



Article

Estimating Concurrent Probabilities of Compound Extremes: An Analysis of Temperature and Rainfall Events in the Limpopo Lowveld Region of South Africa

Caston Sigauke ^{*,†}  and Thakhani Ravele [†] 

Department of Mathematical and Computational Sciences, University of Venda Private Bag X5050, Thohoyandou 0950, South Africa; thakhani.ravele@mvulaunivenac.onmicrosoft.com

* Correspondence: caston.sigauke@univen.ac.za; Tel.: +27-15-962-8135

† These authors contributed equally to this work.

Abstract: In recent years, there has been increasing interest in the joint modelling of compound extreme events such as high temperatures and low rainfall. The increase in the frequency of occurrence of these events in many regions has necessitated the development of models for estimating the concurrent probabilities of such compound extreme events. The current study discusses an application of copula models in predicting the concurrent probabilities of compound low rainfall and high-temperature events using data from the Lowveld region of the Limpopo province in South Africa. The second stage discussed two indicators for monitoring compound high temperature and low rainfall events. Empirical results from the study show that elevations ranging from 100–350 m, 350–700 m and 700–1200 m exhibit varying probabilities of experiencing drought, with mild droughts having approximately 64%, 66%, and 65% chances, moderate droughts around 36%, 39%, and 38%, and severe droughts at approximately 16%, 19%, and 18%, respectively. Furthermore, the logistic regression models incorporating the southern oscillation index as a covariate yielded comparable results of copula-based models. The methodology discussed in this paper is robust and can be applied to similar datasets in any regional setting globally. These findings could be useful to disaster management decision makers, helping them formulate effective mitigation strategies and emergency response plans.



Citation: Sigauke, C.; Ravele, T. Estimating Concurrent Probabilities of Compound Extremes: An Analysis of Temperature and Rainfall Events in the Limpopo Lowveld Region of South Africa. *Atmosphere* **2024**, *15*, 557. <https://doi.org/10.3390/atmos15050557>

Academic Editor: Masoud Rostami

Received: 30 March 2024

Revised: 27 April 2024

Accepted: 28 April 2024

Published: 30 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: bivariate extremes; copulas; drought; joint extreme events; rainfall deficit; temperature

1. Introduction

1.1. Overview

The current study discusses applying copula models to predict the concurrent probabilities of compound low-rainfall and high-temperature events using data from the Lowveld region of the Limpopo province in South Africa. In order to develop effective climate adaptation and risk-management strategies, it is therefore important to accurately predict the joint behaviour of high temperature and low rainfall. There is a strong correlation between the frequency and impact of extreme events, such as heat waves and droughts. However, extreme events can have a greater impact when they occur in a particular order or sequence. Drought and heat waves occurring simultaneously have a greater effect than univariate counterparts [1]. A simultaneous occurrence of extreme events can significantly impact the ecosystem and society. Climate change and variability can be mitigated by predicting changes in concurrent climate extremes [2].

A few studies have examined concurrent climatic extremes, but most analysed changes in a single climate variable. Climate extremes, such as extreme temperature and rainfall, are critical in determining drought severity and risk. In order to manage and mitigate natural hazards, it is imperative to understand rainfall and temperature trends, their probabilistic characteristics, and how they relate to future climate changes [2].

1.2. A Survey of the Related Literature

Due to the increased threat to human society and ecosystems caused by extreme weather events, there has been a growing interest in the joint modelling of high temperatures and low rainfall. Several approaches are used in modelling such events, but the copula modelling framework has gained much interest.

An overview of copula modelling is discussed in detail by [3]. The authors investigated the dependence between two random variables using copulas. Although the paper emphasised inference and testing procedures, the authors also presented an application of the proposed methodology to modelling Harricana River data.

Bivariate extreme value-copula models are powerful in modelling the joint distribution of extreme compound events such as temperature and rainfall extremes [4]. The bivariate extreme value-copula model has several advantages. It can capture the tail dependence between compound extreme events such as temperature and rainfall. In addition, it allows for a more flexible joint distribution modelling, including nonlinear relationships between extreme compound events [4]. In support of bivariate extreme-value copulas, ref. [5] argues that extreme-value copulas are among the most commonly used copula families since they can capture asymmetry well and are also known to be very flexible.

A recent study in modelling drought risk using bivariate spatial extreme is that of [6]. The authors used temperature and rainfall data to model meteorological drought. Max-stable processes were used in the study to capture the spatio-temporal dependencies of temperature and rainfall data from the Limpopo Lowveld region of South Africa. Results from this study showed that the Schlather model with various covariance functions was a good fit for both data sets compared to the Smith model based on the Gaussian covariance function. However, in this study, the authors did not estimate concurrent probabilities.

In another study, ref. [7], the author used the multivariate frequency analysis to quantify drought risk in the contiguous United States (CONUS). This was carried out by analysing the temperature and rainfall data of CONUS. Results from this study showed that the dependence between low rainfall and high temperature could be positive, negative, or insignificant and that there were no major changes in the last three years. Serinaldi [7] argues that the probability of occurrence of the compound event depends largely on the variables selected and how they are combined.

Furthermore, ref. [8] used Indian data to investigate the concurrence of meteorological droughts and heatwaves. Both variables' extremes are modelled using the peaks over threshold method. Empirical results from this study suggest that there could be an increase in the frequency of concurrent meteorological droughts and heatwaves in India. Zscheischler and Seneviratne [9] investigated how the dependence structure between meteorological variables affects the frequency of occurrence of multivariate extremes. They argue that to fully understand the changes in climate extremes, including their impacts and the designing of adaptation strategies, it is important to use the multivariate modelling framework.

A review of the different approaches used in the characterisation and modelling compound extremes in hydroclimatology is given by [10]. The approaches discussed include the indicator approach, empirical approach, multivariate distribution, quantile regression, and Markov chain model. The authors highlight the limitations of the data available for modelling extremes and the challenges of modelling asymmetric tail dependencies of multiple events. In another study, ref. [11] conducted a comparative analysis of traditional empirical methods and copula models to estimate the probability of compound climate extremes, i.e., hot, dry and windy events, using data from the central United States of America. In a separate study, ref. [12] used copula models to establish the characteristics and the probability of the occurrence of different combinations of water discharge and several water quality indicators. Empirical results from this study showed that the Gaussian copula is the best function for describing the joint distribution of water discharge and water quality.

McKee et al. [13] utilised the standardised rainfall index (SPI) to classify droughts into four primary categories. Specifically, they defined mild droughts when SPI falls within the range of 0 to -0.99 , moderate droughts for SPI between -1 and -1.49 , severe droughts for SPI in the range of -1.5 to -1.99 , and extreme droughts for SPI less than or equal to -2 . The authors argued that for SPI values of -2 , -1 , 0 , 1 , and 2 , there are associated probabilities that the SPI will be less than or equal to the values above, namely, 0.02 , 0.16 , 0.5 , 0.84 , and 0.92 , respectively.

Drought is recognised as a complex phenomenon. Esit and Yuce [14] in their study argue that a comprehensive analysis of drought necessitates modelling it with multiple variables. The authors used the SPI to characterise drought and utilised various bivariate copula functions in their study, considering different elevation levels. Carrillo et al. [15] support the modelling of drought considering different elevation levels and claim that considering different elevation levels is important. They argue that in regions characterised by complex topography, including elevation gradients can significantly contribute to an improved understanding of drought modelling.

Using the SPI values for two sub-seasons of the rain season, October to December and January to March, ref. [16] assessed the impact of elevation on the severity of drought and frequency of occurrence using South African data from the Free State province over the period 1960–2013. Empirical results showed that highland areas had the highest frequency of droughts. However, the authors noted that extreme droughts occurred in the low-lying areas. It also stated that variations in altitude have notable impacts on the severity of drought at the onset of the summer compared to the late summer season. In a related study using two drought indices, ref. [17] assessed meteorological drought and wet conditions using data from the KwaZulu-Natal province in South Africa. This study showed increased drought frequency and severity with the most extreme dry periods experienced between the 1992–1993 and 2015–2016 summer seasons.

1.3. Research Highlights

Based on the literature survey in Section 1.2, the highlights and contributions of this study are as follows:

- Elevations ranging from 100–350 m, 350–700 m, and 700–1200 m show varying probabilities of experiencing drought, with mild droughts having approximately 64%, 66%, and 65% chances, moderate droughts around 36%, 39%, and 38%, and severe droughts at approximately 16%, 19%, and 18% respectively.
- Specific elevation clusters exhibit distinct frequency probabilities for mild and moderate drought occurrences, such as 0.43 and 0.03 for 100–350 m, 0.32 and 0.02 for 350–700 m, and 0.02 for mild drought at 700–1200 m elevations.
- Logistic regression models incorporating the Southern Oscillation Index (SOI) as a covariate yielded comparable results to copula-based models, demonstrating strong predictive performance for compound low rainfall and high temperatures during the 2015/2016 season.
- The monitoring system captured the major drought years in Southern Africa between 1970 and 2020, which are the 1982/1983, 1991/1992, 2002/2003, 2015/2016, and 2019/2020 seasons.

The remainder of the paper is structured as follows: Section 2 elaborates on the methodology. Section 3 presents the empirical results, followed by the discussion in Section 4, and the conclusion in Section 5.

2. Methods

It is well known that rainfall is dependent on temperature. As temperature increases, rainfall also increases. However, there is a decrease in rainfall for high temperatures, resulting in meteorological drought. Due to the dependence between these two meteorological variables, their joint distribution is best described by copula functions. The present study is an extension of the work done by [6].

2.1. Bivariate Copulas

A copula C is a joint distribution function of standard uniform random variables and is given as [18].

$$C(u_1, \dots, u_n) = P(U_1 \leq u_1, \dots, U_d \leq u_n), \quad (1)$$

where $U_i = U(0, 1) \forall i = 1, \dots, n$. Copulas capture the dependence structure of random variables very well, separately from the marginal distributions. Let Y denote rainfall and X temperature. The distribution function of X and Y is given as ([18])

$$G(y, x) = C(F_Y(y), F_X(x)) = C(v, u), \quad (2)$$

where $U = F_X(x)$ and $V = F_Y(y)$ are standard uniform variables. There are three main classes of copulas: the elliptical, Archimedean, and extreme value copulas. In this study, the elliptical and Archimedean copulas were used. These copulas are known to be flexible in modelling the joint distribution of variables with different marginal distributions.

Drought risk is typically associated with inadequate rainfall. Low rainfall decreases soil moisture, impacting vegetation, agriculture, and water supplies. Temperature, on the other hand, affects evaporation rates, potentially intensifying drought conditions. Considering this, $P(V < v, U > u)$ would provide a more comprehensive understanding of the joint impact of rain and temperature on drought risk. From Figure 1, region 4 is our region of interest in which we need to calculate $P(V < v, U > u)$.

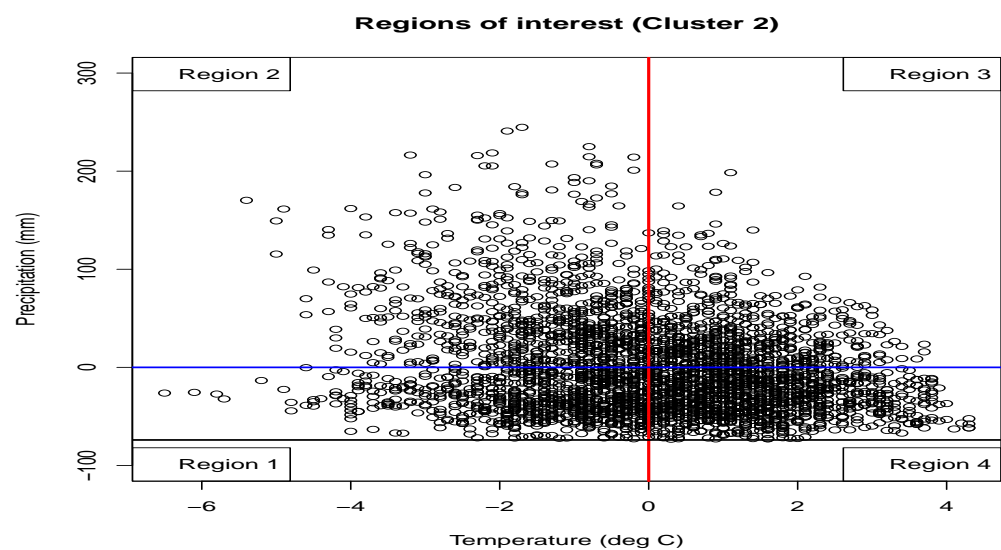


Figure 1. Scatter plot of temperature and rainfall showing the regions of interest for cluster 2. Rainfall and temperature anomalies were used. The blue horizontal line refers to a the reference rainfall while the red line is the reference temperature both used for comparison.

The joint probability $P(V < v, U > u)$ quantifies the probability of experiencing low rainfall with high temperatures above a certain threshold. This probability captures the joint impact of rain and temperature on drought risk by considering situations where both contribute to drought conditions. It helps assess the likelihood of drought risk with insufficient rainfall and high temperatures.

On the other hand, the conditional probability $P(V < v, U > u)$ focuses on rainfall alone, given a specific temperature threshold. While it provides information on the likelihood of low rainfall during high-temperature periods, it does not explicitly capture the joint impact of both variables. Therefore, $P(V < v, U > u)$ is better suited for describing the joint impact of rain and temperature on drought risk as it considers the combined

effects of these variables, providing a more holistic understanding of the conditions leading to drought.

2.1.1. Elliptical Copulas

Elliptical copulas are a class of copulas used in multivariate statistics to model the dependence structure between random variables. These copulas are based on elliptical distributions, which include well-known distributions like the multivariate normal (Gaussian) distribution, the Student’s t-distribution, and the multivariate Laplace distribution. Elliptical copulas are useful for modelling symmetric, linear, and well-behaved dependencies. Two common models within the elliptical copulas are the Gaussian and the t copulas ([19]).

The choice between Gaussian and t copulas depends on the specific characteristics of the data and the modelling objectives. Gaussian copulas are simpler but may not capture extreme events well, while t copulas offer more flexibility but come with greater complexity and estimation challenges ([19]).

The Gaussian Copula

The Gaussian copula is defined as ([19])

$$C_P^{Gauss} = \Phi(\Phi^{-1}(u_1), \dots, \Phi(\Phi^{-1}(u_d))), \tag{3}$$

where $\Phi(\cdot)$ is the standard univariate normal distribution function (DF) and $\Phi_P(\cdot)$ represents the joint CDF of X .

The t Copula

The t copula is defined as ([19])

$$C_{v,P}^t(u) = t_{v,P}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_d)), \tag{4}$$

where P is a correlation matrix, $t_{v,P}$ is the joint DF of $X \sim t_d(v, 0, P)$ and t_v is the standard univariate DF of a t -distribution with v degrees of freedom.

2.1.2. Archimedean Copulas

Archimedean copulas have been widely used because they are convenient and easy to use ([20]), comprise different families, and possess several nice properties ([21]). The Archimedean copula produces a much better dependency model due to its more tractable mathematical properties. Archimedean copulas contain sufficient dependence models for modelling upper and lower tail dependences. Naifar [18] defines the bivariate Archimedean copula equation as given in Equation (5).

$$C_{arch}(u, v) = \phi^{-1}(\phi(u) + \phi(v)), \tag{5}$$

where ϕ denotes a generator function of the copula, if for all $0 \leq u_1, u_2 \leq 1$. Within the Archimedean copula, there are a variety of different dependency structures. These structures simplify the construction of bivariate distributions in many families ([18]). Two Archimedean copula functions are considered in this study: Frank and Gumbel.

The Frank Copula

Frank copula was first introduced by [22] and is defined by,

$$C_{\theta}^{Frank}(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right), \tag{6}$$

where θ denotes the dependence parameter. The upper and lower Fréchet–Hoeffding bounds can be determined using the θ . Modelling data with weak tail dependence is suitable for this model since it is not tail-dependent ($\lambda_u = \lambda_L = 0$).

The Gumbel Copula

Gumbel copula was first introduced by [23] and is defined by Equation (7).

$$C_{\theta}^{Gumbel}(u, v) = \exp\left(-\left[(-\ln u)^{\theta} + (-\ln v)^{\theta}\right]^{\frac{1}{\theta}}\right); \quad 0 \leq u, v \leq 1, \quad (7)$$

where the parameter $\theta \in [1, \infty)$ controls the degree of dependence between u and v . If $\theta = 1$, the bivariate Gumbel copula converges to complete independence, and if $\theta \rightarrow 0$, perfect independence is achieved. The bivariate Gumbel parameter (θ) and Kendall's tau (τ) are integrated by the following formula:

$$\tau_k = 1 - \theta^{-1}. \quad (8)$$

Bivariate Gumbel copula upper (λ^U) and lower (λ^L) tail dependence estimation is carried out by the following functions:

$$\lambda^U = 2 - 2^{\frac{1}{\theta}} \text{ and } \lambda^L = 0.$$

Estimation of the parameters of the copula models will be completed using the maximum likelihood method.

Table 1 presents the relationship between bivariate elliptical and Archimedean copula parameters (θ), tail-dependence coefficients: Kendall's tau (τ) and upper and lower tail dependence, λ^U and λ^L , respectively.

Table 1. Bivariate Archimedean copulas.

Family	θ	Kendal's Tau (τ)	Upper Tail (λ^U)	Lower Tail (λ^L)
Gaussian				
t				
Frank	$-\infty < \theta < \infty$	$1 - \frac{4}{\theta} [D_j(\theta)]$	0	0
Gumbel	$\theta \geq 1$	$\frac{\theta-1}{\theta}$	$2 - 2^{\frac{1}{\theta}}$	0

2.2. Joint and Conditional Distributions

2.2.1. Joint Probability Distribution

The joint probability distribution of low rainfall and high temperature is given in Equation (9).

$$P(V < v, U > u) = P(V \leq v) - P(V \leq v, U \leq u) \quad (9)$$

2.2.2. Conditional Distribution

The conditional distribution of compound high temperature and low rainfall is given in Equation (10).

$$\begin{aligned} P(V < v|U > u) &= \frac{P(V < v, U > u)}{P(U > u)} \\ \text{But } P(V < v, U > u) &= P(V \leq v) - P(V \leq v, U \leq u) \text{ Joint Probability} \\ \implies P(V < v|U > u) &= \frac{P(V \leq v) - P(V \leq v, U \leq u)}{P(U > u)} \\ &= \frac{P(V \leq v) - P(V \leq v, U \leq u)}{1 - P(U \leq u)} \\ &= \frac{C(v, 1) - C(v, u)}{1 - C(1, u)} \\ &= \frac{v - C(v, u)}{1 - u} \end{aligned} \quad (10)$$

2.3. Monitoring Compound Dry and Hot Events

We use two indicators to monitor compound low rainfall and high temperature events. The first is the standardised compound event indicator (SCEI) discussed in [24]. This indicator is based on the joint probability of the two weather variables, rainfall and temperature, in this study.

$$SCEI = \Phi^{-1}(F(P(Y \leq y, X > x))) \tag{11}$$

According to [24], lower SCEI values indicate more severe occurrences of compound dry and hot events.

The second indicator assesses the concurrence of low rainfall and high temperature for specific regions. For given thresholds, p_o and t_o of rainfall and temperature, respectively, the occurrence of a compound low rainfall and high temperature event can be defined as given in Equation (12) ([24]).

$$I = \begin{cases} 1, & \text{if } P \leq p_o, T \geq t_o, \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

Table 2 summarises the drought classification discussed by [24]. We present only part of the table given in [24].

Table 2. SPI Drought classification ([24]).

SPI Values	Drought Category	Time in Category
0 to −0.99	mild drought	24%
−1.00 to −1.49	moderate drought	9.2%
1.50 to −1.99	severe drought	4.4%
≤ −2.00	extreme drought	2.3%

2.4. Prediction of The Occurrences of Compound Events

To predict the occurrence of compound events, i.e., when $Z = 1$, using the Southern Oscillation Index (SOI) as a covariate, we applied the logistic regression model outlined by [24] and represented by Equation (13).

$$\ln \left[\frac{\theta}{1 - \theta} \right] = \alpha + \beta x, \tag{13}$$

where the probability of occurrence $P(Z = 1|x)$ is denoted by θ , α represents the constant, β is the regression coefficient, and x denotes a covariate, which, in this context, is the SOI. Consequently, the one-month-ahead forecast of the probability of compound low rainfall and high temperature events occurring (i.e., $Z = 1$) can be expressed as shown in Equation (14).

$$P(Z_{t+1} = 1|x) = \frac{1}{1 + \exp[-(\alpha + \beta x_t)]} \tag{14}$$

The aim is to evaluate the influence of SOI on the prediction of compound event occurrences, specifically when $Z = 1$.

3. Empirical Results

3.1. Data and Study Area

The study area for this research is found in the Lowveld region of the Limpopo Province in South Africa, situated between latitude -22° and -24° and longitude 30° and 33° . This area is known to experience warm temperatures throughout the year and has an average annual rainfall of 500 mm from October to March ([25]). In this study, we use

monthly rainfall and temperature data. We consider only summer months, i.e., October to March.

The top and bottom panels of Figure 2 show maps illustrating the average annual precipitation and the average maximum temperature across the study area, respectively. These maps also identify clusters corresponding to annual average precipitation and maximum temperature within the study region.

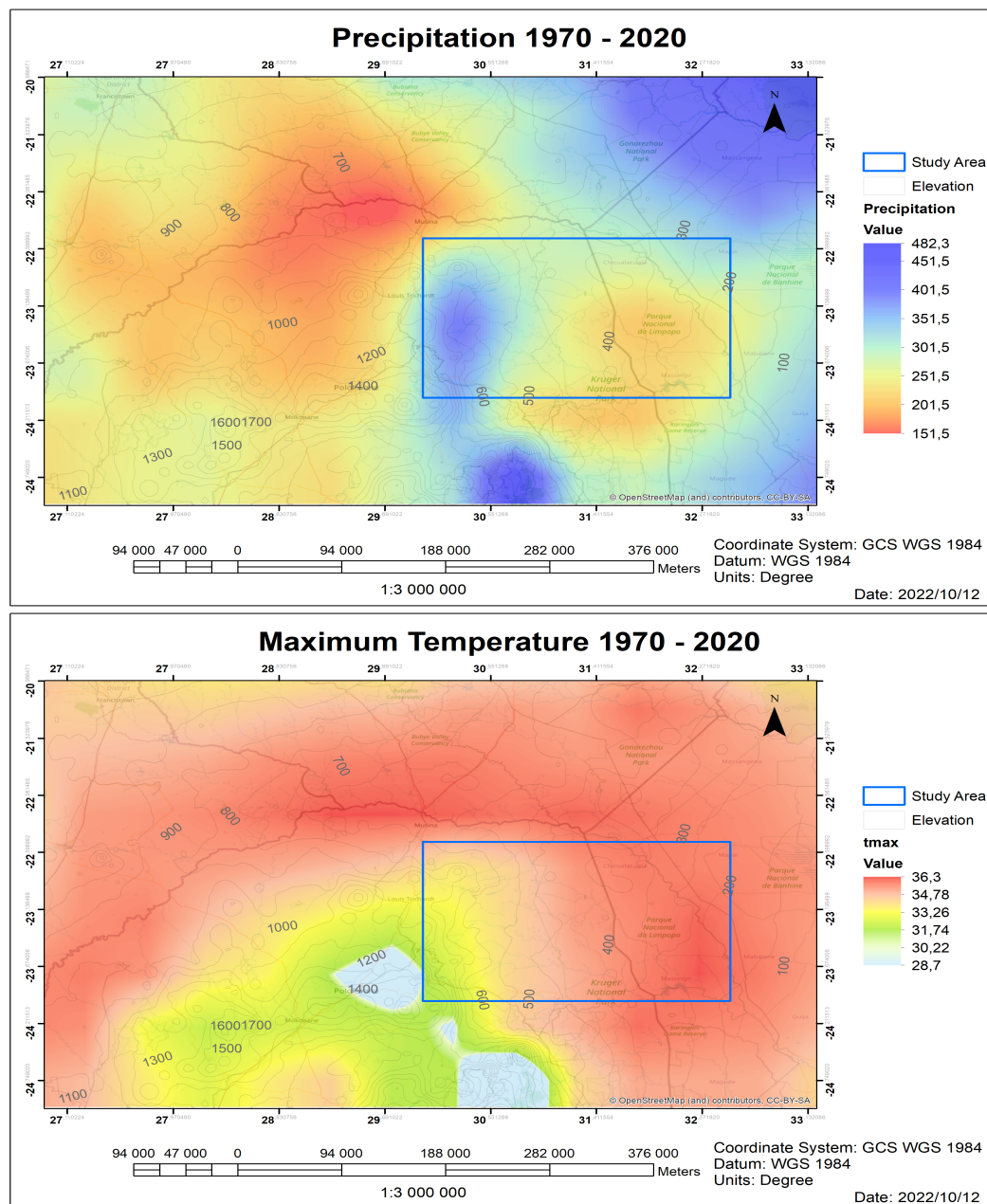


Figure 2. Top panel: Annual average rainfall. Bottom panel: Annual maximum temperature. Source: [6].

The study area is split into three subregions based on the following elevation (altitude) intervals: subregion one 100–350 m, subregion two 350–700 m, and subregion three 700–1200 m, respectively. The average temperatures in these three subregions are: 32.9°, 31.9° and 29.4°, respectively. The average rainfall in these three subregions 1–3 during the rainy season, October to March of each year, are 73.4 mm, 73.9 mm, and 110.1 mm, respectively.

This study uses anomalies of both temperature and rainfall data. Using anomalies can help remove long-term trends, seasonality, and other confounding factors, making it easier

to analyse and model the joint distribution. The metadata, a summary of the grid points, is given in Table A1. The three groups based on elevation are as follows:

- Cluster 1 (elevation 100–350 m) has the following grid points: r4c6, r2c6, r3c6, r1c6, r4c5, r3c5, r2c5, r2c4, r1c4, r1c5 and r4c4 with $n_2 = 3366$ observations.
- Cluster 2 (elevation 350–700 m) has the following grid points: r2c3, r3c4, r4c3, r3c3, r1c3, r2c2, r3c2, r4c2, r1c2, r1c1 and r2c1 with $n_2 = 3366$ observations.
- Cluster 3 (elevation 700–1200 m) has the following grid points: r4c1 and r3c1 with $n_2 = 612$ observations.

We fitted theoretical distributions to precipitation and temperature data. A summary of the results is given in Table 3.

Table 3. Distribution fitting to the data.

Distribution	Elevation 100–350 m		Elevation 350–700 m		Elevation 700–1200 m	
	Temp	Rain	Temp	Rain	Temp	Rain
Gamma	shape = 2561.8 (13.7) rate = 17.1 (0.42)	shape = 2.46 (0.056) rate = 0.03 (0.0008)	shape = 335.68 (8.18) rate = 10.51 (0.26)	Shape = 2.155 (0.0489) rate = 0.029 (0.0007)	shape = 346.85 (19.82) rate = 11.81 (0.68)	
Weibull						shape = 1.67 (0.02) scale = 123.35

3.2. Exploratory Data Analysis

Figure 3 shows histograms (diagonal) superimposed with kernel densities, pairwise scatter plots (bottom left), and pairwise Kendall’s rank correlation coefficient (top right) of temperature and rainfall data for cluster 2 (elevation 350–700 m). The figures for clusters 1 and 3 are given in Figures A1 and A2, respectively.

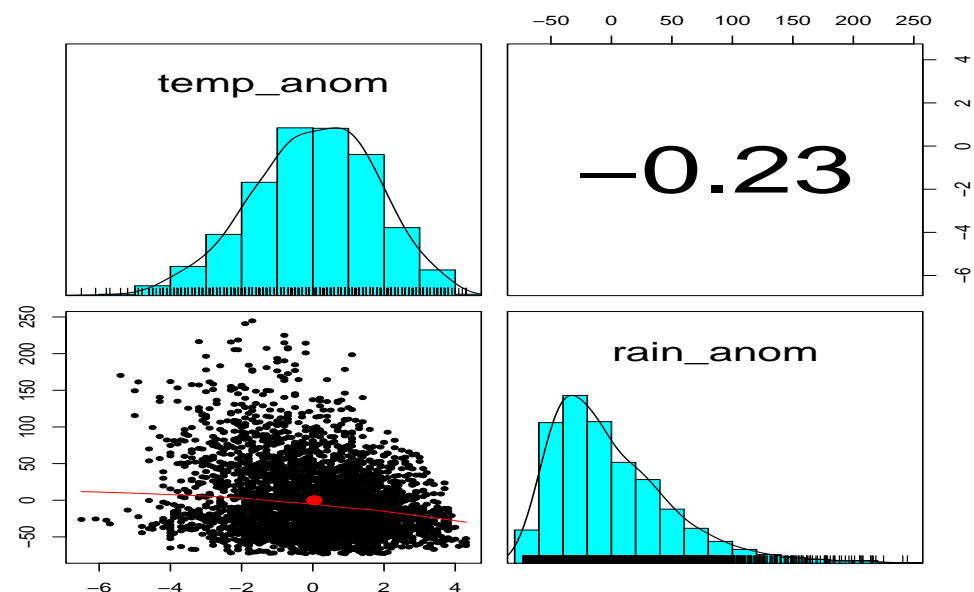


Figure 3. Cluster 2: Histograms (diagonal), a scatter plot of the rainfall and temperature (bottom left). The red dot represents the reference point (0,0). Kendall’s rank correlation coefficient (top right) of temperature and rainfall data for Cluster 2 (elevation 350–700 m).

Based on Kendall’s rank correlation coefficient values for Cluster 1 (elevation 100–350 m), Cluster 2 (elevation 350–700 m), and Cluster 3 (elevation 700–1200 m), rainfall and temperature tend to be more negatively correlated for cluster 2 and a very weak correlation in cluster 1.

Figure 4 shows plots of rainfall, SOI, and temperature, accompanied by box plots displaying their distributions for Cluster 1, corresponding to elevations ranging from 100 to 350 m. High rainfall is observed during the December to February period.

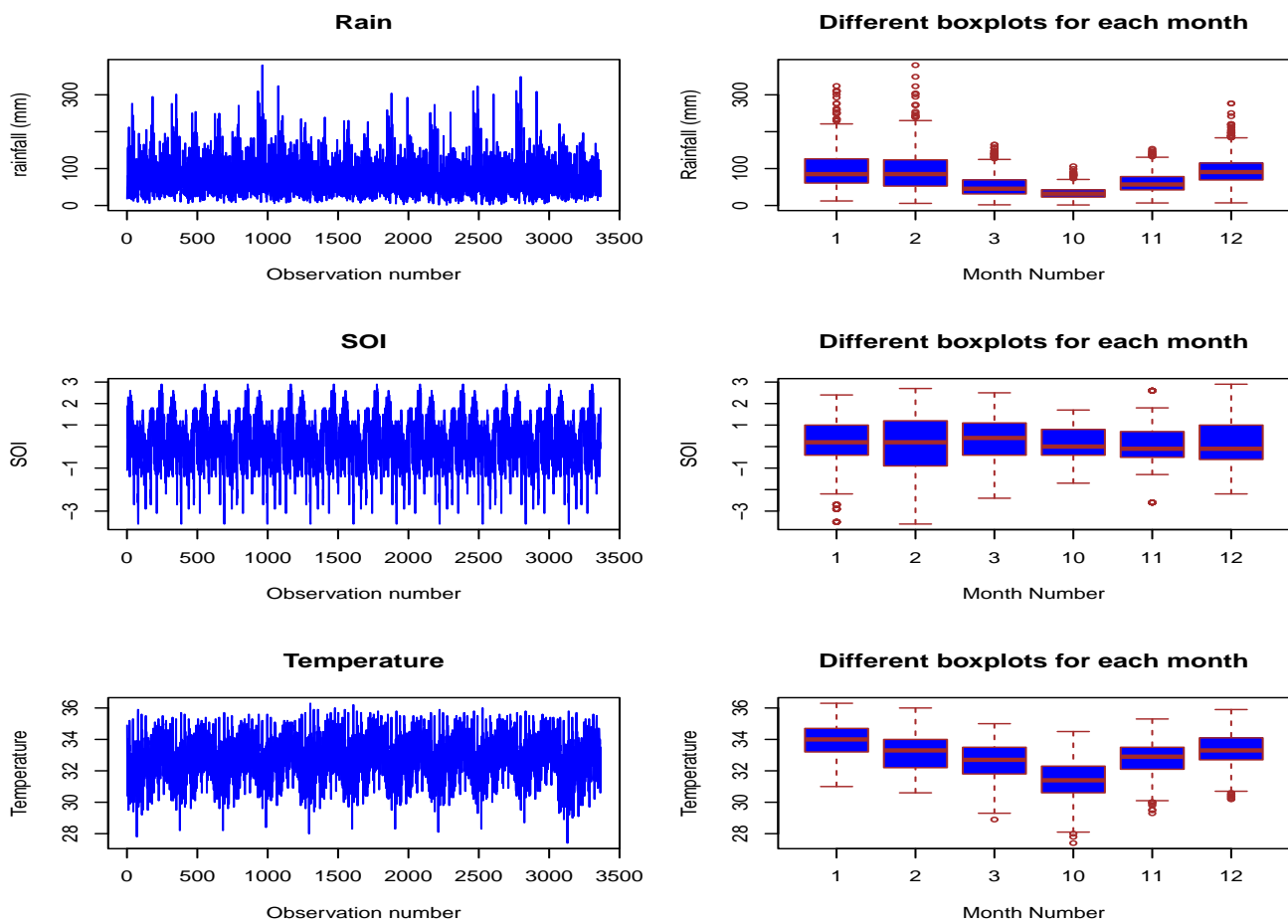


Figure 4. Plots for Cluster 1 (elevation 100–350 m).

3.3. Results

A formal goodness-of-fit test was carried out. The choice of the significance level (α) depends on the trade-off between type I and type II errors. A lower α , such as 0.01 compared to 0.05, reduces the risk of type I error, i.e., false positives. In this study, $\alpha = 0.01$ will be used. The parameter estimates for the best-fitting copula function for all three clusters are given in Table 4.

Table 4. (a) Parameter estimation for the copula functions: Cluster 1 (Elevation 100–350 m); (b) Parameter estimation for the copula functions Cluster 2 (Elevation 350–700 m); (c) Parameter estimation for the copula functions Cluster 3 (Elevation 700–1200 m).

Family	Copula	Estimate (ρ)	ℓ	AIC	BIC	λ^U	λ^L	τ
(a)								
Arhimedean	Frank	0.1710	1.384	−0.7680	5.353	0	0	0.0190
(b)								
Arhimedean	Frank	−1.1137	57.45	−112.89	−106.77	0	0	−0.1224
(c)								
Arhimedean	Frank	−0.8171	5.61	−9.220	−4.803	0	0	−0.0902

Estimating Concurrent Probabilities

Approximate temperature ranges for the different levels of drought severity in the Limpopo Lowveld region of South Africa might be as follows:

- Mild Drought: Slightly above-average temperatures, typically 1–3 °C above the long-term average.
- Moderate Drought: Elevated temperatures, often in the 3–5 °C range above the long-term average.
- Severe Drought: High temperatures, commonly exceeding 5 °C above the long-term average.
- Extreme Drought: High temperatures, potentially reaching 7 °C or more above the long-term average.

Average rainfall and temperature in the study area during the rainy season, October to March, are 74 mm and 32 °C, respectively. The drought characterisation based on the rainfall and temperature ranges of values is given in Table 5. These ranges were based on the information in Table 2 third column, i.e., time in the category given as a percentage.

As discussed in Section 2.1, the joint probability considers the simultaneous occurrence of low rainfall and high temperature, key factors influencing drought conditions. It provides insights into the combined effect of these variables, indicating when both conditions align to contribute to drought risk.

Table 5. Drought characterisation based on rainfall and temperature ranges of values.

Drought Characterisation	Rainfall (mm)	Temperature (Degrees Celsius)
Mild	59 < Y ≤ 74	32 < X ≤ 35
Moderate	52 < Y ≤ 59	35 < X ≤ 37
Severe	44 < Y ≤ 52	37 < X ≤ 40
Extreme	Y ≤ 44	X > 40

Tables 6–8 show the results from the computation of the joint and conditional probabilities for the three elevation groups: 100–350 m, 350–700 m, and 700–1200 m, respectively. The conditional probabilities are significantly higher compared to the joint probabilities. This result is expected since there is some dependency between rainfall and temperature, meaning the occurrence of one variable affects the probability of another.

Table 9 summarises the number of low-rainfall and high-temperature occurrences for the four drought categories. The values in set braces denote the number of occurrences followed by the probability of the occurrence. In square brackets are the probabilities from the logistic regression based on SOI as the covariate. The comparable probabilities suggest that SOI is a good predictor of the joint probabilities of the compound low rainfall and high temperatures.

Table 6. (a) Joint and conditional probabilities based on the Frank copula (Elevation 100–350 m); (b) $P(Y < y, X > x) = \frac{\sum_{y,x} [Y < y, X > x]}{n}$ and $P(Y < y | X > x) = \frac{\sum_{y,x} [Y < y, X > x]}{\sum_{y,x} [X > x]}$ (Elevation 100–350 m).

Joint Probability	Conditional Probability	Drought
(a)		
$P(Y < 74, X > 32) = 0.6378$	$P(Y < 74 X > 32) = 0.7973$	Mild
$P(Y < 59, X > 35) = 0.3551$	$P(Y < 59 X > 32) = 0.5959$	Moderate
$P(Y < 52, X > 37) = 0.1551$	$P(Y < 52 X > 32) = 0.3959$	Severe
$P(Y < 44, X > 40) = 0.0378$	$P(Y < 44 X > 32) = 0.1972$	Extreme
(b)		
$P(Y < 74, X > 32) = 0.6316$	$P(Y < 74 X > 32) = 0.7889$	Mild
$P(Y < 59, X > 35) = 0.3523$	$P(Y < 59 X > 32) = 0.5996$	Moderate
$P(Y < 52, X > 37) = 0.1554$	$P(Y < 52 X > 32) = 0.3688$	Severe
$P(Y < 44, X > 40) = 0.0487$	$P(Y < 44 X > 32) = 0.1859$	Extreme

Table 7. (a) Joint and conditional probabilities based on the Frank copula (Elevation 350–700 m); (b) $P(Y < y, X > x) = \frac{\sum_{v_{xy}[Y < y, X > x]}]{n}$ and $P(Y < y|X > x) = \frac{\sum_{v_{xy}[Y < y, X > x]}]{\sum_{v_x[X > x]}}$ (Elevation 100–350 m).

Joint Probability	Conditional Probability	Drought
(a)		
$P(Y < 74, X > 32) = 0.6566$	$P(Y < 74 X > 32) = 0.8207$	Mild
$P(Y < 59, X > 35) = 0.3898$	$P(Y < 59 X > 32) = 0.6273$	Moderate
$P(Y < 52, X > 37) = 0.1898$	$P(Y < 52 X > 32) = 0.4261$	Severe
$P(Y < 44, X > 40) = 0.0566$	$P(Y < 44 X > 32) = 0.2180$	Extreme
(b)		
$P(Y < 74, X > 32) = 0.6480$	$P(Y < 74 X > 32) = 0.8099$	Mild
$P(Y < 59, X > 35) = 0.3734$	$P(Y < 59 X > 32) = 0.6045$	Moderate
$P(Y < 52, X > 37) = 0.1916$	$P(Y < 52 X > 32) = 0.4181$	Severe
$P(Y < 44, X > 40) = 0.0663$	$P(Y < 44 X > 32) = 0.2094$	Extreme

Table 8. (a) Joint and conditional probabilities based on the Frank copula (Elevation 700–1200 m); (b) $P(Y < y, X > x) = \frac{\sum_{v_{xy}[Y < y, X > x]}]{n}$ and $P(Y < y|X > x) = \frac{\sum_{v_{xy}[Y < y, X > x]}]{\sum_{v_x[X > x]}}$ (Elevation 700–1200 m).

Joint Probability	Conditional Probability	Drought
(a)		
$P(Y < 74, X > 32) = 0.6509$	$P(Y < 74 X > 32) = 0.8136$	Mild
$P(Y < 59, X > 35) = 0.3835$	$P(Y < 59 X > 32) = 0.6197$	Moderate
$P(Y < 52, X > 37) = 0.1835$	$P(Y < 52 X > 32) = 0.4191$	Severe
$P(Y < 44, X > 40) = 0.0509$	$P(Y < 44 X > 32) = 0.2123$	Extreme
(b)		
$P(Y < 74, X > 32) = 0.1144$	$P(Y < 74 X > 32) = 0.7873$	Mild
$P(Y < 59, X > 35) = 0.0683$	$P(Y < 59 X > 32) = 0.5991$	Moderate
$P(Y < 52, X > 37) = 0.0315$	$P(Y < 52 X > 32) = 0.3988$	Severe
$P(Y < 44, X > 40) = 0.0119$	$P(Y < 44 X > 32) = 0.2168$	Extreme

Table 9. (a) Number of occurrences for Cluster 1 (Elevation: 100–350 m ($n = 3366$)); (b) Number of occurrences for Cluster 2 (Elevation: 350–700 m ($n = 3366$)); (c) Number of occurrences for Cluster 3 (Elevation: 700–1200 m ($n = 612$)).

Indicator	Mild	Moderate	Severe	Extreme
(a)				
	$P(Y < 74, X > 32)$	$P(Y < 59, X > 35)$	$P(Y < 52, X > 37)$	$P(Y < 44, X > 40)$
$I = 1$	(1464) 0.4349 [0.3039]	(99) 0.0294 [0.0023]	(0) 0 [0]	(0) 0 [0]
$I = 0$	(1902) 0.5651 [0.6961]	(3267) 0.9706 [0.9977]	(3366) 1 [1]	(3366) 1 [1]
(b)				
	$P(Y < 74, X > 32)$	$P(Y < 59, X > 35)$	$P(Y < 52, X > 37)$	$P(Y < 44, X > 40)$
$I = 1$	(1088) 0.3232 [0.1982]	(72) 0.0214 [0.0027]	(0) 0 [0]	(0) 0 [0]
$I = 0$	(2278) 0.6768 [0.8018]	(3294) 0.9786 [0.9973]	(3366) 1 [1]	(3366) 1 [1]
(c)				
	$P(Y < 74, X > 32)$	$P(Y < 59, X > 35)$	$P(Y < 52, X > 37)$	$P(Y < 44, X > 40)$
$I = 1$	(14) 0.0229 [0.0021]	(0) 0 [0]	(0) 0 [0]	(0) 0 [0]
$I = 0$	(598) 0.9771 [0.9979]	(612) 1 [1]	(612) 1 [1]	(612) 1 [1]

The number of occurrences for mild drought $P(Y \leq 74, X \geq 32)$ (Cluster 1), $P(Y \leq 74, X \geq 32)$ (Cluster 2) and $P(Y \leq 74, X \geq 32)$ (Cluster 3) are shown in Figures 5–7, respectively. The monitoring system captured the major drought years in Southern Africa between 1970 and 2020, which are 1982/1983, 1991/1992, 2002/2003, 2015/2016, and the 2019/2020 seasons [26]. Out of the 1088 incidents of mild drought from cluster 2, 85 were for the 2015/2016 period, which was seen as the most severe drought for the sampling period 2000–2020. These results are consistent with those of [26].

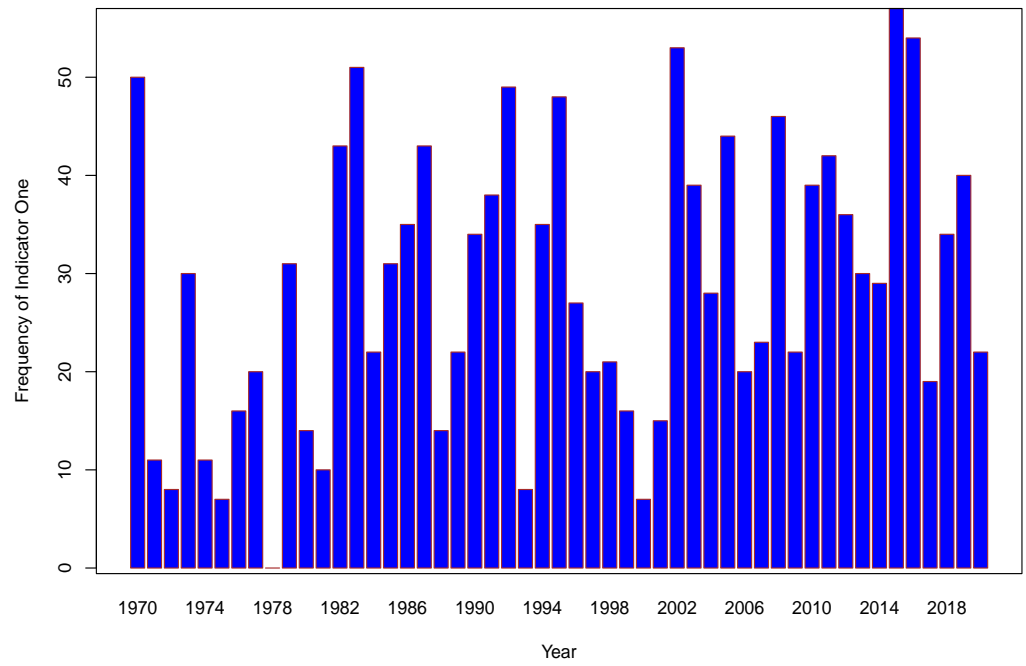


Figure 5. Frequency of indicator ones for mild drought $P(Y \leq 74, X \geq 32)$. Cluster 1 (Elevation: 100–350 m).

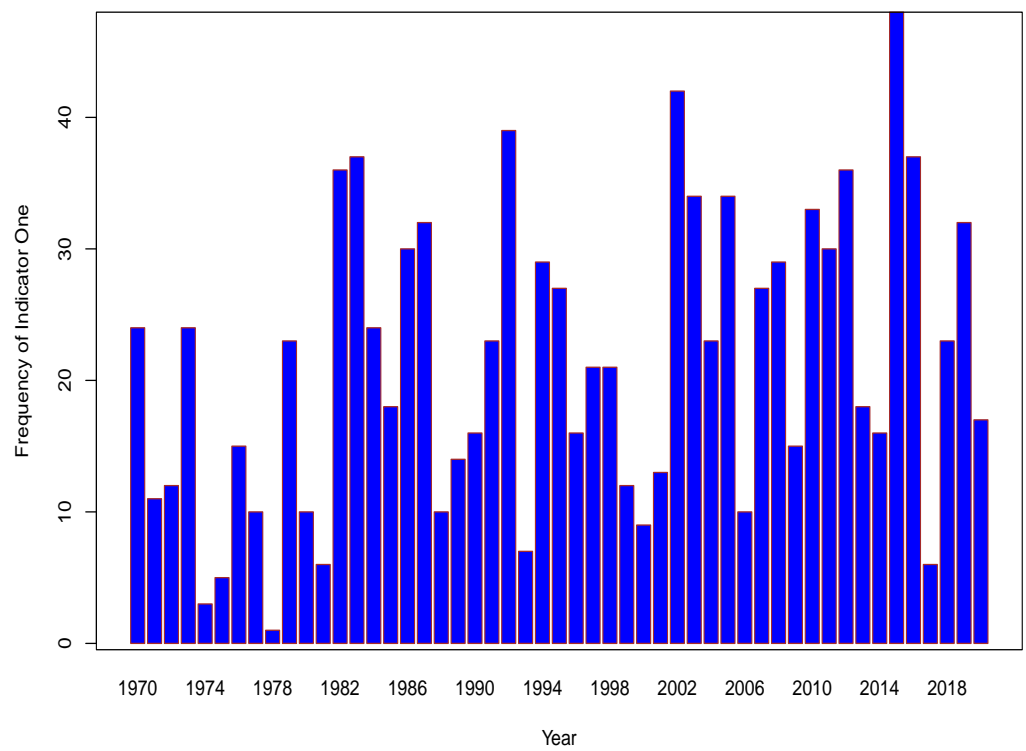


Figure 6. Frequency of indicator ones for mild drought $P(Y \leq 74, X \geq 32)$. Cluster 2 (Elevation: 350–700 m).

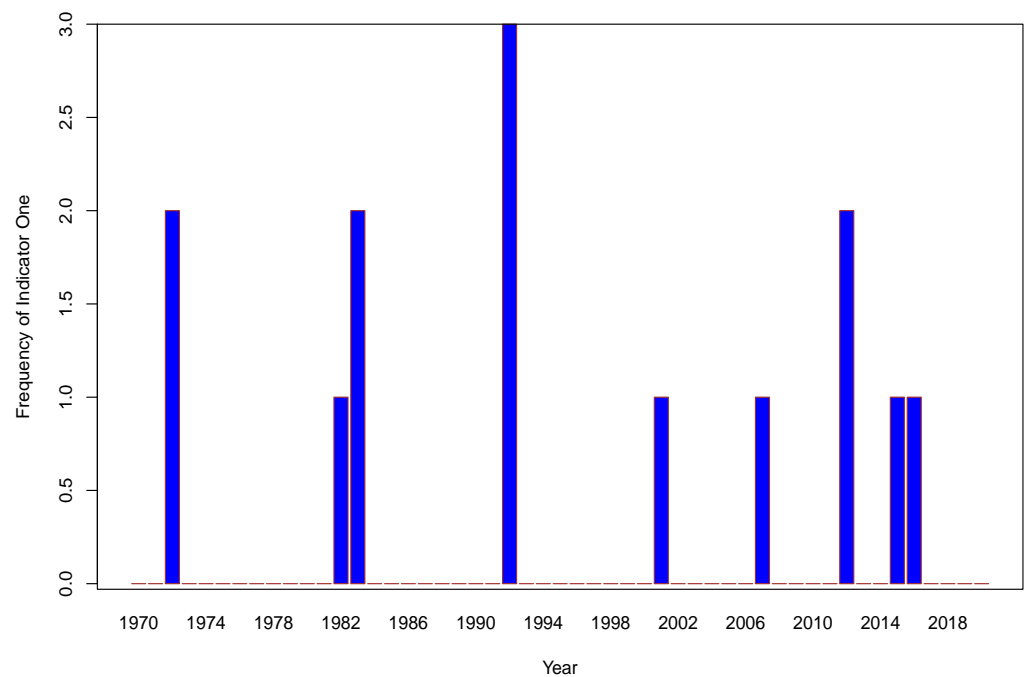


Figure 7. Frequency of indicator ones for mild drought $P(Y \leq 74, X \geq 32)$. Cluster 3 (Elevation: 700–1200 m).

4. Discussion

The current study was motivated by work discussed in [6]. It focused on analysing the concurrent occurrence of low rainfall and high temperatures in the Lowveld region of South Africa’s Limpopo province. The primary aim was to evaluate how elevation influences the joint modelling of drought risk involving low rainfall and high temperatures. This involved using elliptical and Archimedean copulas for joint modelling and estimating the probabilities of getting low rainfall and high temperatures.

The findings of the study revealed that in elevations ranging from 100–350 m, 350–700 m, and 700–1200 m, there were approximately 64%, 66%, and 65% chances of experiencing mild droughts, with corresponding risks of moderate drought at around 36%, 39% and 38%, and severe drought at approximately 16%, 19%, and 18%, respectively. An analysis of mild and moderate drought occurrences for different elevation clusters indicated specific frequency probabilities, such as 0.43 and 0.03 for elevation 100–350 m and 0.32 and 0.02 for elevation 350–700 m, with zero probabilities for severe and extreme droughts. At 700–1200 m elevations, the probability was 0.02 for mild drought, with no occurrence of moderate, severe, or extreme droughts.

Furthermore, these frequency probabilities were compared with those obtained from a logistic regression model utilising the Southern Oscillation Index (SOI) as the sole covariate, demonstrating comparable results. The models exhibited strong predictive performance regarding compound low rainfall and high temperatures during the 2015/2016 season, as depicted in Figures 5–7. These results are consistent with those of [24].

One of the limitations of the study is that it did not include other indicators such as the standardised precipitation index (SPI), Palmer drought severity index (PDSI), normalised difference vegetation index (NDVI), drought severity and coverage index (DSCI), or the standardised precipitation evapotranspiration index (SPEI), among others. It is known that integrating multiple drought indicators provides a comprehensive assessment of drought risk and severity. This will be carried out in future research.

5. Conclusions

The study investigated the increasing occurrence of extreme compound events, focusing on concurrent low rainfall and high temperatures in South Africa’s Lowveld region in the Limpopo province. It assessed how elevation influences drought risk, using copulas for joint modelling and estimating concurrent probabilities. Findings from the study show varying probabilities of mild, moderate, and severe droughts across elevation ranges, with comparisons to a logistic regression model using the southern oscillation index. The study’s modelling framework provides insights into the complex relationship between high temperatures and low rainfall, offering valuable implications for disaster management and suggesting robust methodologies applicable globally.

Author Contributions: Conceptualization, C.S.; Methodology, C.S. and T.R.; Software, T.R.; Validation, C.S. and T.R.; Formal analysis, C.S. and T.R.; Investigation, C.S. and T.R.; Data curation, C.S.; Writing—original draft, C.S. and T.R.; Writing—review & editing, C.S. and T.R. These authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The analytic data can be downloaded from <https://github.com/csigauke> (accessed on 31 October 2023).

Acknowledgments: The authors sincerely thank the anonymous reviewers for their helpful comments and suggestions on this paper.

Conflicts of Interest: There are no conflicts of interest.

Appendix A

Appendix A.1. Tables

Table A1 gives a summary of the metadata.

Table A1. Meta data. Source: Nemukula et al. [6]

Grid ID	r1c1	r1c2	r1c3	r1c4	r1c5	r1c6
Latitude	−22.5	−22.5	−22.5	−22.5	−22.5	−22.5
Longitude	30	30.5	31	31.5	32	32.5
Grid ID	r2c1	r2c2	r2c3	r2c4	r2c5	r2c6
Latitude	−23	−23	−23	−23	−23	−23
Longitude	30	30.5	31	31.5	32	32.5
Grid ID	r3c1	r3c2	r3c3	r3c4	r3c5	r3c6
Latitude	−23.5	−23.5	−23.5	−23.5	−23.5	−23.5
Longitude	30	30.5	31	31.5	32	32.5
Grid ID	r4c1	r4c2	r4c3	r4c4	r4c5	r4c6
Latitude	−24	−24	−24	−24	−24	−24
Longitude	30	30.5	31	31.5	32	32.5

Appendix A.2. Figures

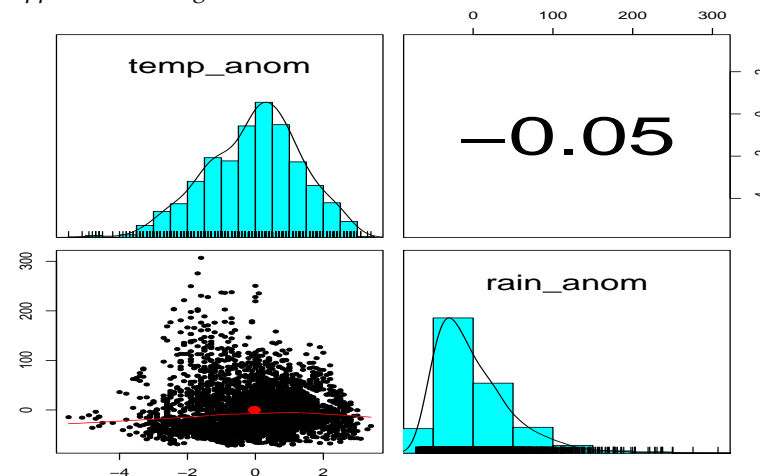


Figure A1. Histograms (diagonal), a scatter plot of the rainfall and temperature (**bottom left**). The red dot represents the reference point (0,0). Kendall's rank correlation coefficient (**top right**) of temperature and rainfall data for Cluster 1 (elevation 100–350 m).

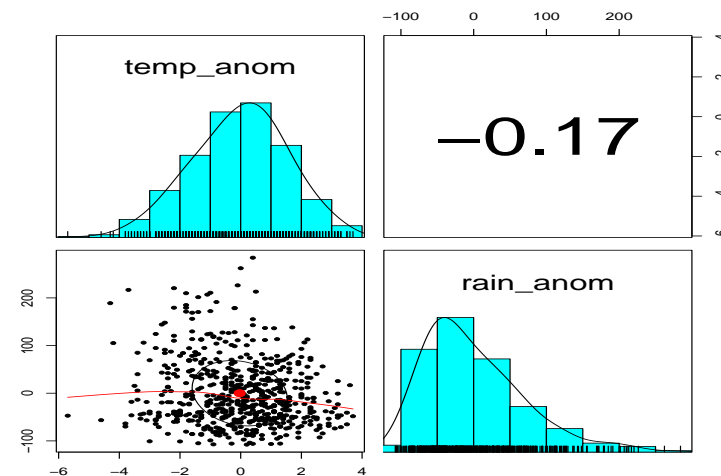


Figure A2. Histograms (diagonal), a scatter plot of the rainfall and temperature (**bottom left**). The red dot represents the reference point (0,0). Kendall's rank correlation coefficient (**top right**) of temperature and rainfall data for Cluster 3 (elevation 700–1200 m).

References

1. Sedlmeier, K.; Mieruch, S.; Schädler, G.; Kottmeier, C. Compound extremes in a changing climate—A Markov chain approach. *Nonlinear Process. Geophys.* **2016**, *23*, 375–390. [[CrossRef](#)]
2. Miao, C.; Sun, Q.; Duan, Q.; Wang, Y. Joint analysis of changes in temperature and rainfall on the Loess Plateau during the period 1961–2011. *Clim. Dyn.* **2016**, *47*, 3221–3234. [[CrossRef](#)]
3. Genest, C.; Favre, A.-C. Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. *J. Hydrol. Eng.* **2007**, *12*, 347–368. [[CrossRef](#)]
4. Gudendorf, G.; Segers, J. *Extreme-Value Copulas*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 1–21. [[CrossRef](#)]
5. Serrano, J.F. Semiparametric bivariate extreme-value copulas. *arXiv* **2022**, arXiv:2109.11307. [[CrossRef](#)]
6. Nemukula, M.M.; Sigauke, C.; Chikoore, H.; Bere, A. Modelling Drought Risk Using Bivariate Spatial Extremes: Application to the Limpopo Lowveld Region of South Africa. *Climate* **2023**, *11*, 46. [[CrossRef](#)]
7. Serinaldi, F. Can we tell more than we can know? The limits of bivariate drought analyses in the United States. *Stoch. Environ. Res. Risk Assess.* **2016**, *30*, 1691–1704. [[CrossRef](#)]
8. Sharma, S.; Mujumdar, P. Increasing frequency and spatial extent of concurrent meteorological droughts and heatwaves in India. *Sci. Rep.* **2017**, *7*, 15582. [[CrossRef](#)] [[PubMed](#)]
9. Zscheischler, J.; Seneviratne, S.I. Dependence of drivers affects risks associated with compound events. *Sci. Adv.* **2017**, *3*, e1700263. [[CrossRef](#)]
10. Hao, Z.; Singh, V.P.; Hao, F. Compound Extremes in Hydroclimatology: A Review. *Water* **2018**, *10*, 718. [[CrossRef](#)]

11. Tavakol, A.; Rahmani, V.; Harrington, J., Jr. Probability of compound climate extremes in a changing climate: A copula-based study of hot, dry, and windy events in the central United States. *Environ. Res. Lett.* **2020**, *15*, 104058. [[CrossRef](#)]
12. Liu, Y.; Cheng, Y.; Zhang, X.; Li, X.; Cao, S. Combined Exceedance Probability Assessment of Water Quality Indicators Based on Multivariate Joint Probability Distribution in Urban Rivers. *Water* **2018**, *10*, 971. [[CrossRef](#)]
13. McKee, T.B.; Doesken, N.J.; Kleist, J. The Relationship of Drought Frequency and Duration to Time Scales. In Proceedings of the 8th Conference on Applied Climatology, Anaheim, CA, USA, 17–22 January 1993; pp. 179–184.
14. Esit, M.; Yuce, M.I. Copula-based bivariate drought severity and duration frequency analysis considering spatial–temporal variability in the Ceyhan Basin, Turkey. *Theor. Appl. Climatol.* **2023**, *151*, 1113–1131. [[CrossRef](#)]
15. Carrillo, J.; Hernández-Barrera, S.; Expósito, F.J.; Díaz, J.P.; González, A.; Pérez, J.C. The uneven impact of climate change on drought with elevation in the Canary Islands. *NPJ Clim. Atmos. Sci.* **2023**, *6*, 31. [[CrossRef](#)]
16. Mbiriri, M.; Mukwada, G.; Manatsa, D. Influence of Altitude on the Spatiotemporal Variations of Meteorological Droughts in Mountain Regions of the Free State Province, South Africa (1960–2013). *Adv. Meteorol.* **2018**, *2018*, 5206151. [[CrossRef](#)]
17. Ndlovu, M.S.; Demlie, M. Assessment of Meteorological Drought and Wet Conditions Using Two Drought Indices Across KwaZulu-Natal Province, South Africa. *Atmosphere* **2020**, *11*, 623. [[CrossRef](#)]
18. Naifar, N. Modelling dependence structure with Archimedean copulas and applications to the iTraxx CDS index. *J. Comput. Appl. Math.* **2011**, *235*, 2459–2466. [[CrossRef](#)]
19. Embrechts, P.; Lindskog, F.; McNeil, A. Modelling dependence with copulas. *Rapp. Tech. Dep. Math. Inst. Fed. Technol. Zur. Zur.* **2001**, *14*, 1–50.
20. Zhang, Q.; Li, J.; Singh, V.P. Application of Archimedean copulas in the analysis of the rainfall extremes: Effects of rainfall changes. *Theor. Appl. Climatol.* **2012**, *107*, 255–264. [[CrossRef](#)]
21. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer Series in Statistics Springer; Springer: Berlin/Heidelberg, Germany, 2006.
22. Frank, M.J. On the simultaneous associativity of $F(x,y)$ and $x + y - F(x,y)$. *Aequ. Math.* **1979**, *19*, 194–226. [[CrossRef](#)]
23. Gumbel, E.J. Distributions des valeurs extremes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* **1960**, *9*, 171–173. Available online: <https://hal.science/hal-04092830> (accessed on 2 February 2024).
24. Hao, Z.; Hao, F.; Xia, Y.; Singh, V.P.; Zhang, X. A monitoring and prediction system for compound dry and hot events. *Environ. Res. Lett.* **2019**, *14*, 114034. [[CrossRef](#)]
25. Chikoore, H.; Jury, M.R. South African drought, deconstructed. *Weather Clim. Extrem.* **2021**, *33*, 100334. [[CrossRef](#)]
26. Bhaga, T.D.; Dube, T.; Shekede, M.D.; Shoko, C. Impacts of Climate Variability and Drought on Surface Water Resources in Sub-Saharan Africa Using Remote Sensing: A Review. *Remote Sens.* **2020**, *12*, 4184. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.