East Java Climatology Station, Malang Geophysical Station, Banyuwangi Meteorological Station, Pasuruan Geophysical Station, Juanda Meteorological Station, Nganjuk Geophysical Station, Sangkapura Meteorological Station, Trunojoyo Meteorological Station, and Perak Meteorological Station 1.

8. Research procedure

The steps to determine the air temperature prediction in East Java Province using return levels with spatial extreme value theory with a Gaussian copula approach are to identify and provide descriptive statistics of data obtained from several observation stations. Then, extreme data values are determined by grouping data using the BM method. From the results of the BM method, several rainfall blocks are obtained. Then, the maximum value of each block formed is taken, and the GEV parameters are univariately using MLE. Next, calculate dependencies between spatial locations using the Gaussian copula approach transforming the GEV-distributed BM data into a copula and estimating the parameters of the Gaussian copula using the trend surface model with MLE. Finally, find the return level value of extreme events that will occur in the future.

Meanwhile, to find out the air temperature prediction in East Java Province using return levels with spatial extreme value theory with the student t-copula approach, the following steps are taken by performing identification and descriptive statistics of data obtained from several observation stations. Then, extreme data values are determined by grouping data using the BM method. Several rainfall blocks were obtained from the results of the BM method. Rainfall blocks. Then, the maximum value of each block formed is taken, and the GEV parameters are univariately using MLE. Next, calculate dependencies between spatial locations using the student t-copula approach and transform the BM data with GEV distribution into a copula. Lastly, estimate Gaussian copula parameters using the trend surface model with PMLE and find the return level value of extreme events that will occur in the future. Then Find the best model is derived from the case of air temperature in East Java Province, the AIC value and interpretation of results.

C. RESULT AND DISCUSSION

Based on their theoretical qualities and capacity to accurately represent the features of extreme value distributions, the block maxima and Generalized Extreme Value (GEV) distributions were chosen to describe extreme temperature data in East Java. The two techniques work together to analyze East Java's extreme temperature data. Seasonal extremes can be found using the block maxima approach, and the GEV distribution offers an adaptable framework for simulating the tail behavior of these extreme values. This study can learn a great deal about the frequency and intensity of extreme temperature episodes in the area by combining these techniques.

The first step of this analysis identifies the characteristics of air temperature data in East Java to understand air temperature patterns. The correlation of each station has been presented on Figure 1. Based on information from BMKG, Indonesia has an air temperature pattern that follows the monsoonal season, which means that the dry and rainy seasons have apparent differences. East Java Province has nine observation stations to record air temperature. The data collected is air temperature data at the observation location. Then, two periods of air temperature in one year are formed. Each block will contain temperature data for six months based on the monsoonal pattern, namely the April-May-June-July-August-September block (AMJJAS) and the October-November-December-January-February-March block (ONDJFM). The maximum air temperature value will be retrieved for each block using the Block Maxima (BM) method. As a result, from 2013 to 2022, 20 blocks will be formed, thereby obtaining extreme values. Based on each location's extreme air temperature values, descriptive statistical data is generated as shown in Figure 1.

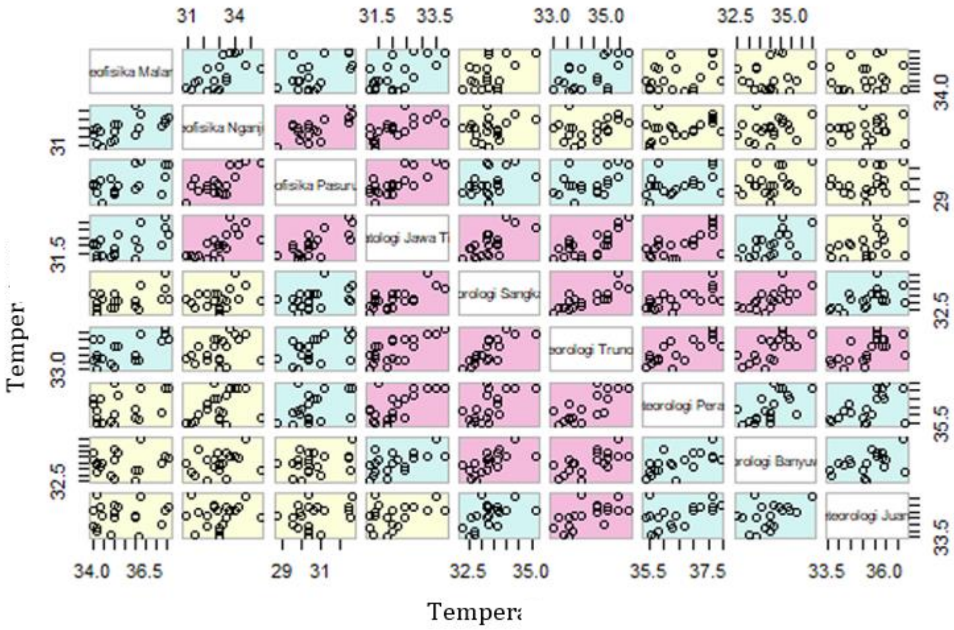

Figure 1. Correlation plot of extreme air temperature between observation sites

Table 1. Statistics Descriptive

Station	Mean	Min	Max	Median
East Java Climatology	32.09	31.20	33.80	32.00
Malang Geophysical	35.42	34.00	37.60	35.00
Banyuwangi Meteorological	34.03	32.50	36.00	34.00
Pasuruan Geophysical	30.65	28.80	32.60	30.40
Juanda Meteorological	35.20	33.60	36.80	35.30
Nganjuk Geophysical	32.92	30.80	35.60	33.00
Sangkapura Meteorological	33.30	32.20	35.20	33.25
Trunojoyo Meteorological	34.31	33.00	35.80	34.30
Perak Meteorological 1	36.52	35.40	37.90	36.35

After obtaining descriptive statistical data of extreme air temperature using the BM, it can be concluded that the distribution of extreme values taken by the BM will follow the GEV distribution. After the extreme data obtained from the BM has been validated following the GEV distribution, GEV parameter estimation is carried out using MLE method univariately. In this stage, we will estimate the parameters of the GEV distribution, which consists of three parameters. The following are the results of parameter estimation based on the location.

Table 2. Estimated value of GEV distribution parameters

Station	Location (μ)	Scale (σ)	Shape (ξ)
East Java Climatology	31.689	0.510	0.213
Malang Geophysical	34.674	0.765	0.391
Banyuwangi Meteorological	33.706	0.804	-0.210
Pasuruan Geophysical	30.190	0.882	-0.069
Juanda Meteorological	34.937	0.920	-0.385
Nganjuk Geophysical	32.480	1.107	-0.217
Sangkapura Meteorological	33.006	0.548	-0.042
Trunojoyo Meteorological	34.033	0.781	-0.296
Perak Meteorological 1	36.209	0.779	-0.240

The value of Table 2 can be entered into the GEV CDF equation for Banyuwangi Meteorological, which results in the following equation, where location (μ) shows a larger value which means a higher overall level of extreme values or at a temperature of 33.706°, while scale (σ) shows that the extreme temperatures are relatively spread out with a standard deviation of 0.5° and shape (ξ) shows that the distribution is weakly tailed, which indicates a lower probability of very hot temperatures of 0.210.

$$F(x_{ij};\mu,\sigma,\xi) = F(x_{ij}, 33.706, 0.804, -0.210) = exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}$$

Table 2 indicates that the shape parameter estimates at the seven air temperature observation stations have negative values and are almost the same. This indicates a similarity in the characteristics of air temperature at that location and homogeneity in air temperature characteristics. In addition, when the estimated values of the shape parameter are negative ($\xi < 0$), it indicates that the distribution of extreme values follows the Reversed Weibull type of the GEV distribution. Since the data covers many locations, spatial factors affect the observed air temperature. It was explained that when extreme events occur at different locations, an approach is needed to describe the dependencies between these locations. Two approaches that can be used are the Gaussian copula and the student's t-copula.

The first step in the Gaussian copula approach is transforming the extreme value distribution into a Frechet distribution. This is done because data with a heavy tail ($\xi < 0$) is more suitable for using the copula approach. Then the transformations will be performed using the z method for all extreme values at each air temperature observation location. After transforming into the Frechet distribution. The copula transformation results for each extreme value obtained through the BM method will be obtained using similar calculations. After obtaining the copula transformation results, the parameter $\hat{\beta}$ is estimated and used in the trend surface model. To estimate the parameter $\hat{\beta}$, we use the Gaussian. From the Gaussian copula model, the parameter estimate $\hat{\beta}$ will be obtained as shown in Table 3.

Table 3. Parameter estimation values $\hat{\beta}$ Gaussian Copula

Parameter $\hat{\beta}$	Value
$\beta_{1,0}$	-21.067
$\beta_{1,1}$	0.502
$\beta_{1,2}$	0.341
$\beta_{2,0}$	9.532
$\beta_{2,1}$	-0.105
$\beta_{2,2}$	-0.538
$\beta_{3,0}$	-0.220

By using the parameter estimates from the Gaussian copula model, the following trend surface model is obtained. The trend surface model, the spatial dependency between locations can be calculated using each observation station's longitude and latitude coordinates (Ribatet, 2015). Since the trend surface model requires the location coordinates of each observation station, the following table illustrates the location coordinates, as shown in Table 4.

Table 4. Location coordinates of air temperature observation stations

Station	Latitude	Longitude
East Java Climatology	-7.901	112.598
Malang Geophysical	-8.150	112.450
Banyuwangi Meteorological	-8.215	114.355
Pasuruan Geophysical	-7.705	112.635
Juanda Meteorological	-7.385	112.783
Nganjuk Geophysical	-7.735	111.767
Sangkapura Meteorological	-5.851	112.658
Trunojoyo Meteorological	-7.040	113.914
Perak Meteorological 1	-7.224	112.724

Based on the table above, we can calculate the estimation of Gaussian copula parameters between locations as listed in the Table 5.

Table 5. Gaussian copula parameter estimation values

Station	Location (μ)	Scale (σ)	Shape (ξ)
East Java Climatology	32.725	2.003	-0.220
Malang Geophysical	32.566	2.152	-0.220
Banyuwangi Meteorological	33.500	1.988	-0.220
Pasuruan Geophysical	32.811	1.893	-0.220
Juanda Meteorological	32.994	1.705	-0.220
Nganjuk Geophysical	32.365	2.000	-0.220
Sangkapura Meteorological	33.454	0.893	-0.220
Trunojoyo Meteorological	33.679	1.401	-0.220
Perak Meteorological 1	33.019	1.625	-0.220

The results of the above table show that the shape parameter value (ξ) has a constant value at each air temperature measurement station. This is because measuring the shape parameter in air temperature does not consider other factors, such as wind speed or the altitude at which the air temperature is measured. Then, the results of spatial parameter estimation using the Gaussian copula approach can be used to calculate the return level of air temperature, referred to as air temperature prediction. For the calculation of the return level value of air temperature with a return period of 2 years in the future, as shown in Table 6.

Table 6. Return level of Gaussian copula approach

Station	2023-2024	2025-2026	2027-2028
East Java Climatology	34.908	35.982	36.508
Malang Geophysical	34.912	36.067	36.631
Banyuwangi Meteorological	35.666	36.733	37.255
Pasuruan Geophysical	34.874	35.890	36.387
Juanda Meteorological	34.852	35.768	36.215
Nganjuk Geophysical	34.545	35.618	36.143
Sangkapura Meteorological	34.427	34.906	35.140
Trunojoyo Meteorological	35.206	35.958	36.326
Perak Meteorological 1	34.790	35.662	36.088

Based on the Table 6, it can be concluded that in the biennial periods of 2023-2024, 2025-2026, and 2027-2028, the air temperature around the Banyuwangi Meteorological temperature observation station in Banyuwangi will reach the highest level. The second approach is the student's t-copula. Starting by transforming the extreme value distribution into a Frechet distribution. The transformation results show in Table 2. The copula transformation results for each extreme value obtained through the BM method will be obtained using similar calculations. After obtaining the copula transformation results, the parameter $\hat{\beta}$ is estimated and used in the trend surface model. Thus, the parameter estimates $\hat{\beta}$ will be obtained, which can be seen in the Table 7.

Table 7. Parameter estimation values $\hat{\beta}$ Student's t-Copula

Parameter $\widehat{\boldsymbol{\beta}}$	Value
$\beta_{1,0}$	-16.333
$\beta_{1,1}$	0.403
$\beta_{1,2}$	-0.463
$\beta_{2,0}$	4.206
$\beta_{2,1}$	-0.038
$\beta_{2,2}$	-0.437
$\beta_{3,0}$	-0.284

So, from the estimation of the student's t-copula parameter, we get a trend surface model, the spatial dependency between locations can be showed on Table 4, the student's t-copula parameter estimation using the trend surface model can be calculated as follows.

Table 8. Student's t-copula parameter estimation values

Station	Location (μ)	Scale (σ)	Shape (ξ)
East Java Climatology	32.742	3.378	-0.284
Malang Geophysical	32.797	3.492	-0.284
Banyuwangi Meteorological	33.596	3.448	-0.284
Pasuruan Geophysical	32.666	3.291	-0.284
Juanda Meteorological	32.577	3.145	-0.284
Nganjuk Geophysical	32.330	3.337	-0.284
Sangkapura Meteorological	31.817	2.480	-0.284
Trunojoyo Meteorological	32.874	2.952	-0.284
Perak Meteorological 1	32.479	3.077	-0.284

The results of the above Table 8 show that the shape parameter value (ξ) has a constant value at each air temperature measurement station. This is because the measurement of shape parameters in air temperature does not consider other factors, such as wind speed or altitude, where air temperature is measured. The results of spatial parameter estimation using the Gaussian copula approach can be used to calculate the return level of air temperature, referred to as air temperature prediction. Return level or return value in the case of air temperature can be calculated involves the spatial parameter estimation results obtained from the student's t-copula approach. For the calculation of the return level value of air temperature with a return period of 2 years in the future. Then, the return value at each station can be show in the Table 9.

Table 9. Return level of Student's t-copula approach

Station	2023-2024	2025-2026	2027-2028
East Java Climatology	36.287	37.923	38.692
Malang Geophysical	36.463	38.154	38.950
Banyuwangi Meteorological	37.215	38.885	39.671
Pasuruan Geophysical	36.119	37.713	38.463
Juanda Meteorological	35.878	37.402	38.119
Nganjuk Geophysical	35.832	37.448	38.208
Sangkapura Meteorological	34.420	35.622	36.187
Trunojoyo Meteorological	35.972	37.402	38.074
Perak Meteorological 1	35.709	37.199	37.900

Based on the Table 9, it can be concluded that in the biennial period of 2023-2024, 2025-2026, and 2027-2028, the highest air temperature around Banyuwangi Meteorological temperature observation station in Banyuwangi. The next step is to determine the best results from the Gaussian models used, namely the Gaussian copula and student's t-copula approaches, the Akaike's information criterion (AIC) method is used. The approach with the lowest in this method is considered the best approach. Therefore, the following are the result values of the AIC method, as shown in Table 10.

Table 10. AIC Results

Approach	AIC Value
Gaussian Copula	728.017
Student's t-Copula	815.161

While the Student copula is theoretically often considered the first choice for modeling extreme value phenomena, this study found that the Gaussian copula can provide better results in specific contexts, such as the analysis of extreme temperature data in East Java. The choice of copula model can significantly affect the accuracy of temperature predictions and, consequently, the effectiveness of climate change mitigation efforts. By carefully considering the characteristics of the data used in this research, that the data used is limited data due to the availability of data at climatology stations and the objectives of the research, researchers can choose the most appropriate copula model for analysis by considering the value of prediction accuracy with AIC.

From the Table 10, it can be concluded that the Gaussian copula approach is better than the student's t-copula approach. This is because the AIC value of the Gaussian copula approach is much lower than the student's t-copula approach. Previous research on spatial extreme values with max-stable and Gaussian copula approaches showed that the results of modelling spatial extreme values with the Gaussian copula approach are better than the max-stable approach (Artha & Sofro, 2019). Another research using the student's t-copula approach to modelling spatial extreme values (Fauziyah & Purnomo, 2020). In this study, researchers compared the Gaussian copula approach with the student's t-copula approach to modelling spatial extreme values. The results of this study show that spatial extreme values modelling with the Gaussian copula approach are better than the student's t-copula in predicting air temperature in East Java Province. Given the many negative impacts that can occur due to an increase in extreme air temperature. The results of this research can be used as information and reference material to anticipate negative impacts that may occur in the future.

Predicting temperature accurately can be significantly impacted by the copula model selection, especially in extreme events. Forecasts of future temperature rises may become more accurate if the Gaussian copula accurately represents the underlying data structure. Precise temperature forecasts are necessary to create efficient methods of mitigating climate change, such as developing plans for adaptation to handle the effects of climate change (infrastructure upgrades or modifications to agricultural methods), evaluating the influence of various mitigation measures on temperature trends, and making necessary adjustments to strategy.

D. CONCLUSION AND SUGGESTIONS

Based on the analysis, the predicted return level in this research uses a return period of 2 years. The results of the Gaussian copula approach to calculate the extreme air temperature return level in East Java show that the highest air temperature is expected to occur around the Banyuwangi temperature observation station located in Banyuwangi in the sequential periods of 2023-2024, 2025-2026, and 2027-2028. The return level for extreme air temperature events in East Java with the student's t-copula approach provides results showing that the highest air temperature is expected to occur in the exact location as the Gaussian copula approach, namely around the Banyuwangi temperature observation station located in Banyuwangi. However, the resulting return levels have different air temperature values in the sequential 2023-2024, 2025-2026, and 2027-2028. Thus, the best model produced is the Gaussian copula approach because it has a smaller AIC value than the student's t-copula approach, which is 8.0174. The actual data exhibits heavier tails or extreme events. The Student's t copula may provide more accurate predictions where the air temperature data in the study exhibits a relatively normal distribution with limited extreme events. So, the choice of model can also influence the precision of predictions. A model that better captures the underlying data structure will likely provide more precise estimates. From the conclusions obtained, air temperature prediction can be developed with more locations and outputs. Future research can be explored using a dynamic copula that allows time-varying dependence and incorporates spatial dependence.

REFERENCES

Ahmed, M., Ahmad, S., & Kheir, A. M. S. (2023). Climate change: an overview. *Global Agricultural Production: Resilience to Climate Change*, 1–30. https://doi.org/10.1007/978-3-031-14973-3_1.

Artha, A. A., & Sofro, A. (2019). Rainfall prediction in east java using spatial extreme value theory. *Journal of Physics: Conference Series*, *1417*(1), 012010. https://doi.org/10.1088/1742-6596/1417/1/012010.

Boulaguiem, Y., Zscheischler, J., Vignotto, E., van der Wiel, K., & Engelke, S. (2022). Modeling and simulating spatial extremes by combining extreme value theory with generative adversarial networks. *Environmental Data Science*, *1*, e5. https://doi.org/10.1017/eds.2022.4.

Carreau, J., & Toulemonde, G. (2020). Extra-parametrized extreme value copula: Extension to a spatial framework. *Spatial Statistics*, *40*, https://doi.org/10.1016/j.spasta.2020.100410.

Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*, *11*(3), https://doi.org/10.1002/wics.1460.

Cressie, N., & Moores, M. T. (2023). Spatial statistics. In *Encyclopedia of mathematical geosciences* (pp. 1362–1373). Springer. https://doi.org/10.1007/978-3-030-85040-1_31.

Czado, C., & Nagler, T. (2022). Vine copula based modeling. *Annual Review of Statistics and Its Application*, *9*(1), 453–477. https://doi.org/10.1146/annurev-statistics-040220-101153.

Engelke, S., & Hitz, A. S. (2020). Graphical models for extremes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(4), 871–932. https://doi.org/10.1111/rssb.12355.

Fauziyah, M., & Purnomo, J. D. T. (2020). Spatial extreme modeling using student t copula approach in Ngawi Regency. *Journal of Physics: Conference Series*, *1538*(1), 012051. https://doi.org/10.1088/1742-6596/1538/1/012051.

Geenens, G. (2020). Copula modeling for discrete random vectors. *Dependence Modeling*, *8*(1), 417–440. https://doi.org/10.1515/demo-2020-0022

Gimeno-Sotelo, L., & Gimeno, L. (2022). Concurrent extreme events of atmospheric moisture transport and precipitation: the role of atmospheric rivers. *Authorea Preprints*. https://doi.org/10.1002/essoar.10508750.3.

Guermah, T., & Rassoul, A. (2020). Study of extreme rainfalls using extreme value theory (case study: Khemis-Miliana region, Algeria). *Communications in Statistics: Case Studies, Data Analysis and Applications*, *6*(3), 364–379. https://doi.org/10.1080/23737484.2020.1789901.

Hussain, F., Li, Y., Arun, A., & Haque, M. M. (2022). A hybrid modelling framework of machine learning and extreme value theory for crash risk estimation using traffic conflicts. *Analytic Methods in Accident Research*, *36*, https://doi.org/10.1016/j.amar.2022.100248.

Hussain, S. I., & Li, S. (2022). Dependence structure between oil and other commodity futures in China based on extreme value theory and copulas. *The World Economy*, *45*(1), 317–335. https://doi.org/10.1111/twec.13123.

Melina, Sukono, Napitupulu, H., & Mohamed, N. (2023). A conceptual model of investment-risk prediction in the stock market using extreme value theory with machine learning: a semisystematic literature review. *Risks*, *11*(3), 60. https://doi.org/10.3390/risks11030060.

Min, J. L. J., & Halim, S. A. (2020). Rainfall modelling using generalized extreme value distribution with cyclic covariate. *Mathematics and Statistics*, *8*(6), 762–772. https://doi.org/10.13189/ms.2020.080617.

Naveau, P., Hannart, A., & Ribes, A. (2020). Statistical methods for extreme event attribution in climate science. *Annual Review of Statistics and Its Application*, *7*(1), 89–110. https://doi.org/10.1146/annurev-statistics-031219-041314.

Nazeri Tahroudi, M., Ramezani, Y., De Michele, C., & Mirabbasi, R. (2021). Flood routing via a copula-based approach. *Hydrology Research*, *52*(6), 1294–1308. https://doi.org/10.2166/nh.2021.008.

Orsini, F., Gecchele, G., Gastaldi, M., & Rossi, R. (2020). Large-scale road safety evaluation using extreme value theory. *IET Intelligent Transport Systems*, *14*(9), 1004–1012. https://doi.org/10.1049/iet-its.2019.0633.

Portet, S. (2020). A primer on model selection using the Akaike Information Criterion. *Infectious Disease Modelling*, *5*, 111–128. https://doi.org/10.1016/j.idm.2019.12.010.

Ramadhani, F. A., Nurrohmah, S., & Novita, M. (2017). Extreme value theory (EVT) application on estimating the distribution of maxima. *AIP Conference Proceedings*, *1862*(1). https://doi.org/10.1063/1.4991260.

Rypkema, D., & Tuljapurkar, S. (2021). Modeling extreme climatic events using the generalized extreme value (GEV) distribution. In *Handbook of Statistics* (Vol. 44, pp. 39–71). Elsevier. https://doi.org/10.1016/bs.host.2020.12.002.

Szigeti, M., Ferenci, T., & Kovács, L. (2020). The use of block maxima method of extreme value statistics to characterise blood glucose curves. *2020 IEEE 15th International Conference of System of Systems Engineering (SoSE)*, 433–438. https://doi.org/10.1109/SoSE50414.2020.9130427.

Tian, A., Shu, X., Guo, J., Li, H., Ye, R., & Ren, P. (2023). Statistical modeling and dependence analysis for tide level via multivariate extreme value distribution method. *Ocean Engineering*, *286*, https://doi.org/10.1016/j.oceaneng.2023.115616.

Towler, E., Llewellyn, D., Prein, A., & Gilleland, E. (2020). Extreme-value analysis for the characterization of extremes in water resources: A generalized workflow and case study on New Mexico monsoon precipitation. *Weather and Climate Extremes*, *29*, https://doi.org/10.1016/j.wace.2020.100260.

Upadhyay, R. K. (2020). Markers for global climate change and its impact on social, biological and ecological systems: A review. *American Journal of Climate Change*, *9*(03), 159. https://doi.org/10.4236/ajcc.2020.93012.

Zhang, Z., Huser, R., Opitz, T., & Wadsworth, J. (2022). Modeling spatial extremes using normal mean-variance mixtures. *Extremes*, *25*(2), 175–197. https://doi.org/10.1007/s10687-021-00434-2.

Zhong, M., Zeng, T., Jiang, T., Wu, H., Chen, X., & Hong, Y. (2021). A copula-based multivariate probability analysis for flash flood risk under the compound effect of soil moisture and rainfall. *Water Resources Management*, *35*, 83–98. https://doi.org/10.1007/s11269-020-02709-y.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., & Mahecha, M. D. (2020). A typology of compound weather and climate events. *Nature Reviews Earth & Environment*, *1*(7), 333–347. https://doi.org/10.1038/s43017-020-0060-z.