University of Birjand

Water Harvesting Research Vol. 7, No. 2, Autumn & Winter 2024, Original paper p. 258-276

10.22077/JWHR.2024.8502.1159

Bivariate Frequency Analysis of Extreme Sea Level with Rainfall and Temperature in New York City

Ali Razmi[a], Emad Mahjoobi[b&*], Heydar Ali Mardani-Fard[c], Zahra Zahmatkesh[d]

[a]PhD Candidate, Department of Water and Environmental Engineering, Shahrood University of Technology, Shahrood, Iran.
[b]Assistant Professor, Department of Water and Environmental Engineering, Shahrood University of Technology, Shahrood, Iran.
[c]Assistant Professor, Department of Mathematics, Yasouj University, Yasouj, Iran.
[d]Post-Doctoral Fellow, Department of Civil Engineering, McMaster University, Hamilton, Ontario, L8S4L8, Canada.

*Corresponding Author, E-mail address: emahjoobi@shahroodut.ac.ir
Received: 30 October 2024/ **Revised**: 05 December 2024/ **Accepted**: 20 December 2024

Abstract

Climate change negatively impacts hydrologic patterns, affecting rainfall, temperature extremes, and sea level rise. Long-term averages of these variables may shift over time due to climate change effects. This study conducted trend analysis on rainfall, maximum and minimum temperature, and water level data from Manhattan, Central Park, and Battery Park stations to identify significant changes in means. The Partial Mann-Kendall test was employed for trend analysis. Frequency analysis utilized common probability distribution functions, including Generalized Extreme Value (GEV), normal, log-normal, and Log-Pearson distributions, with goodness-of-fit tests like Kolmogorov-Smirnov to identify the most suitable distributions. While flood frequency analysis typically examines rainfall and water levels separately, their combination can significantly influence floodplain delineation. This study aimed to enhance flood frequency analysis by considering joint probability distributions for rainfall and storm surge. The correlations between variables and joint probabilities of extreme water levels and temperatures were explored to assess the potential impacts of global warming on sea level flooding. Copula functions determined the joint probabilities of water levels with rainfall and temperature across various recurrence intervals. The trend analysis results indicated an increase in long-term averages due to climate change. The GEV distribution emerged as the most appropriate function for extreme climate variables. This joint probability distribution analysis underscored the necessity of incorporating both rainfall and water level data in flood frequency assessments.
Keywords: Climate change, Climate variables, Copula, Joint probability, Partial Mann-Kendall.

1. Introduction

Rainfall and water levels in coastal areas are among the most crucial climate variables that can lead to severe flooding and damage. In recent years, significant flood events have occurred worldwide, including in central Europe in 2002 and 2005, and on the East Coast of the United States in 2011 and 2012, resulting in substantial damages (Kron, 2005; Karamouz et al., 2016). In the US, there have been noticeable and increasing trends in annual flood damage since 1934 (Pielke et al., 2002). It is evident that the both the intensity and frequency of the floods have risen over the past century (Pielke et al., 2002; Svensson et al., 2005; Karamouz et al., 2017). The growth of population and infrastructure in flood-prone areas has resulted in increased damages and costs, prompting a need to better understand the primary factors driving such changes. Recent studies have further emphasized the impact of climate change on hydrological extremes. For instance, Zhou et al. (2021) demonstrated that climate change significantly alters precipitation patterns, leading to increased flood risks in urban areas. Similarly, AghaKouchak et al. (2020) highlighted the concurrent occurrence of extreme weather events, such as heavy rainfall and high temperatures, which exacerbate flooding risks.

Climate change has been shown to have a negative impact on the hydrologic cycle, which could be linked to the heightened frequency and intensity of extreme hydrologic events (Milly et al., 2002; Robson, 2002; Razmi, 2012; Goharian et al., 2016). The analysis of hydrologic variables plays a critical role in watershed management and hydrology. Regional frequency analysis is applied in flood-prone areas to estimate design floods (Burn, 1990; Karamouz et al., 2014).

Procedures for hydrological frequency analysis are crucial for predicting extreme hydrological events. Recent advancements in multivariate frequency analysis techniques, particularly the application of copula functions, have improved our ability to model the joint probabilities of extreme events. For example, Li et al. (2023) employed copula methods to analyze the joint distribution of rainfall and temperature, revealing critical insights into the dependencies between these variables and their implications for flood frequency analysis.

Frequency analysis involves determining the likelihood of extreme events occurring (Gilroy and McCuen, 2012). The main objective of frequency analysis is to establish a relationship between the magnitude of these events and their frequency of occurrence, using probability distributions (Chebana et al., 2013). The analysis of hydrological variables relies on several assumptions including data independency, homogeneity and stationarity. These assumptions must be verified before the modeling process in univariate analysis (Chebana et al., 2013).

Considering the significance of managing flood risks in flood-prone areas, including coastal regions, it is necessary to calculate exceedance probabilities of rainfall and water levels. By utilizing a suitable flood frequency analysis, an optimal design can be attained (Saf, 2008). Recent studies have emphasized the importance of joint probability distributions in flood risk assessments. For instance, Gräler et al. (2013) conducted a comprehensive review of multivariate return periods in hydrology, emphasizing the necessity of incorporating copula-based approaches to better understand the interdependencies among hydrological variables.

Exceedance probabilities are typically determined through extreme value analysis of a dataset comprising historical observations (Mudersbach and Jensen, 2010). In the case of hydrologic extremes like extreme temperature, rainfall and water levels, the extreme value theory (EVT) provides a solid theoretical framework, extensively discussed in literature (Hawkes et al., 2008; Gumbel, 2012).

The integration of copula functions in this context allows for a more nuanced understanding of the relationships between multiple hydrological variables, which is crucial for effective flood risk management. Recent advancements in copula methodologies have facilitated the modeling of complex dependencies among variables, enabling more accurate predictions of extreme events and their impacts on flood risk (Favre et al., 2021; Xu et al., 2014).

Over the past decade, there has been a growing interest in conducting multivariate analysis on hydrological variables. Trend analysis of these variables is commonly conducted to evaluate the impact of human activities on the environment, taking into account natural fluctuations in temperature, rainfall, and other factors that may influence the variables being studied (Libiseller and Grimvall, 2002). Understanding these trends is crucial for identifying spatial and temporal changes in climatic data, as well as for gaining insights into the state of water resources for future development and sustainable management, and for creating plans to stabilize environmental conditions (Fathian et al., 2016).

In a study by Libiseller and Grimvall (2002), the effectiveness of the Partial Mann-Kendall (PMK) test was assessed in examining water quality trends. The PMK test is a trend analysis method that determines the critical region based on the conditional distribution of one MK statistic for monotone trends, given a set of other MK statistics. Chebana et al. (2013) focused on multivariate hydrological frequency analysis and highlighted the importance of utilizing nonparametric tests for detecting monotonic trends in water quality data. Their findings indicated that using both univariate and multivariate tests can help in capturing different trend signals and selecting appropriate models.

Fathian et al. (2016) investigated trends in hydrological and climatic time series data from the Urmia Lake basin in Iran using four variations of the Mann-Kendall approach: 1) the original Mann-Kendall test; 2) the Mann-Kendall test that considers the impact of lag-1 autocorrelation; 3) the Mann-Kendall test that takes into account autocorrelations or sample size effects; and 4) the Mann-Kendall test that considering the Hurst coefficient.

While flooding can occur in urban areas far from the coasts due to intense surface streamflow, coastal areas face a higher risk of flooding due to the combination of storm events and extreme water levels (Razmi, 2012; Karamouz et al., 2017). Researchers have suggested using joint probability distributions to account for the simultaneous occurrence of rainfall and storm events with extreme water levels in coastal areas (Song et al., 2004; Cong and Brady, 2012; Golian et al., 2012; Graeler et al., 2013; Roussas, 2013).

In many of these studies, Copula joint probability distributions have been employed. For instance, Xu et al. (2014) utilized Copula for analyzing the joint occurrences of extreme precipitation and storm tide. By examining the joint return period of extreme precipitation and storm tide, they were able to propose a design standard for enhancing future flooding preparedness.

This paper aims to establish a framework for conducting frequency analysis on hydrological variables, specifically temperature, rainfall, and water levels. The study utilizes historical extreme climate data from southern Manhattan, a coastal area of New York City, and employs common analysis methods such as the MK test to identify trends in the extreme data.

The joint probability of water levels with rainfall and temperature is calculated for various recurrence intervals ranging from two to 1,000 years. These probabilities are then compared with the severity of individual events to assess the impact of different combinations of climate variables on the hydrologic system.

The paper is structured as follows: The next section introduces and presents the study area and historical observed data. The methodology is then described by presenting the proposed flowchart. Following this, the results are provided and discussed. Finally, a summary and conclusion are given.

2. Materials and Methods

2.1. Study area

New York, the largest state in the United States, includes the borough of Manhattan, which is the most densely populated area of New York City (NYC) with the area of 87.46 km^2 and a population of 1.619 million in 2012. Manhattan is surrounded by several waterways and faces the threat of flooding. Since the 17th century, more than 20 storm events have impacted NYC. Recent storm events on east coast of the United States that have affected Manhattan include hurricane Irene in 2011 and super storm Sandy in 2012.

Irene was one of the most destructive and costliest cyclones o hit NYC, while Sandy was the largest storm ever recorded in the Atlantic Basin. Data from the Battery Park station in southern Manhattan (Fig. 1) shows that the recorded water level from Sandy reached up to 5.24 m based on the Station Datum (STD).

Trend detection and frequency analysis are conducted on data related to rainfall, temperature, and water levels in Manhattan. The data is collected from the Central Park and the Battery Park stations, as shown in Table 1.

2.2. Methodology

The flowchart illustrating the proposed methodology of this study is presented in figure 2. The methodology consists of three primary stages: data collection, trend analysis, and frequency analysis. A detailed description of these stages is provided in this section.

2.3. Data collection and processing

Rainfall, water level, and temperature data time series on a daily basis (representing cumulative precipitation within a 24-hour period), including maximum and minimum temperature values, are utilized in this study for statistical examination. The data is analyzed to generate a time series of extreme annual data, which is subsequently employed for additional analysis.

2.4. Trend analysis

Climate change has been proven to have the potential to impact various aspects of the hydrological cycle, such as rainfall, temperature and water levels. Changes in these hydrological variables can affect flood timing, magnitude and frequency, particularly in coastal areas (Laz and Rahman, 2014). Therefore, it is necessary to investigate trends in hydrological variables when designing hydraulic structures for future operation.

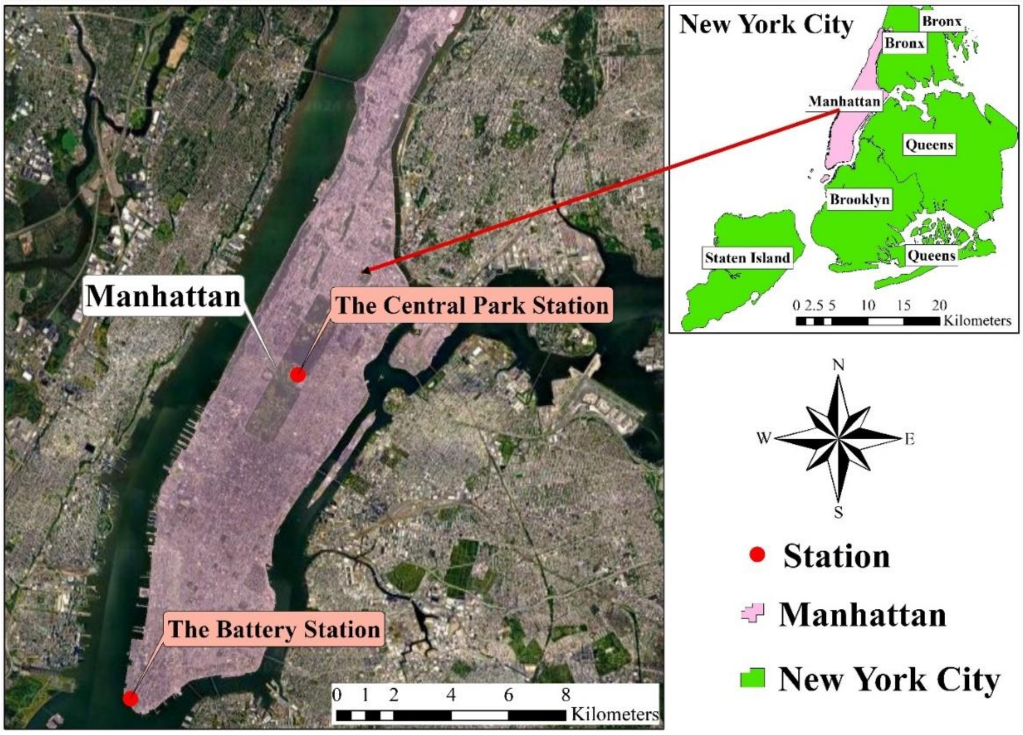

Fig. 1. Location of the Battery Park and Central Park stations in the Manhattan

Table 1. Data sets and characteristics of the stations

Station	Longitude	Latitude	Elevation (m)	Record period	Data
Battery Park	74°00'48"	40°42'00"	15.24	1920-2020	Water level (m)
Central Park	73°57′55″	40°46′56″	39.6	1920-2020	Rainfall (mm), Maximum and Minimum Temperature (ºC)

Several methods exist for testing trends in data. The non-parametric rank-based Mann-Kendall (MK) test is commonly used in climate, hydrology, and water resources studies (Mann, 1945; Kendall, 1975; Burn et al., 2004). Non-parametric tests are suitable for non-normally, distributed, nonlinear and censored data commonly seen in hydrological and metrological time series (Laz and Rahman, 2014). It is important to note that the MK test could be uni- or multi- variate. When weather conditions may influence the time series of the variable being considered, the most appropriate choice is the Partial MK (PMK) test.

The PMK test is a statistical method used to identify trends in a primary variable, such as water level, while accounting for the influence of another variable, like rainfall or temperature. This is important because environmental factors can affect the variable of interest. For example, if we want to analyze the trend in water levels, we might also consider how rainfall patterns have changed over time. The PMK test adjusts for these influences by examining the relationship between the water level and the covariate (rainfall or temperature) (Wahlin and Grimvall, 2010).

This adjustment helps us understand whether the observed trend in water levels is due to changes in rainfall or if it is a separate trend. The test calculates a statistic based on the ranks of the data, which allows us to determine if the trend is statistically significant (Libiseller and Grimvall, 2002).

In our methodology, we first apply the PMK test to analyze the water level data in relation to rainfall.

This means we look at how changes in rainfall over time might influence water levels. Next, we conduct a similar analysis using temperature as the covariate. By doing this, we can see if temperature changes also affect water levels. This two-step approach allows us to isolate the effects of each variable on water levels, providing a clearer picture of the trends.

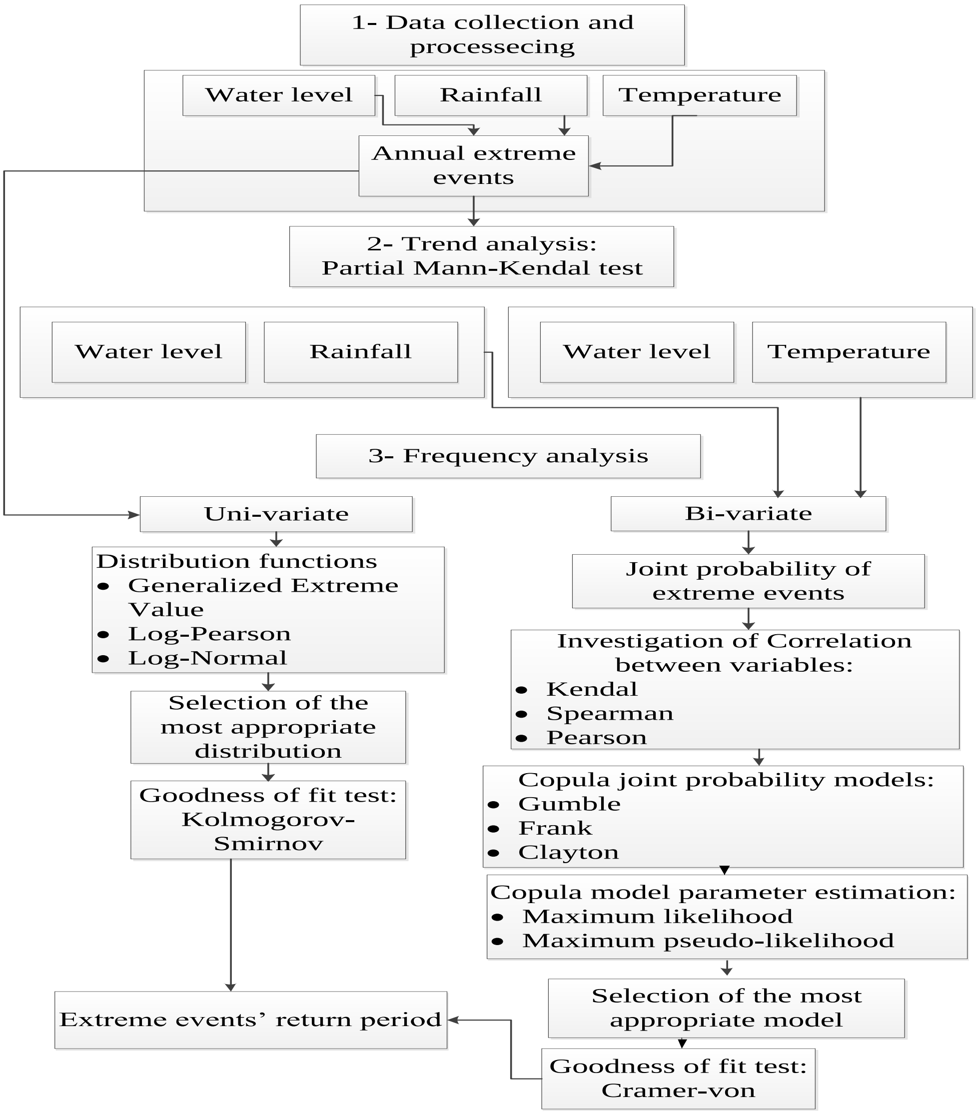

Fig. 2. Methodology flowchart to estimate joint return period of extreme hydrologic variables

The univariate Mann-Kendall test detects monotonic trends. This test involves pairwise comparisons of all observations within a time series dataset. This test is commonly used to assess changes in the central value (median) of water level over time. In the context of this test, the null hypothesis (H0) posits that there is no trend in the data. Conversely, the alternate hypothesis (H1) indicates the presence of a trend. The MK test statistic for a time series $\{x_k, k = 1, 2, \ldots, n\}$ is calculated as:

$$s = \sum_{i<j}^{n} sgn(x_j - x_i) \quad (1)$$

$$sgn(x_j - x_i) = \begin{cases} +1 & \text{if } (x_j - x_i) > 0 \\ 0 & if\ (x_j - x_i) = 0 \\ -1 & if\ (x_j - x_i) < 0 \end{cases} \quad (2)$$

where n represents the number of observations in a data time series, x_i and x_j refer to the i^{th} and j^{th} observations. If there are no ties among the observations, and no trend exists, the test statistic follows an asymptotically normal distribution. The expected value of the test statistic is zero, and its variance can be calculated using a specific formula:

$$E(s) = 0$$
$$Var\,(s) = \frac{n(n-1)(2n+5)}{18} \quad (3)$$

The significance levels (p-values) obtained are typically determined under the assumption that the statistic s follows an approximately normal distribution, given the aforementioned conditions, and that the null hypothesis is true. This implies that all permutations of the observed data are equally likely.

If the response variable is recorded across several classes (e.g., *w*), the seasonal MK test, also known as the Hirsch-Slack test, is computed. To accomplish this, the data are first divided into ω sub-series, with each series corresponding to a specific season. The test statistic for this seasonal analysis is calculated by summing over all classes as follows:

$$s = \sum_{j=1}^{w} \sum_{k<l} sgn(x_{lj} - x_{kj}) \quad (4)$$
$$j = 1, \dots, w$$

where s is asymptotically normally distributed with a mean of zero and a variance of:

$$VAR(s) = \frac{1}{18}\{n(n-1)(2n+5) - \sum_{j=1}^{w} t_j\,(t_j - 1)(2t_j + 5)\} \quad (5)$$

The test statistic is estimated as:

$$Z_m = \begin{cases} \frac{s-1}{\sqrt{VAR(s)}} & if\ s > 0 \\ 0 & if\ \ s = 0 \\ \frac{s+1}{\sqrt{VAR(s)}} & if\ \ s < 0 \end{cases} \quad (6)$$

where *w* represents the number of classes, t_j denotes the number of data points in the j^{th} class. Negative values of Z_M indicate a decreasing trend, while positive values indicate an increasing trend in the time series data. In our analysis, we test the null hypothesis, which states that there is no trend in the data. We reject this hypothesis if our calculated statistic, Z_M, is significantly different from zero.

A significance level of 0.05 means that there is only a 5% chance that we would see this result if the null hypothesis were true. If $|\,Z_M\,| > 1.96$, we conclude that there is a significant trend in the data. This statistical approach helps us determine whether the changes we observe in water levels are meaningful or simply due to random fluctuations (Burn et al., 2004).

The covariance between two MK statistics is obtained as follows:

$$Cov(s_j s_g) = \frac{[s_{jg} + 4\sum_{m=1}^{n} R_{mj} R_{mg} - n(n_j+1)(n_g+1)]}{3} \quad (7)$$

where n_j and n_g reoresent the number of non-missing observations for calss *j* and *g*, respectively, and:

$$s_{jg} = \sum_{m<n} sign[(x_{nj} - x_{mj})(x_{ng} - x_{mg})] \quad (8)$$

Matrix *R* corresponds to the observations in the time series dataset, where the non-missing observations for each class are ranked among themselves. The rank of the m^{th} element within the i^{th} class is denoted by:

$$R_{mj} = \frac{[n_j + 1 + \sum_{k=1}^{n} sign(x_{mj} - x_{kj})]}{2} \quad (9)$$

To assing the midrank of $\frac{n_j+1}{2}$ to missing values, the term $sign(x_{mj} - x_{kj})$ is set to zero if either x_{mj} or x_{kj} is missing.

For the PMK test, observations for each variable (in this paper, one dependent variable, such as water level, and one independent variable, such as rainfall or temperature) are divided to several classes to account for seasonality, if necessary. For each sequence, the MK statistic is calculated and represented by the vector T. The matrix Γ corresponds to the variances and covariances between the variables. The test statistic can be summed over the classes for either the dependent variable alone or for both the dependent and the independent variables. The test statistic $T = [T_\alpha, T_\beta]^T$ where α and β represent the response (dependent) and explanatory (independent) variables, respectively, is asymptotically multivariate normally distributed with a mean vector $\mu = \begin{bmatrix}\mu_\alpha \\ \mu_\beta\end{bmatrix}$ and a covariance matrix $\begin{bmatrix}\Gamma_{\alpha\alpha} & \Gamma_{\alpha\beta} \\ \Gamma_{\beta\alpha} & \Gamma_{\beta\beta}\end{bmatrix}_{\delta\times\delta}$. Therefore, the conditional distribution of S_α given that $S_{\beta=} s_\beta$ is multivariate normal with:

$$E(S_\alpha | S_{\beta=} s_\beta) = \mu_\alpha + \Gamma_{\alpha\beta}\Gamma_{\beta\beta}^{-1}(s_\beta - \mu_\beta) \quad (10)$$
$$Var(S_\alpha | S_{\beta=} s_\beta) = \Gamma_{\alpha\alpha} - \Gamma_{\alpha\beta}\Gamma_{\beta\beta}^{-1}\Gamma_{\beta\alpha} \quad (11)$$

Under the null hypothesis of no trend in any of the variables, the mean vector is given by

$\mu = \begin{bmatrix}\mu_\alpha \\ \mu_\beta\end{bmatrix} = \begin{bmatrix}0 \\ 0\end{bmatrix}$, which simplifies Equation 10 to:

$$E(S_\alpha | S_{\beta=} s_\beta) = \Gamma_{\alpha\beta} \Gamma_{\beta\beta}^{-1}(s_\beta) \quad (12)$$

The PMK test statistic is expressed as follows (Libiseller, 2002):

$$Z_{PMK} = \frac{S_\alpha - E(S_\alpha | S_{\beta=} s_\beta)}{\sqrt{Var(S_\alpha | S_{\beta=} s_\beta)}} \quad (13)$$

which is normally distributed with a mean of zero and a standard deviation of one (Wahlin and Grimvall, 2010).

2.5. Frequency analysis

2.5.1. Univariate frequency analysis

Standard design approaches in hydrology often rely on univariate frequency analysis (FA) methods (Graeler et al., 2013). FA involves four key steps: 1) conducting descriptive and exploratory analyses, including outlier detection; 2) verifying assumptions like stationarity, homogeneity, and independence; 3) modeling and estimation; and 4) performing risk analysis. In hydrological studies, the assumptions of data independence and stationarity are foundational for traditional frequency analysis methods.

Independence implies that the occurrence of one event does not influence another, while stationarity suggests that the statistical properties of the data do not change over time. Many studies have historically relied on these assumptions to simplify the modeling of extreme events, particularly in univariate and bivariate analyses (Gilroy and McCuen, 2012; Karamouz et al., 2014). In univariate FA methods, frequency analysis is conducted on a specific hydrologic variable, such as rainfall, without considering the probability occurrence of other variables.

Various distribution functions have been proposed and utilized for rainfall, water level, and temperature. In this study, three distribution functions -log-normal, Log-Pearson Type III, and generalized extreme value (GEV) are examined for the selected climate variables.

2.5.2. Log-normal probability distribution function

The Log-Normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. A random variable that follows a log-normal distribution can only take positive real values. The CDF of a random variable x that adheres to the Log-Normal distribution is expressed as follow:

$$F(x, \mu_n, \sigma_n) = \frac{1}{2}\left[1 + erf\left(\frac{\ln x - \mu_n}{\sigma_n\sqrt{2}}\right)\right] \quad (14)$$

In this equation, μ_n and σ_n represent the mean and standard deviation of the natural logarithm of the variable, respectively, and these parameters define the distribution (William, 2003).

2.5.3. Log-Pearson type iii probability distribution function

The Log-Pearson Type III distribution is employed to calculate the frequency of maximum events when all events are log-normally distributed. Events are log-normally distributed when they result from the product of a large number of independent random variables (Foster, 1924). One advantage of this technique is that allows for the extrapolation of values for events with return periods that extend well beyond the range of observed data. This method is the standard approach utilized by federal agencies in the United States (Bedient and Huber, 2002).

The Log-Pearson Type III distribution is analogous to the normal distribution, characterized by three parameters: shape (α), scale (σ), and location (μ). If Γ denotes the gamma function, then the CDF for this distribution, given a random variable x, is expressed as follows:

$$F(x, \alpha, \mu, \sigma) = \frac{\Gamma_{(\ln(x)-\mu)/\sigma}(\alpha)}{\Gamma(\alpha)} \quad (15)$$

2.5.4. Generalized extreme value probability distribution function

The GEV distribution is a probability distribution function developed within the framework of extreme value theory. This distribution encompasses three sub-distributions: Gumbel, Fréchet, and Weibull, allowing it to model a wide range of extreme behaviors depending on the tail characteristics of the data. This flexibility is particularly advantageous when dealing with climatic data, which can exhibit varying degrees of tail heaviness (Katz et al., 2002). The selection of the GEV distribution arises from its superiority in fitting various climatic datasets, which is particularly relevant in the context of hydrological and meteorological extremes. The GEV distribution has been extensively applied in hydrological studies for flood frequency analysis, demonstrating its effectiveness in predicting the occurrence of rare events. Recent studies have reinforced its applicability in various climatic contexts, highlighting its robustness in capturing the dynamics of extreme hydrological phenomena (Salvadori et al., 2007; Karamouz et al., 2014).

The GEV distribution is characterized by three parameters: a shape parameter (ξ), a location parameter (μ), and a scale parameter (σ). This distribution can capture the tail behavior of extreme values effectively. For example, when assessing annual maximum daily rainfall, the GEV's flexibility allows it to adapt to datasets that show both heavy tails and trends, which are common in climatic data impacted by climate change (Kharin et al., 2013). The CDF of the GEV distribution for the considered variable (x) is calculated as follows:

$$F(x,\zeta,\mu,\sigma) = exp\left\{-\left[1+\zeta(\frac{x-\mu}{\sigma})\right]^{-\frac{1}{\zeta}}\right\} \quad (16)$$

2.6. Goodness-of-fit test to select the most appropriate distribution

To assess the compatibility of the time series data for each variable with the specified probability distribution functions, the nonparametric Kolmogorov-Smirnov (KS) goodness of fit test is employed (Chakravart et al., 1967). This test evaluates how well the distribution functions align with the climate data. The KS statistic (KSD) quantifies the distance between the empirical distribution function of the data ($F_i(x)$) and the cumulative distribution function of the distribution of interest. The KS test is framed in terms of two hypotheses and a statistic (KSD).

The null hypothesis posits that the data conform to the theoretical distribution, while the alternative hypothesis asserts that they do not. The KSD statistic for a given cumulative distribution function is calculated as follows:

$$KSD_i = \sup_x |F_i(x) - F(x)| \quad (17)$$

where $\sup_x$ represents the supremum of the set of the distances. For extreme annual data within each time series, collected over a continues time period and sorted in a descending order such that $x_1 < x_2 <\ldots< x_n$, the empirical cumulative distribution function for the i^{th} data point is defined as:

$$F_i(x) = \frac{1}{i} \times n(i) \qquad 1 \le i \le n \quad (18)$$

where $n(i)$ is the number of data points that are less than x_i. For the empirical distribution, if the value of the variable originates from a specified theoretical distribution, then KSD_iconverges to 0. This null hypothesis is rejected at the chosen significance level (α) if the KSD exceeds the critical value obtained from Kolmogorov-Smirnov table (Chakravart et al., 1967).

2.7. Extreme values' return period

After applying the KS test to identify the most suitable distribution that best fits the historical extreme annual data, the annual exceedance probability (E_P)-defined as the probability that an event is equaled or exceeded in any given year-is determined. To achieve this, the CDF of the fitted distribution is utilized. The E_P for an event with a magnitude of x_D is estimated as follows:

$$E_P = P\{x \ge x_D\} = 1 - P\{x < x_D\} = 1 - F(x_D) \quad (19)$$

where P represents the probability and F denotes the CDF of the fitted distribution. The relationship between the return period (T) and the exceedance probability (E_P) is expressed as:

$$T = \frac{1}{E_P} \quad (20)$$

A T-year extreme event is defined as an event that is equaled or exceeded, on average, once every T-year. The optimal distribution for the historical data is employed to estimate extreme values of the variable across different return periods.

2.8. Multivariate frequency analysis

The application of simple univariate FA methods may result in a significant underestimation of the risk associated with a given extreme event. In hydrology, the analysis of multivariate events is particularly important. A wide range of methods has been employed for univariate FA of extreme events; however, multivariate FA is seldom implemented. One reason for this may be the limited availability of multivariate models that adequately represent extreme values. Commonly used multivariate distributions in hydrologic studies include normal, bivariate exponential, bivariate gamma and bivariate extreme value distributions. These methods have several drawbacks, including: (1) the requirement for identical families for each marginal distribution, (2) ambiguity in extending beyond the bivariate case, and (3) the use of parameters from marginal distributions to model the dependence between random variables (Favre et al., 2004).

In contrast, a multivariate distribution that does not exhibit these drawbacks is known as a copula (Sklar, 1996). Copulas account for all dependencies between two random variables of interest. A critical step before fitting copulas is to investigate whether there is a correlation between the variables. Various tests can be employed for this purpose.

In this study, three distinct tests are employed to assess whether there is a correlation among water level and rainfall, as well as between water level and temperature.

2.8.1. Kendall correlation coefficient

The correlation coefficient between variables x and y, based on the Kendal test, is evaluated as follows:

$$\tau_n = \binom{n}{2}^{-1} \sum_{i<j} sign\,[(x_i - x_j)(y_i - y_j)] \quad i,j = 1,2,\dots,n \tag{21}$$

where n represents the number of data points. The sign of the correlation coefficient is determined as follows: it is +1 if $[(x_i - x_j)(y_i - y_j)]$ >0, indicating a positive correlation, and -1 if $\left[(x_i - x_j)(y_i - y_j)\right] < 0$, indicating a negative correlation.

2.8.2. Spearman correlation coefficient

The Spearman's correlation coefficient for the u^{th} component of the data is defined as follows:

$$S^{(u)} = \sum_{i=1}^{n} \left(i - \frac{n+1}{2}\right)\left(rank\left(x_i^{(u)}\right) - \frac{n+1}{2}\right) \quad u = 1,\dots,d \tag{22}$$

The $rank\left(x_i^{(u)}\right)$ is denoted as rank of $x_i^{(u)}$ in the time series of observed data $x_1^{(u)}, \dots, x_n^{(u)}$.

2.8.3. Pearson correlation coefficient

For any given pair of variables x and y, the Pearson correlation coefficient is calculated as follows:

$$r_{xy} = \frac{(\sum x - \bar{x})(\sum y - \bar{y})}{\sqrt{\sum x^2 \sum y^2}} \tag{23}$$

where $\bar{x}$ and $\bar{y}$ represent the long-term averages of the data time series for variables x and y, respectively.

2.9. Copula joint probability distribution

Copulas are defined as "multivariate distribution functions that link joint probability distributions to their one-dimensional marginal distributions" (Nelsen, 2006). Marginal distributions that are uniform within the interval [0, 1] can be connected using copula multivariate distribution functions. To achieve this, the marginal univariate distributions must first be identified within that interval. Subsequently, a copula function correlates the variables by constructing a multivariate distribution utilizing the copula parameter (θ).

Copulas are particularly valuable for implementing efficient algorithms for simulating joint distributions (Favre et al., 2004). One of the primary advantages of copulas is that they allow for the separate examination of the marginal properties and dependence structure of the variables (Xu et al., 2014). To obtain a joint CDF F(x,y) for random variables, *X* and *Y*, with marginal distributions defined as u_1 =F_X(x)=Pr(X≤x) and u_2=F_Y(y)=Pr(Y≤y), a copula function *C* is utilized:

$$F(x,y) = C(u_1, u_2, \theta) = C(F_X(x), F_Y(y)) \tag{24}$$

where *F(x,y)=Pr(X≤x,Y≤y)* represents the joint probability distribution function of *X* and *Y* (Sklar, 1996). If $F_X(x)$ and $F_Y(y)$ are continuous, then the copula function *C* is unique. To estimate the probability density function (PDF) of $F_X(x)$ and $F_Y(y)$, the following expression is employed:

$$f(x,y) = D\big(F_X(x), F_Y(y)\big) f_X(x), f_Y(y) = D(u_1, u_2; \theta) f_X(x), f_Y(y) \tag{25}$$

where *f(x, y)* is the bivariate PDF, and $f_X(x)$ and $f_Y(y)$ are the marginal PDFs of X and Y. The term D (u_1, $u_2;\theta$) represents a bivariate copula density function:

$$D(u_1, u_2; \theta) = \frac{\partial^2 C(u_1, u_2; \theta)}{\partial u_1 \partial u_2} \quad (26)$$

There are various types of copulas, including elliptical copulas, Archimedean copulas, and copulas with quadratic sections.

2.9.1. Archimedean copulas

In contrast to copulas based on multivariate distributions, there are several copulas that exhibit a relatively simple form. These copulas belong to the class of Archimedean copulas, which interpolate between specific dependency structures, namely counter monotonicity, independence, and comonotonicity. Archimedean copulas possess several advantageous features, such as symmetric and associativity. Furthermore, the computation of measures of dependence is simplified in Archimedean copulas. However, a limitation of these copulas is that they are symmetric (Favre et al., 2004).

In this study, four commonly used Archimedean copulas- Gumbel, Gaussian, Clayton and Frank-are employed to analyze pairs of water level and rainfall, as well as water level and temperature. These copulas are compared to identify the best fit for the pairs of considered variables. The expressions for these copulas are provided in Table. It should be noted that all of these functions are one-parameter copulas (Liu et al., 2015).

Table 2. The applied bivariate copula functions

Copula	Copula function, C (u1, u2;θ)
Gaussian	$\int_{-\infty}^{\varphi^{-1}(u_1)} \int_{-\infty}^{\varphi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{1/2}} exp\left[-\frac{x^2 - 2\rho xy + y^2}{2(1-\theta^2)}\right] dydx$
Clayton	$(u_1{}^{-\theta} + u_2{}^{-\theta} - 1)^{-1/\theta}$
Frank	$-\frac{1}{\theta}\ln\left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right]$
Gumbel	$C(u_1, u_2) = exp\{-\left((-\ln(u_1))^{\theta} + (-\ln(u_2))^{\theta}\right)^{1/\theta}\}$

2.9.2. Selection of the most appropriate Archimedean copula

A critical step in multivariate FA using copulas is to identify the most appropriate function that best fits the data. For this purpose, various goodness of fit test can be employed. In this study, the Cramér-von Mises (CvM) test is utilized to determine the best copula function to fitting the joint data time series. The CvM statistic (S_n) is expressed as follows:

$$S_n = \int_u \Delta C_n(u)^2 \, d\, C_n(u) \quad (27)$$

$$\Delta C_n = \sqrt{n}(C_n - C_{\theta_n}) \quad (28)$$

where C_n is the empirical copula, n is the sample size, and C_{θ_n} is the parametric copula estimated for a sample size n. The null hypothesis states that the parametric copula adequately fits the data (i.e., H_0: $C_n \in C_{\theta_n}$). P-values greater than the significance level indicate acceptance of the null hypothesis, otherwise, it is rejected (Genest and Werker, 2002).

Therefore, among the considered copula functions, the one with the highest p-value (i.e., the smallest S_n) is preferred and selected for further analysis (Salvadori and Michele., 2004; Genest et al., 2009; Madadgar and Moradkhani, 2013; Madadgar, 2013).

2.9.3. Estimation of the model's parameter: maximum likelihood estimation (MLE)

Likelihood-based methods are frequently employed for parameter estimation. In this study, MLE is utilized to estimate the parameters of the selected copula model. The likelihood is expressed as follows:

$$l(\theta) = \prod_{t=1}^{n} g(X_t|\theta) \quad (29)$$

where g represents the probability density function of the variable X_t. n is the number of data points, and θ denotes the model parameter. The maximum likelihood estimate (MLE) of θ is defined as the value of θ that maximizes the likelihood function, denoted as l (θ) (Wang et al., 2009).

2.9.4. Estimation of the model's parameter: maximum pseudo-likelihood (MPLE)

Pseudo-likelihood is an approximation of the joint probability distribution of a time series of a random variable and is calculated as follows:

$$\ln l(\theta) = \sum_{i=1}^{n} \ln c_\theta(\widehat{U}_\iota) \quad (30)$$

where $\widehat{U}_\iota = (\widehat{U}_{i,1}, \dots, \widehat{U}_{i,d})$ represents the pseudo observations obtained from the time series of $X_i = (X_{i,1}, \dots, X_{i,d})$ using the formula: $\widehat{U}_{i,j} = \frac{R(X_{i,j})}{n+1}$. Here, $R(X_{i,j})$ denotes the rank of $X_{i,j}$ among the values $X_{i,1}, \dots, X_{i,d}$. The term c_θ refers to the selected copula model with parameter θ (Shih and Louis, 1995; Kojadinovic and Yan, 2010).

2.10. Return period of joint extreme events

The multivariate return period is derived using the concept of copulas (Nelsen, 2006; Cong and Brady, 2012; Li et al., 2013; Keerthirathne and Perera, 2015). For two variables, *X* and *Y* (e.g., rainfall and water level or temperature and water level), the joint survival distribution can be defined as:

$$\bar{F}(x, y) = \Pr(X > x, Y > y) = \bar{C}(\bar{F}_X(x), \bar{F}_Y(y)) \quad (31)$$

where $\bar{F}_X = 1 - F_X$ and $\bar{F}_Y = 1 - F_Y$ are the marginal survival functions of X and Y, respectively, and $\bar{C}$ represents the survival copula. By utilizing the survival critical layer with probability t (where $t \in (0,1)$), we can obtain a survival critical layer on which the data sets for *X* and *Y* share the same probability level *t* (Burn, 1990; Schoelzel and Friederichs, 2008):

$$L_t^{\bar{F}} = \{x, y \in R^d : \bar{F}(x, y) = t\} \quad (32)$$

The survival Kendal return period of simultaneous occurrence of X and Y, denoted as $T_{X,Y}$, is then calculated as follows:

$$T_{X,Y} = \frac{\mu_T}{1 - \bar{T}(t) = \Pr(\bar{F}(x, y) \geq t)} = \frac{\mu_T}{\Pr(\bar{C}(\bar{F}_X(x), \bar{F}_Y(y)) \geq t)} \quad (33)$$

where μ_T is the average interarrival time of the concurrence of X and Y. By inverting the Kendal survival function at the probability level $p = 1 - \frac{\mu}{T}$, the survival layer corresponding to return period T can be estimated as:

$$\bar{q} = \bar{q}(p) = \bar{T}^{-1}(p) \quad (34)$$

where $\bar{q}$ is the survival Kendall quantile of order p. The survival critical layer, denoted as $L_t^{\bar{F}}$, corresponds to the quantile $\bar{q}$, indicating that the combined variables X and Y have a joint return period T (Salvadori and De Michele, 2004; Salvadori et al., 2007; Salvadori et al., 2011; Salvadori et al., 2013; Graeler et al., 2013; Liu et al., 2015).

3. Results and Discussion

Figure 3 illustrates the variation in the time series data. This figure indicates an increasing trend in the data, with the exception of the time series of maximum annual temperature. Subsequently, trend analysis is conducted on the time series of individual variables. Following this, the effects of rainfall and temperature on the trend of water level are examined.

3.1. Trend analysis of climate variables

The results of the Mann-Kendal (MK) test at the 5 % confidence level for the climate variables under consideration are presented in Table 3. In this table, "MAWL" denotes maximum annual water level, and "MDP" represents maximum daily precipitation. Additionally, "Tmin" and "Tmax" correspond to minimum and maximum temperatures, respectively. Based on the Z statistic and P-values, the hypothesis of trend's existence is accepted for all the climate variables, with the exception of Tmax.

3.2. PMK test for multivariate trend analysis

A trend analysis of water level, designated as the dependent variable, was conducted with rainfall and temperature serving as the independent variables over the period from 1920 to 2015. The results are presented in Table 4.

Based on the results presented in Table 3, the value of the test statistics (i.e., Z) for the climate variables exceeds 1.96, indicating that the application of the MK test demonstrates the existence of a trend in the data. However, as shown in Table 4, the results from the PMK test indicate that precipitation has the strongest influence on water levels compared to temperature.

This finding suggests that as rainfall patterns change, they have a more pronounced effect on the water levels in our study area. For instance, if we observe an increase in heavy rainfall events, we can expect to see corresponding increases in water levels, which could lead to flooding. This relationship underscores the importance of monitoring rainfall trends when assessing flood risks. This conclusion is supported by the observation that, when analyzing trends, the inclusion of precipitation alongside water level resulted in the greatest reduction in the value of the test statistic compared to the Z value for water level alone in Table 3.

Following precipitation, minimum and maximum temperatures are ranked in descending order regarding their impact on water level trend.

3.3.Univariate frequency analysis

The Kolmogorov-Smirnov (KS) statistic is utilized to determine the most suitable probability distribution function for fitting climate variables. The results are summarized in Table 5. The KS statistic for the GEV was significantly lower, suggesting a closer alignment with the empirical distribution of the observed data. Recent studies have also demonstrated that the GEV distribution often outperforms other commonly used distributions (e.g., Normal, Log-Normal, and Exponential) in accurately describing extreme value behavior (Su et al., 2020; Zeng et al., 2023). Using the fitted distributions (GEV), values of climate variables in different return periods are estimated and reported in Table 6.

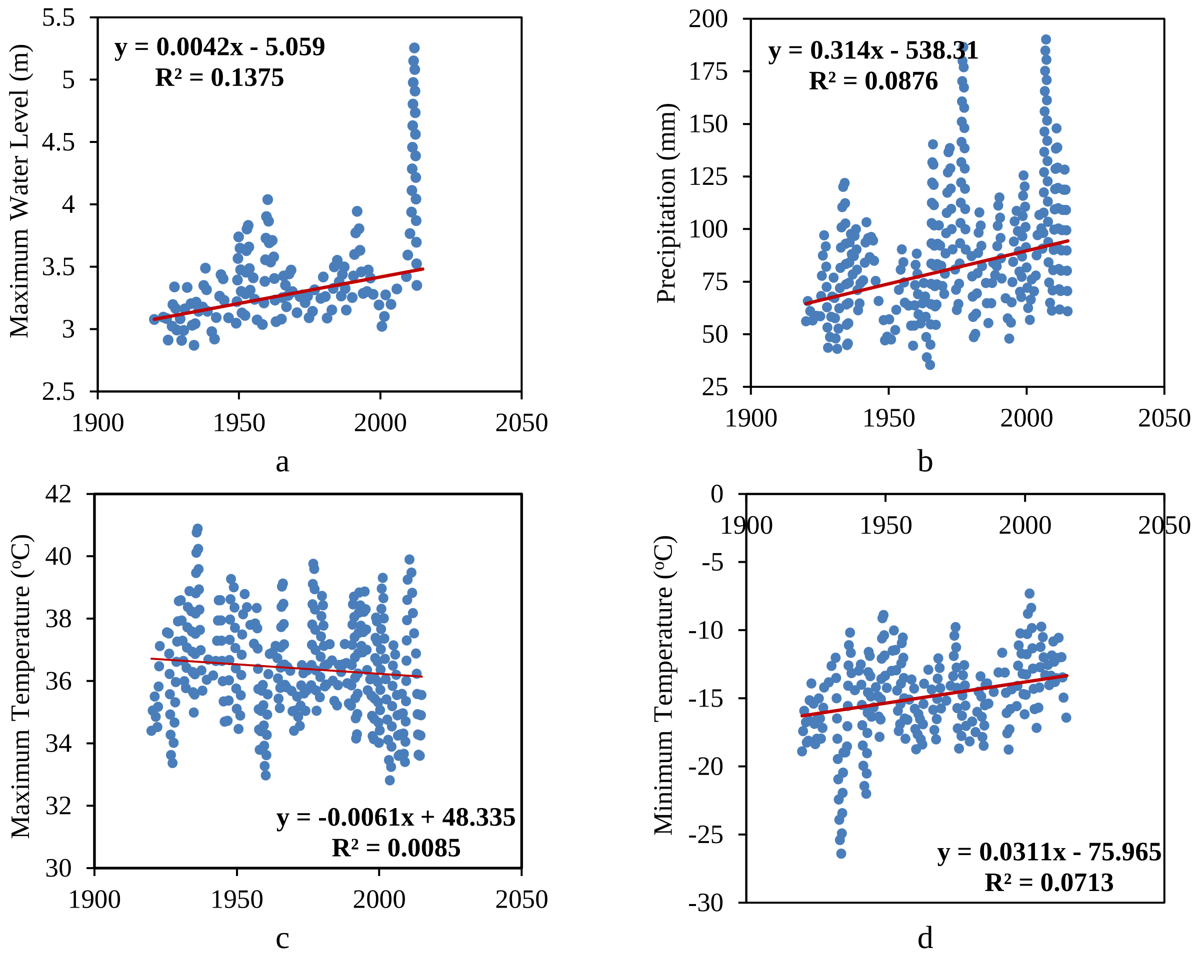

Fig. 3. Variation of a: Maximum annual water level (MAWL), b: Maximum daily precipitation (MDP), c: Maximum temperature (Tmax), and d: Minimum temperature (Tmin)

Table 3. Mann-Kendall test results

Data	Z	Kendall's tau	S	Var(S)	p-value (Two-tailed)	alpha	Sen's slope
MAWL (m)	4.37	0.305	1379	99659.66	< 0.0001	0.05	0.003
MDP (mm)	2.74	0.190	865	99791	0.006	0.05	0.253
Tmin (ºC)	2.28	0.162	718	99286	0.023	0.05	0.026
Tmax (ºC)	-0.96	-0.069	-302	98826.66	0.338	0.05	0

Table 4. PMK test results for multivariate trend analysis

Data	Z	p-value	S	Var (S)
MAWL & MDP	3.71	0.0002	1116.29	90455.5
MAWL & Tmin	4.08	4.47E-05	1274.93	97574.1
MaWL & Tmax	4.32	1.59E-05	1359.49	99247.4

Table 5. K-S test results to investigate different probability distribution functions to climate data

Data	Distribution	Log-pearon Type 3	Log normal	GEV
MAWL (m)	Parameters	α=0.86363, β=0.09407, γ=1.1026	σ=0.08697, μ=1.1839	k=0.17253, σ=0.17921, μ=3.1403
	K-S statistic	0.12341	0.14972	0.07212
MDP (mm)	Parameters	α=120.09, β=0.03191, γ=0.481	σ=0.34782, μ=4.3135	k=0.04415, σ=21.817, μ=65.873
	K-S statistic	0.05669	0.06533	0.05857
Tmin (ºC)	Parameters	No fit	σ=0.02466, μ=4.8831	k=-0.36259, σ=3.3004, μ=-15.812
	K-S statistic	No fit	0.09082	0.0866
Tmax (ºC)	Parameters	α=1697.8, β=0.00122, γ=1.5229	σ=0.05, μ=3.5941	k=-0.22776, σ=1.7921, μ=35.731
	K-S statistic	0.07399	0.07612	0.07168

Table 6. Univariate frequency analysis of climate variables

Return Period (years)	MAWL (m)	MDP (mm)	Tmin (C)	Tmax (C)
2	3.2	73.93	-14.67	36.36
5	3.44	99.7	-11.99	38
10	3.63	117.49	-10.73	38.88
25	3.9	140.82	-9.56	39.8
50	4.1	158.77	-8.92	40.36
100	4.39	177.15	-8.42	40.84
200	4.69	196.04	-8.04	41.24
500	5.13	221.85	-7.66	41.68
600	5.23	227.11	-7.6	41.76
1000	5.52	242.07	-7.45	41.98

3.4. Bivariate frequency analysis

Archimedean Copulas are among the most commonly employed models for the multivariate frequency analysis of hydrologic variables. To utilize copulas effectively, it is essential to investigate the correlation between the variables. Various statistical tests can be applied for this purpose. The results of the correlation analysis between the variables are presented in Table 7.

Based on this table, it is observed that the correlation between water level and both precipitation and minimum temperature is positive, as indicated by the calculated coefficients. Consequently, three copula models will be examined for the joint probability of these variables. It is important to note that when Kendall's coefficients are negative, the Clayton and Gumbel copula functions are not applicable.

Table 7. Correlation coefficient among climate variables

Correlation coefficient	MAWL and MDP	MAWL and Tmin	MAWL and Tmax
Kendal	0.205	0.105	-0.05
Spearman	0.305	0.145	-0.064
Pearson	0.136	0.219	0.022

3.5. Parameter estimation for copula models

The Probability Maximum Likelihood Estimation (PMLE) method is employed to accurately estimate the parameters of the copula models. This approach ensures that the estimation of the model parameters is not influenced by the model itself. The copula model with the maximum value of the Maximum Likelihood Estimation (MLE) is considered the most suitable for joint probability analysis.

3.6. Selection of the most appropriate copula model

The Cramer-von Mises method is utilized to identify the most appropriate copula model. This method incorporates bootstrap and Monte Carlo techniques to generate random data that aligns with the estimated copula. According to this method, the model with the minimum value of Sn will be selected. Furthermore, it is essential that the model is not rejected based on the goodness-of-fit test. The results of this analysis are presented in Table 8.

Since all the copulas under consideration belong to the Archimedean family, there is no significant difference in the performance of the models. However, because the Frank copula does not have the limitations associated with other copulas - specifically, its applicability when the correlation between variables is negative - it has been selected for further analysis (Singh and Strupczewski, 2002; AghaKouchak et al., 2014; Cheng et al., 2016).

3.7. Joint return period

Figures 4 to 6 illustrate the joint return periods of precipitation and water level, minimum temperature and water level, as well as maximum temperature and water level.

In our study, we employed a copula-based approach to analyze the joint probabilities of extreme water levels, rainfall, and temperature, which allows for the modeling of dependencies between variables without the strict assumptions of independence and stationarity.

This approach is supported by recent advancements in multivariate frequency analysis, which recognize the interdependencies among hydrological variables (Favre et al., 2004; Xu et al., 2014).

For example, Xu et al. (2014) conducted a joint probability analysis of extreme precipitation and storm tide, acknowledging that the correlation between these variables can significantly affect flood risk assessments. Their findings underscore the importance of considering variable interdependencies, particularly in coastal regions where extreme events are influenced by multiple climatic factors. In contrast, studies that have assumed independence and stationarity may overlook critical interactions between variables, potentially leading to underestimations of flood risks. For instance, Karamouz et al. (2014) highlighted that static bivariate frequency analyses, which do not account for changing climatic conditions, may yield misleading results in flood risk assessments.

Table 8. Fitted copula models to the joint climate variables and parameter estimation for the models

Variable	Copula model Parameter	Parameter estimation method: Inversion Of Spearman's rho	Inversion Of Kendall's tau	Maximum Likelihood	Maximum Pseudo-Likelihood	Max value for MLE	Std. Error	Sn	P-value	Z value
MAWL and MDP	Gaussian	-	0.31	0.30	-	-	-	0.11	0.4	-
	Clayton	0.52	0.51	0.43	0.43	3.95	0.15	0.052	0.004	2.84
	Frank	1.92	1.91	1.86	1.86	4.43	0.68	0.043	0.006	2.733
	Gumbel	1.26	1.25	1.18	1.18	2.32	0.09	0.074	<2e-16	13.14
MAWL and Tmin	Gaussian	-	0.16	0.20	-	-	-	0.12	0.27	-
	Clayton	0.21	0.23	0.20	0.20	2.18	0.12	0.11	0.098	1.65
	Frank	0.88	0.96	1.08	1.08	1.2	0.63	0.092	0.091	1.69
	Gumbel	1.11	1.12	1.1	1.1	0.87	0.07	0.092	<2e-16	15.57
MAWL and Tmax	Gaussian	-	0.21	0.003	-	-	-	15.2	0	-
	Clayton	-0.08	-0.10	-0.21	-0.21	2.29	0.08	0.20	0.009	-2.62
	Frank	-0.38	-0.49	-0.34	-0.34	0.14	0.57	0.15	0.55	-0.59
	Gumbel	-	1.16	1	-	-	-	15.1	0	-

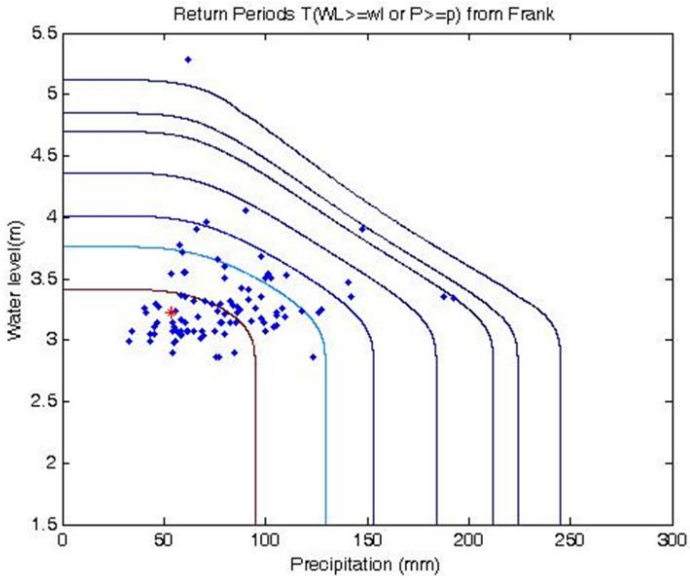

Fig. 4. Joint return period of water level and precipitation

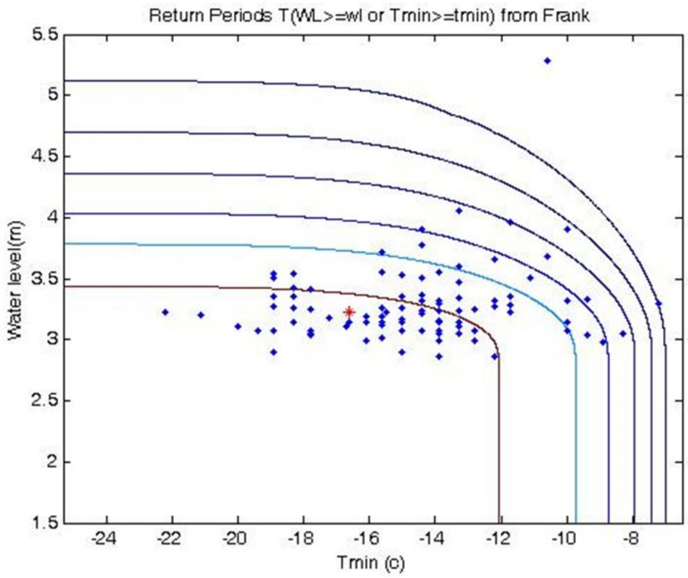

Fig. 5. Joint return period of water level and minimum temperature

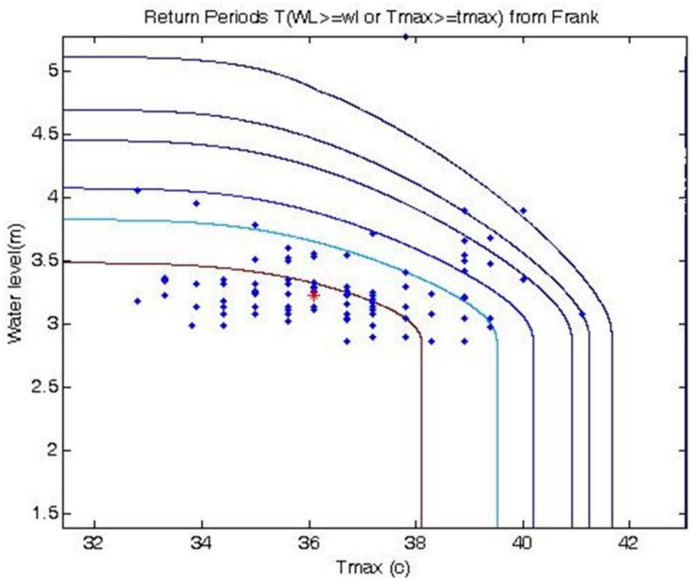

Fig. 6. Joint return period of water level and maximum temperature

4. Conclusion

This study investigates trend analysis in several climate variables under the potential impacts of climate change. The climate variables considered include water level, precipitation, maximum temperature, and minimum temperature. The Manhattan coastal region in New York City is selected as the case study. The required data for analysis were obtained from the Central Park and Battery Park stations. For univariate and multivariate trend analyses, the Mann-Kendall and Partial Mann-Kendall tests were employed, respectively. Several common probability distributions were examined for the frequency analysis of the data.

To identify the most appropriate distribution function, the Kolmogorov-Smirnov goodness-of-fit test was utilized. Subsequently, the joint probability of water level with rainfall and temperature was investigated. To achieve this, the correlation between the climate variables was assessed. Various copula functions were fitted to the data, and the joint probability of observed water levels with both rainfall and temperature data for different return periods was calculated.

The results of the trend analysis indicated an increase in the long-term average of rainfall, water level and minimum temperature, which could be attributed to the impacts of climate change. Additionally, the results suggested that the Generalized Extreme Value (GEV) distribution is the most appropriate distribution for fitting to the considered extreme hydrologic variables. It should be noted that the GEV distribution's compatibility with copula functions further enhances its utility in multivariate analyses, allowing for the assessment of joint probabilities of extreme events, such as the simultaneous occurrence of high rainfall and sea levels.

The findings from the joint probability distribution analysis underscore the necessity of considering rainfall, temperature, and water level in flood frequency analysis, particularly because these climate variables are interrelated.

The results, particularly identifying increasing trends in extreme rainfall, water levels, and temperature, underscore the urgent need for adaptive infrastructure design. As climate change continues to exacerbate the frequency and intensity of extreme weather events, urban planners and engineers must consider these trends in designing flood defenses, drainage systems, and coastal infrastructure. For instance, the application of the GEV distribution in our analysis provides a statistical basis for estimating the return periods of extreme events, which can guide the design of levees and floodwalls to withstand anticipated future conditions.

This study highlights the interconnectedness of rainfall, temperature, and sea level rise, which is critical for urban planning in flood-prone areas like New York City. Policymakers can utilize our findings to inform zoning regulations that restrict development in high-risk areas and promote the use of green infrastructure, such as permeable pavements and green roofs, which can mitigate flooding by enhancing stormwater management. Additionally, integrating our joint probability analysis into urban planning frameworks can help identify areas at risk of simultaneous flooding from rainfall and storm surges, allowing for more effective land-use planning.

The increasing trends in extreme climate variables necessitate the development of comprehensive climate adaptation strategies. Our study's findings can inform local governments and agencies in creating action plans that prioritize resilience against flooding. This includes investing in early warning systems, enhancing emergency response protocols, and conducting regular risk assessments to adapt to changing climatic conditions. Furthermore, our results can support the implementation of nature-based solutions, such as restoring wetlands and creating buffer zones, which can provide natural flood protection and enhance biodiversity.

Consequently, the insights gained from our analysis can serve as a foundation for formulating policies that address the impacts of climate change on flood risk. Engaging stakeholders, including community members, urban planners, and environmental organizations, in discussions about our findings, can foster collaborative approaches to flood management. This engagement is crucial for developing policies that are not only scientifically sound but also socially equitable and economically viable.

5. Disclosure statement

No potential conflict of interest was reported by the authors.

6. References

AghaKouchak, A., Cheng, L., & Mazdiyasni, O. (2020). Concurrent changes in extreme weather events and their implications for flood risk. *Geophysical Research Letters*, 47(12), e2020GL087123.

AghaKouchak, A., Cheng, L., Mazdiyasni, O., & Farahmand, A. (2014). Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophysical Research Letters*, *41*(24), 8847-8852.

Bedient, P. B., & Huber, W. C., (2002). Hydrology and floodplain analysis. *Addison-Wesley*.

Burn, D. H. (1990). Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, *26*(10), 2257-2265.

Burn, D. H., Cunderlik, J. M., & Pietroniro, A. (2004). Hydrological trends and variability in the Liard River basin. *Hydrological Sciences Journal*, *49*(1), 53-67.

Chebana, F., Ouarda, T. B., & Duong, T. C. (2013). Testing for multivariate trends in hydrologic frequency analysis. *Journal of Hydrology*, *486*, 519-530.

Cheng, L., Hoerling, M., AghaKouchak, A., Livneh, B., Quan, X. W., & Eischeid, J. (2016). How has human-induced climate change affected California drought risk?. *Journal of Climate*, *29*(1), 111-120.

Cong, R. G., & Brady, M. (2012). The interdependence between rainfall and temperature: copula analyses. *The Scientific World Journal*, *2012*(1), 405675.

Fathian, F., Dehghan, Z., Bazrkar, M. H., & Eslamian, S. (2016). Trends in hydrological and climatic variables affected by four variations of the Mann-Kendall approach in Urmia Lake basin, Iran. *Hydrological Sciences Journal*, *61*(5), 892-904.

Favre, A. C., El Adlouni, S., Perreault, L., Thiémonge, N., & Bobée, B., (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40(1).

Foster, H. A. (1924). Theoretical frequency curves and their application to engineering problem. *Transactions of the American Society of Civil Engineers*, *87*(1), 142-173.

Genest, C., & Werker, B. J. (2002). Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models. *Distributions with given Marginals and Statistical Modelling*, 103-112.

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, *44*(2), 199-213.

Gilroy, K. L., & McCuen, R. H. (2012). A nonstationary flood frequency analysis method to adjust for future climate change and urbanization. *Journal of Hydrology*, *414*, 40-48.

Goharian, E., Burian, S. J., Bardsley, T., & Strong, C. (2016). Incorporating potential severity into vulnerability assessment of water supply systems under climate change conditions. *Journal of Water Resources Planning and Management*, *142*(2), 04015051.

Golian, S., Saghafian, B., & Farokhnia, A. (2012). Copula-based interpretation of continuous rainfall–runoff simulations of a watershed in northern Iran. *Canadian Journal of Earth Sciences*, *49*(5), 681-691.

Gräler, B., van den BERG, M. J., Vandenberghe, S., Petroselli, A., Grimaldi, S., De Baets, B., & Verhoest, N. E. C. (2013). Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Sciences*, *17*(4), 1281-1296.

Gumbel, E. J. (2012). Statistics of Extremes, Courier Corporation.

Hawkes, P. J., Gonzalez-Marco, D., Sánchez-Arcilla, A., & Prinos, P. (2008). Best practice for the estimation of extremes: A review. *Journal of Hydraulic Research*, *46*(S2), 324-332.

Karamouz, M., Fereshtehpour, M., Ahmadvand, F., & Zahmatkesh, Z. (2016). Coastal flood damage estimator: An alternative to FEMA's HAZUS platform. *Journal of Irrigation and Drainage Engineering*, *142*(6), 04016016.

Karamouz, M., Razmi, A., Nazif, S., & Zahmatkesh, Z. (2017). Integration of inland and coastal storms for flood hazard assessment using a distributed hydrologic model. *Environmental Earth Sciences*, *76*, 1-17.

Karamouz, M., Zahmatkesh, Z., Nazif, S., & Razmi, A. (2014). An evaluation of climate change impacts on extreme sea level variability: Coastal area of New York City. *Water resources management*, *28*, 3697-3714.

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in climate change. *Climate Research*, 21(2), 107-120.

Keerthirathne, D. G. T. C., & Perera, K. (2015). Joint distribution of rainfall and temperature in Anuradhapura, Sri Lanka using copulas.

In *International Research Symposium on Engineering Advancements*, 42-45.

Kendall, M.G., (1975). Rank Correlation Methods, fourth ed. Charles Griffin & Company Ltd., London.

Kharin, V. V., Zwiers, F. W., & Zhang, X. (2013). *Changes in Temperature and Precipitation Extremes in the CMIP5 Models*. Journal of Climate, 26(18), 7128-7141.

Kojadinovic, I., & Yan, J. (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software*, *34*, 1-20.

Kron, W. (2005). Flood risk= hazard• values• vulnerability. *Water International*, *30*(1), 58-68.

Laz, O. U., & Rahman, A. (2014). Trends in extreme rainfall events in Tasmania, Australia. *International Journal of Environmental and Ecological Engineering*, *8*(12), 824-828.

Li, N., Liu, X., & Zhang, P. (2023). "Joint distribution analysis of rainfall and temperature using copulas: Implications for flood frequency analysis." *Water Resources Research*, 59(4), e2022WR032123.

Li, N., Liu, X., Xie, W., Wu, J., & Zhang, P. (2013). The return period analysis of natural disasters with statistical modeling of bivariate joint probability distribution. *Risk Analysis: An International Journal*, *33*(1), 134-145.

Libiseller, C. (2002). A program for the computation of multivariate and partial Mann-Kendall test. *University of Linköping, Sweden*.

Libiseller, C., & Grimvall, A. (2002). Performance of partial Mann–Kendall tests for trend detection in the presence of covariates. *Environmetrics: The official journal of the International Environmetrics Society*, *13*(1), 71-84.

Liu, Z., Zhou, P., Chen, X., & Guan, Y. (2015). A multivariate conditional model for streamflow prediction and spatial precipitation refinement. *Journal of Geophysical Research: Atmospheres*, *120*(19), 10-116.

Madadgar, S. (2013). *Towards improving drought forecasts across different spatial and temporal scales* (Doctoral dissertation, Portland State University).

Madadgar, S., & Moradkhani, H. (2013). Drought analysis under climate change using copula. *Journal of Hydrologic Engineering*, *18*(7), 746-759.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, 245-259.

Milly, P. C. D., Wetherald, R. T., Dunne, K. A., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, *415*(6871), 514-517.

Mudersbach, C., & Jensen, J. (2010). Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline. *Journal of Flood Risk Management*, *3*(1), 52-62.

Nelsen, R. B. (2006). *An introduction to copulas*. Springer.

Pielke, R. A., Downton, M. W., & Miller, J. B. (2002). *Flood damage in the United States, 1926-2000: a reanalysis of National Weather Service estimates*. Boulder, CO: University Corporation for Atmospheric Research.

Razmi, A. (2012). *Assessment of climate change impacts on coastal flood risk management.* (M.Sc. thesis, Azad University, Tehran, Iran).

Robson, A. J. (2002). Evidence for trends in UK flooding. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *360*(1796), 1327-1343.

Roussas, G. G. (2013). *Introduction to probability*. Academic Press.

Saf, B. (2008). Application of index procedures to flood frequency analysis in Turkey 1. *JAWRA Journal of the American Water Resources Association*, *44*(1), 37-47.

Salvadori, G., & De Michele, C. (2004). Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research*, *40*(12).

Salvadori, G., De Michele, C., & Durante, F. (2011). On the return period and design in a multivariate framework. *Hydrology and Earth System Sciences*, *15*(11), 3293-3305.

Salvadori, G., De Michele, C., Kottegoda, N. T., & Rosso, R. (2007). Extremes in nature: an approach using copulas (Vol. 56). *Springer Science & Business Media*.

Salvadori, G., Durante, F., & De Michele, C. (2013). Multivariate return period calculation via survival functions. *Water Resources Research*, *49*(4), 2308-2311.

Schoelzel, C., & Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research–introduction to the copula approach. *Nonlinear Processes in Geophysics*, *15*(5), 761-772.

Shih, J. H., & Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, 1384-1399.

Singh, V. P., & Strupczewski, W. G. (2002). On the status of flood frequency analysis.

Sklar, A. (1996). Random variables, distribution functions, and copulas: a personal look backward and forward. *Lecture Notes-Monograph Series*, 1-14.

Song, J. B., Wei, E. B., & Hou, Y. J. (2004). Joint statistical distribution of two-point sea surface elevations in finite water depth. *Coastal Engineering*, *50*(4), 169-179.

Su, D., Wang, Y., & Chen, Y. (2020). *Generalized Extreme Value Distribution for Modeling Extreme Weather Events: A Case Study of Rainfall in China*. Water, 12(8), 2258.

Svensson, C., Kundzewicz, W. Z., & Maurer, T. (2005). Trend detection in river flow series: 2. Flood and low-flow index series. *Hydrological Sciences Journal*, *50*(5).

Wahlin, K., & Grimvall, A. (2010). Roadmap for assessing regional trends in groundwater quality. *Environmental Monitoring and Assessment*, *165*, 217-231.

Wang, C., Chang, N. B., & Yeh, G. T. (2009). Copula-based flood frequency (COFF) analysis at the confluences of river systems. *Hydrological Processes: An International Journal*, *23*(10), 1471-1486.

William, H. (2003). Econometric analysis fifth edition.

Xu, K., Ma, C., Lian, J., & Bin, L. (2014). Joint probability analysis of extreme precipitation and storm tide in a coastal city under changing environment. *PLoS One*, *9*(10), e109341.

Zeng, L., Wang, H., & Liu, X. (2023). Revisiting the Generalized Extreme Value Distribution for Rainfall Extremes: A Comprehensive Analysis Based on Long-term Observations. *Journal of Hydrology*, 598, 126244.

Zhou, Y., Zhang, Y., & Chen, H. (2021). "Impact of climate change on precipitation patterns and flood risks in urban areas." *Journal of Hydrology*, 598, 126-134.