Statistics and Applications {ISSN 2454-7395 (online)}
Volume 21, No. 1, 2023 (New Series), pp 143–159

Statistical Modeling of Temperature in Krishna District using Copula Analysis

A. Rajini[1] and C. Jayalakshmi[2]
[1]Department of Mathematics and Statistics, Bhavan's Vivekananda College of Science, Humanities and Commerce, Sainikpuri, Secunderabad, Telangana, India
[2]Department of Statistics, Osmania University, Hyderabad, Telangana, India

Received: 12 February 2022; Revised: 20 May 2022; Accepted: 18 June 2022

Abstract

Examining the relationship between temperature and precipitation in the Krishna district and forecast temperature is the purpose of study. Krishna is a district in Andhra Pradesh Plateau region, and it was chosen because it is a densely populated area with significant towns and ports. The district's climate is tropical, with sweltering summers and mild winters. From 1901 to 2019, data was obtained from the Indian Meteorological Department in Pune. Data from 1901 to 1996 was used for training, while data from 1997 to 2019 was used for testing. Through Copula analysis, a model is built keeping in view the relationship between Temperature and Precipitation. The Mean Absolute Percentage Error (MAPE) and Normalized Root Mean Square Error (NRMSE) for the best model in Krishna were determined to be 0.03 and 3.823 for the month of May, which has the highest temperature and precipitation dependency when compared to other months. A similar analysis is carried out for the months in which dependence is significant. It is found that five months interdependency coefficient is insignificant. The data was analyzed using R-software, and IBM SPSS statistics version 25 and the results were interpreted. The best Copula does not have to be the same for different datasets. Based on AIC and BIC criteria, the best Copula for Krishna was Gaussians Copula for the month of April and July, Rotated Gumbel 90 Copula for the month of May and September, Rotated Tawn type 2 270 Copula for the month of June, Rotated Gumbel 270 Copula for the month of August and Rotated Clayton 90 Copula for the month of October. Temperature simulated data was found to be very close to testing results. This article examines how Copula modelling can be used to predict temperature, which helps in planning agriculture and trading commodities. So far, this type of analysis and model fitting is not found in the literature for Krishna district in Andhra Pradesh. The temperature in this location may be accurately predicted using our fitted models.

Key words: Temperature; Precipitation; Copula analysis; Mean absolute percentage error; AIC; BIC.

1. Introduction

The atmospheric conditions like temperature, air pressure and moisture vary from one place to another. Due to changes in the climate, it is tough to predict drought and heavy rains at any time in any corner. Weather forecasts are essential warnings as they help protect life and wealth. Forecasted Temperature helps in planning agriculture and trading of

Corresponding Author: A. Rajini
Email: rajini.ststs@bhavansvc.ac.in

commodities. Temperature is a critical parameter in farming varieties of vegetables, fruits and pulses. Hence, there is a need to carry out continuous research on the influencing factors (Temperature and Precipitation) to meet the demand for an increased population. It has been established in the literature of research studies that Temperature and Precipitation have a deep interdependency.

Some research has been conducted in this direction. Dzupire *et al*. (2020), have used Copula analysis to identify interdependency patterns between Temperature and Precipitation. Lazoglou and Anagnostopoulou (2019) developed a joint distribution for the above two factors using Copula in the Mediterranean region. Similar studies have been carried out by Bezak *et al.*(2018) in Slovenia *et al.* (2020) in China, and Shaukat *et al.* (2020) in Pakistan. Mesbahzadeh *et al.* (2019) modelled Temperature and Precipitation for the Arid region using Copula analysis. In their study, Pandey *et al.* (2018) modelled interdependency between Rainfall and Temperature using Copula. This study was carried out in Agartala (humid region) and Bikaner (Arid region). Zscheischler *et al.* (2017), have inferred from their study that environmental change, Precipitation and Temperature are the significant factors that affect the nature of vulnerability influencing harvest. Crop yields are firmly vulnerable to outrageous atmospheres like a dry spell, floods, and warmth waves. Vergara *et al.* (2008) have examined how catastrophe risk modelling can be used in agriculture as a planning tool to predict the frequency and severity of future weather-related catastrophic events, allowing crop insurance firms and policymakers to better prepare for the financial effect of natural disasters. Zhang and Singh (2007) analyzed hydrology for examining the random factors dependence structure modelled independently by marginal distribution individually; Copula approach has been broadly utilized. Olesen and Bindi (2002) discussed and analyzed Temperature and Precipitation influence the duration of expanding season and plant creation (leaf territory and the photosynthetic productivity), respectively. A lot of literature is available on the impacts of Temperature and Precipitation on crop output. Therefore, we can understand the correlation between Precipitation and Temperature that keeps changing in different time periods. Nelsen (2007) demonstrated the factors between dependence structures Copulas are intended based on uniform marginal values.

These studies tried to forecast rainfall in humid, arid and Mediterranean environments. According to a literature survey, there were few pieces of research on the Plateau region and no temperature predictions using Copula analysis were provided. Our study aims to develop a bi-variate model for monthly average Temperature and Precipitation that can be used to simulate Temperature in the selected regions (Krishna district of Andhra Pradesh, which is a Plateau region). Copula analysis was shown to be the most appropriate methodology in this direction. The study's goal is to develop a Copula model that can be used to estimate Temperature in a specific region.

2. Data and methodology

The study used historical monthly average temperature and monthly average precipitation for 119 years, covering the period from 1901 to 2019, collected by the Indian Meteorological Department for the Krishna district. This district is chosen because it is a densely populated area with significant towns and ports. The district's climate is tropical, with extremely hot summers and mild winters.

2.1. Copula methodology

To work with Copulas, one must be familiar with probability and quantile transformations. As the present work is carried out with the help of packages, a detailed discussion on the hidden procedures is not presented in detail. A brief description is given for the sake of the reader.

Quantile transform: If $U \sim U(0, 1)$ has a standard uniform distribution, then $P(F^{-1}(U) \leq x) = F(x)$ it denotes the generalized inverse.

Probability transform: If X has distribution function F with continuous univariate distribution function, then $F(x) \sim U(0,1)$.

Sklar's theorem is a valuable theorem in a Copula environment. It claims that if F has a joint distribution with marginals, then there exists a unique Copula C such that for all $x_1, x_2, \ldots \ldots, x_d \in R \;\; F(x_1, x_2, \ldots \ldots, x_d) = C\big(F_1(x_1), \ldots \ldots F_d(x_d)\big)$.

Co-monotonicity and Counter-monotonicity Copulas are two essential Copulas. The terms Co-monotonicity and Counter-monotonicity refer to perfect positive and negative dependence, respectively. Intuitively, if a Copula exists and has neither positive nor negative dependence structures, it must be somewhere in the middle. Therefore, every Copula $C(u_1,\ u_2, \ldots \ldots\ u_d)$ has bounds:

$$max\left(\textstyle\sum_{i=1}^{d} u_i + 1 - d, 0\right) \leq C(u) \leq min(u_1,\ u_2, \ldots \ldots\ u_d)$$

and is called Frechet bounds for Copula. As a result, the Frechet upper bound Copula is comonotonicity, while the Frechet Lower bound Copula is counter-monotonicity. Fundamental Copulas identify three sorts of dependent structures; definitions of implicit and explicit Copulas are another method of viewing Copulas. Sklar's theorem is used to extract implicit Copulas from well-known multivariate distributions, but they don't have to result in closed-form expressions. Explicit Copulas, in contrast to implicit Copulas, form closed-form expressions and have Yield Copulas as mathematical structures.

2.2. Bivariate Copula

We limit ourselves to the bivariate situation in this study and highlight the significant properties of d-dimensional Copulas that are relevant to the current work. We have,

$$C : [0,\ 1]^2 \to [0, 1], (u, v) = C(u, v)$$

with properties

1. For all $u, v \in [0, 1]$it holds:
$$C(u\,, 0) = C(0, v) = 0 \text{ } and \text{ } C(u, 1) = u \text{ } and \text{ } C(1, v) = v$$

2. For all $u_1, u_2, v_1, v_2 \in [0, 1]$ with $u_1 \leq u_2 \text{ } and \text{ } v_1 \leq v_2$ it holds:
$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$$

is called a bivariate Copula function.

Let X and Y denote temperature and precipitation, which are continuous in nature, with CDF (Cumulative distribution function) $F_X(x) = P(X \leq x)$and $G_Y(y) = P(Y \leq y)$ respectively.

By the definition of Sklar (1973), the joint probability function is given by

$$P(X \leq x, Y \leq y) = C\big(F(x), F(y)\big)$$

where C is an unique function and is known as Copula *i.e.*, $C(u, v) = P(U \leq u, V \leq v)$ is the distribution of $(U,V) = (F(X), G(Y))$ whose marginal distributions are $U[0,1]$. As contended by Joe (1997) and Nelsen (2007), C portrays the dependence between (X, Y). In literature, many Copula families are accessible whose parameters control the intensity of dependence of the variables (X, Y).

Once the parameters of different Copula are estimated, selecting the Copula which can represent the structure of dependency between the interested variables is very important. Few criteria like Aldrian and Black Information Criteria, are available in the literature to identify the best Copula. Information criteria are received here because they can portray the tradeoff between bias (precision) and variance (intricacy) in model development. To measure the relative goodness of fit of a statistical model we use the Akaike information criterion (AIC). It is defined as

$$AIC = 2k - 2\ ln(L)$$

here k is the Copula parameters; L is the optimized value of the likelihood function of the Copula.

The Bayesian information criterion (BIC) was evolved by Schwarz using Bayesian formalism. It is defined as

$$BIC = -2\ ln(L) + k\ ln(N)$$

here N represents the sample size.

3. Analysis

A. Descriptive statistics of temperature and precipitation

The Krishna district's climate is tropical consisting of sweltering summers and mild winters. A clear seasonal cycle has been observed considering the monthly mean Temperature in Krishna from 1901 to 2019, as shown in Figure 1. Similarly, it is observed that there is a seasonal cycle in the monthly Precipitation in Krishna from 1901 to 2019. Figure 2 displays the average monthly precipitation. Table 1 exhibits the Descriptive statistics of Precipitation.

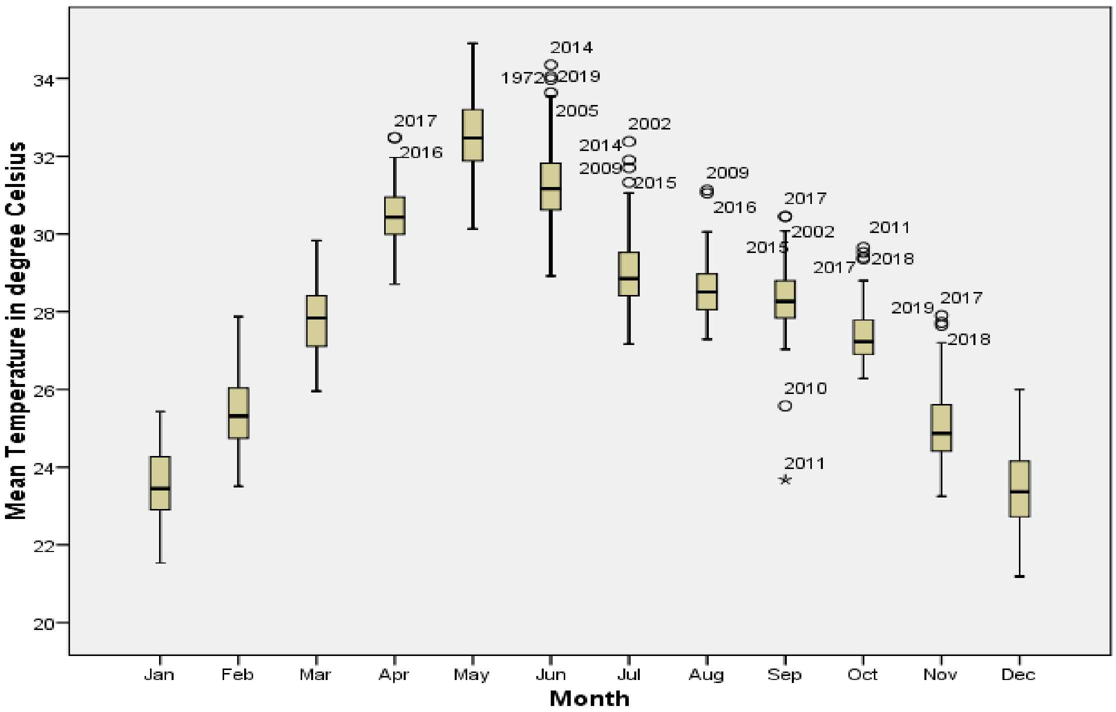


Figure 1: Mean Temperature in Krishna, Andhra Pradesh from 1901 to 2019

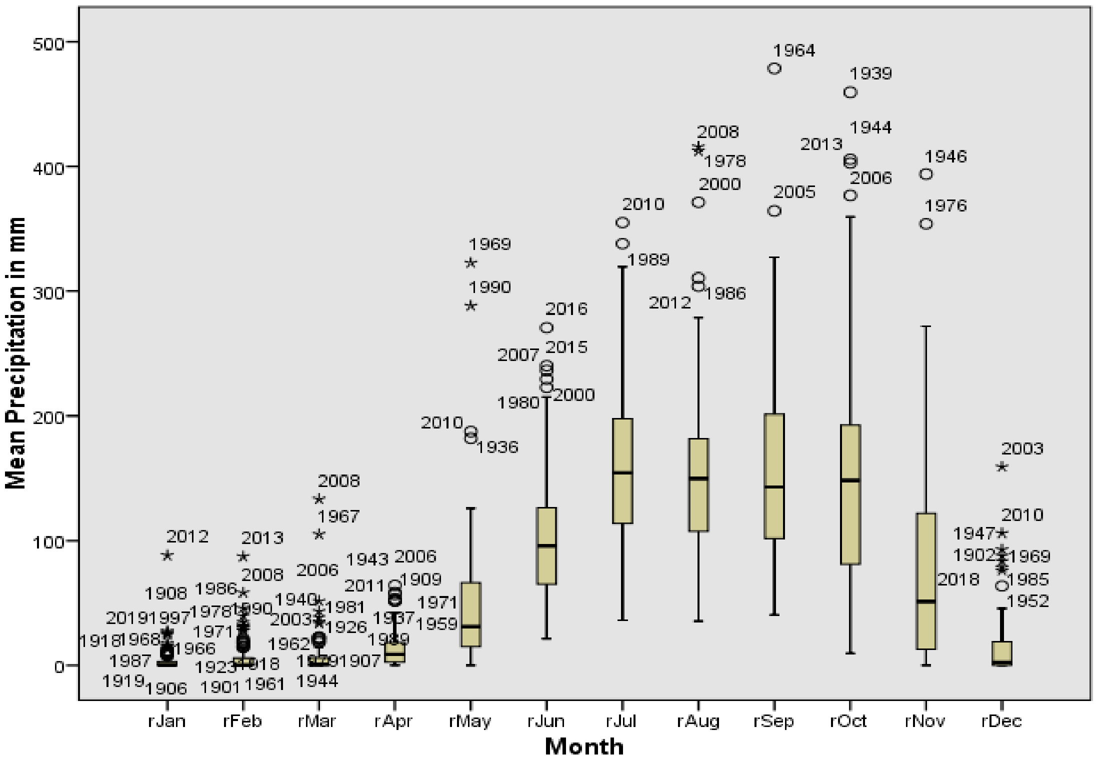


Figure 2: Mean Precipitation (month wise) in Krishna, Andhra Pradesh from 1901 to 2019

Table 1: Descriptive statistics of mean temperature and mean precipitation in Krishna from 1901 to 2019

Month	Temperature in Degree Celsius					Precipitation in mm				
	Min	Max	Mean	Skewness	Kurtosis	Min	Max	Mean	Skewness	Kurtosis
Jan	21.53	25.43	23.55	0.12	–0.60	0.00	88.10	3.81	6.17	49.40
Feb	23.50	27.88	25.42	0.33	–0.40	0.00	87.20	6.48	3.31	14.06
Mar	25.95	29.83	27.82	0.17	–0.66	0.00	133.30	6.69	5.15	31.18
Apr	28.71	32.48	30.47	0.24	–0.11	0.00	64.20	13.89	1.45	1.58
May	30.12	34.90	32.56	0.12	–0.02	0.00	322.67	46.59	2.84	11.57
Jun	28.91	34.35	31.28	0.47	0.54	21.26	270.67	102.31	0.91	0.69
Jul	27.17	32.38	29.04	1.15	1.83	36.25	355.00	160.77	0.60	0.25
Aug	27.29	31.13	28.58	0.88	1.16	35.52	415.80	157.66	1.16	2.06
Sep	23.68	30.45	28.34	–1.09	6.76	40.30	478.71	156.25	1.18	2.52
Oct	26.28	29.65	27.39	0.99	0.97	9.39	459.49	152.17	0.87	0.94
Nov	23.25	27.90	25.09	0.61	0.01	0.00	393.89	75.31	1.46	2.20
Dec	21.18	26.00	23.48	0.52	–0.22	0.00	159.10	14.14	2.96	10.74

B. The association between precipitation and temperature in Krishna district

As the sample data shows a non-Gaussian distribution, the Kendall's tau correlation coefficient is utilized to ascertain the relationship between a month to month Temperature and Precipitation. A negative association has been observed between Precipitation and Temperature during April – October (at the 5% significant level), as given in Table 2.

Table 2: Correlation analysis between temperature and precipitation in Krishna (1901 to 2019)

	January	**February**	**March**	**April**	**May**	**June**
Kendal's Correlation Coefficient	0.147	0.022	– 0.112	– **0.305**	– **0.342**	– **0.285**
***p* - Value**	0.124	0.736	0.082	**0.001**	**0.001**	**0.001**
	July	**August**	**September**	**October**	**November**	**December**
Kendal's Correlation Coefficient	– **0.232**	– **0.173**	– **0.289**	–**0.178**	–0.035	0.07
***p* - Value**	**0.001**	**0.005**	**0.001**	**0.004**	0.577	0.271

C. Estimation of parameters

Initially, a suitable distribution was fitted to Temperature and Precipitation data using R-Software. The best fitted distributions and their parameter estimates are found using this software. Minimum values of AIC and BIC criteria indicate the best fitted distribution. Chi-

Square test is used to determine a variable is likely to come from specified distribution or not. The Maximum Likelihood technique is utilized to estimate the parameters of the best fitted distribution. Table 3 presents the obtained results of Temperature and Precipitation in Krishna. From Table 3, we can observe that all *p*-values are greater than 0.05 (5% level of significance) that is large *p*-values indicate that we can accept the null hypothesis and conclude that data was drawn from a population with the specified distribution.

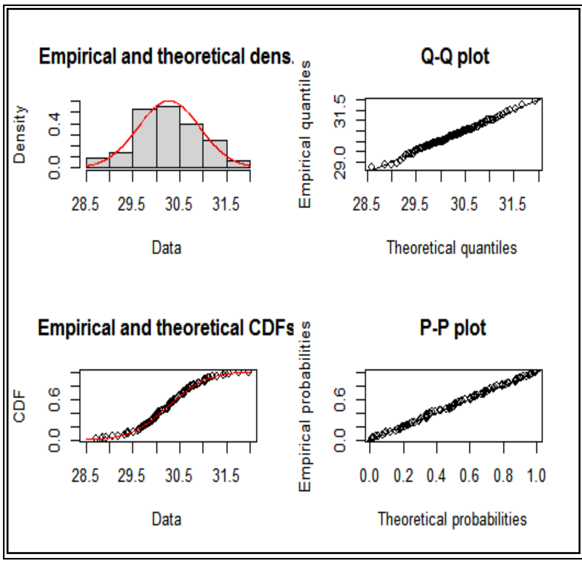
Table 3: Temperature and precipitation parameters estimates

Month	Temperature						Precipitation					
	Distribution	**Parameter**	**Estimate**	***p* - Value**	**AIC**	**BIC**	**Distribution**	**Parameter**	**Estimate**	***p* - Value**	**AIC**	**BIC**
Apr	Normal	μ	30.265	0.432	186.47	191.49	Exponential	μ	14.541	0.945	671.20	673.71
		σ	0.659									
May	Normal	μ	32.34	0.064	247.32	252.43	Exponential	μ	47.142	0.075	624.10	926.65
		σ	0.871									
Jun	Logistic	μ	31.079	0.166	259.82	264.95	Gamma	μ	96.733	0.062	996.89	1002.02
		σ	0.515					σ	0.478			
Jul	Normal	μ	28.776	0.278	188.34	193.47	Gamma	μ	154.48	0.097	1054.65	1059.78
		σ	0.632					σ	0.393			
Aug	Reverse Gumbel	μ	28.112	0.131	161.00	166.13	Gamma	μ	147.78	0.068	1062.25	1067.38
		σ	0.481					σ	0.433			
Sep	Reverse Gumbel	μ	27.927	0.072	156.59	161.72	Inverse Gaussian	μ	155.5	0.826	1070.33	1075.45
		σ	0.469					σ	0.037			
Oct	Normal	μ	27.137	0.305	124.65	129.78	Weibull	μ	171.84	0.071	1125.17	1130.3
		σ	0.453					σ	1.754			

The empirical density function, a simple modification and improvement of the usual histogram, is defined, and its properties are studied. A CDF is the Cumulative Distribution Function. The CDF essentially allows you to plot a feature of the data in order from least to greatest and see the whole feature as if it is distributed across the data set.

Probability plots are the best way to determine whether the data follow a particular distribution. If data follow the straight line on the graph, the distribution fits the data. The Q-Q plot (Quantile – Quantile plot) and P–P plot (Probability–Probability plot) are graphical tools used to determine how well a given data set fits a specific probability distribution that we are testing. Q-Q plot and P–P plot are used to assess how closely two datasets agree, where the two cumulative distribution functions are plotted against each other. If the data points fall along the straight line, we can conclude the data follow that specified distribution.

The graphs from Figure 3 display the frequency curve, cumulative frequency curve, Q-Q and P-P Plots of best fitted distribution of temperature.



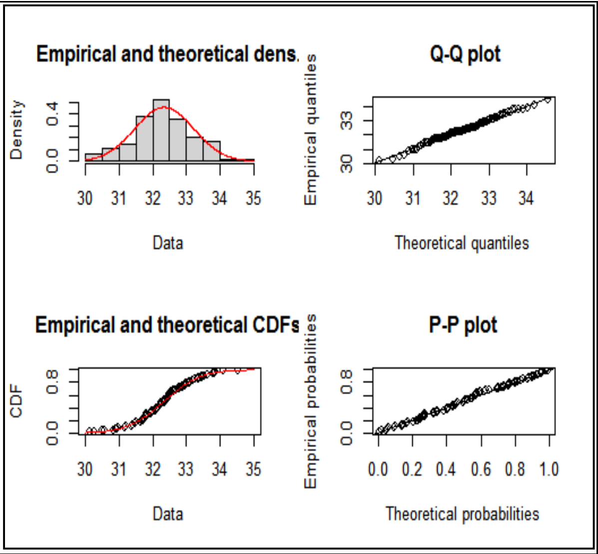
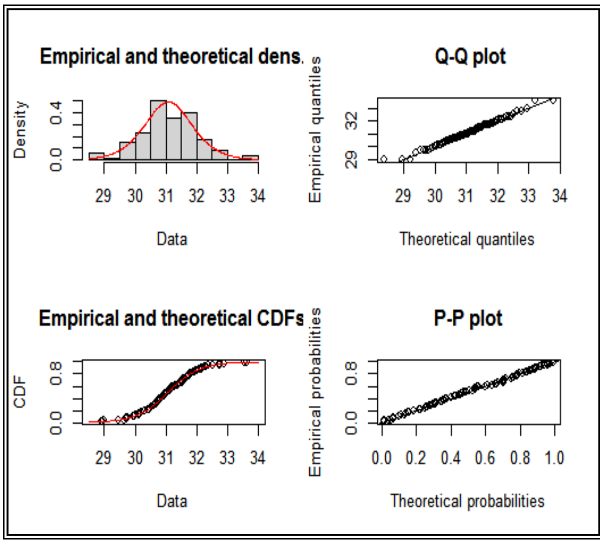


Figure 3(a): Fitted normal distribution of April

Figure 3(b): Fitted normal distribution of May



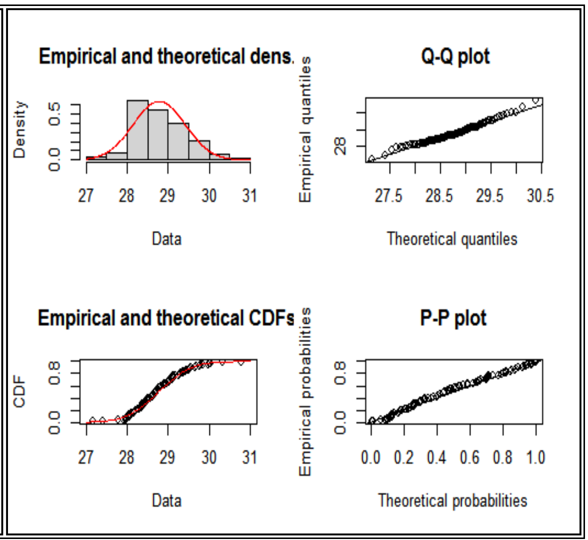


Figure 3(c): Fitted normal distribution of June

Figure 3(d): Fitted normal distribution of July

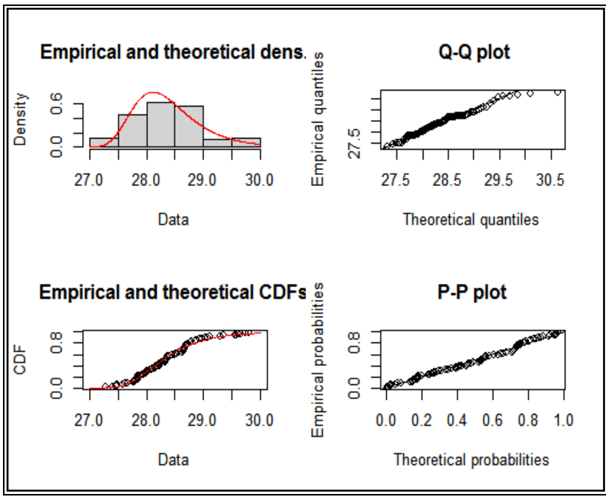


Figure 3(e): Fitted Reverse Gumbel distribution of August

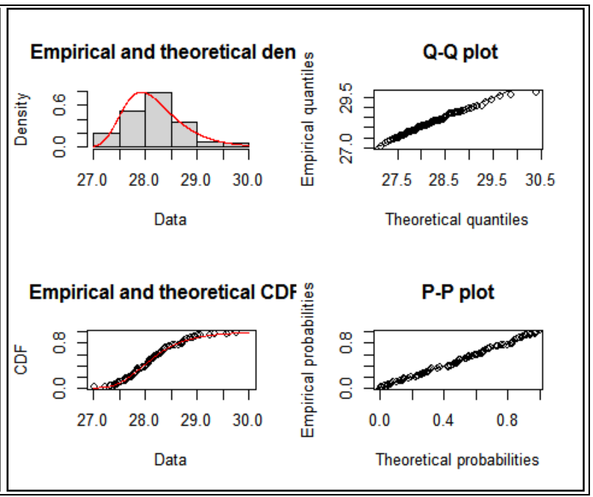


Figure 3(f): Fitted Reverse Gumbel distribution of September

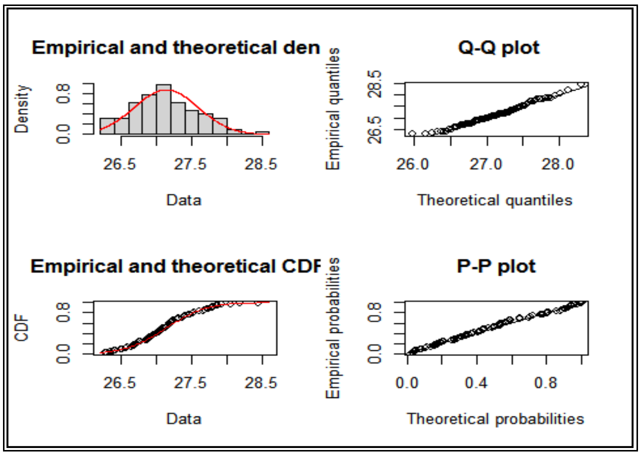


Figure 3(g): Fitted normal distribution of October

Q-Q plots and P-P plots pertaining to precipitation have been shown in Figure 4

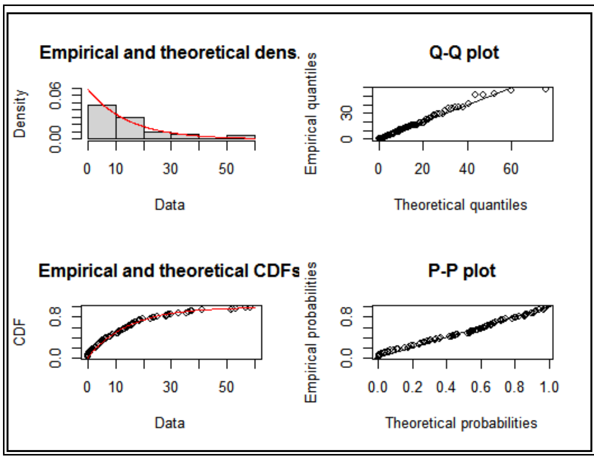


Figure 4(a): Fitted Exponential distribution of April

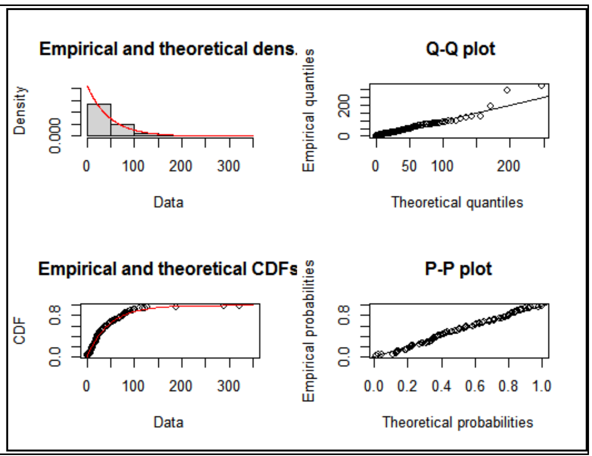


Figure 4(b): Fitted Exponential distribution of May

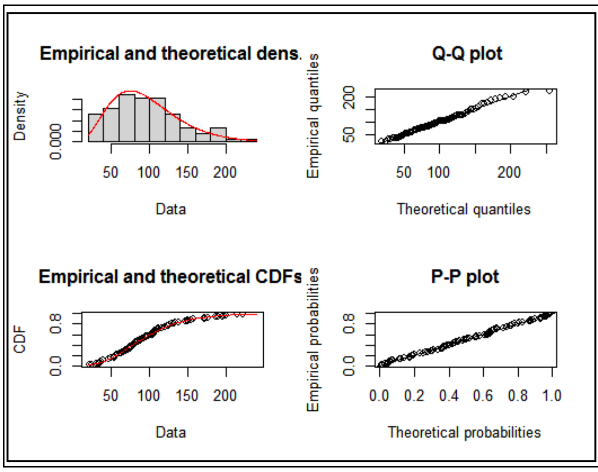


Figure 4(c): Fitted Gamma distribution of June

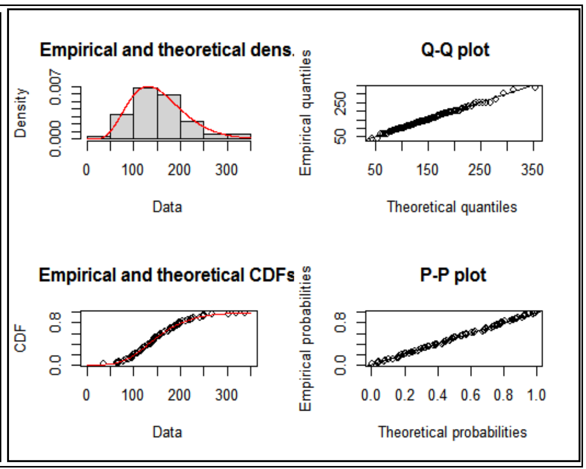


Figure 4(d): Fitted Gamma distribution of July

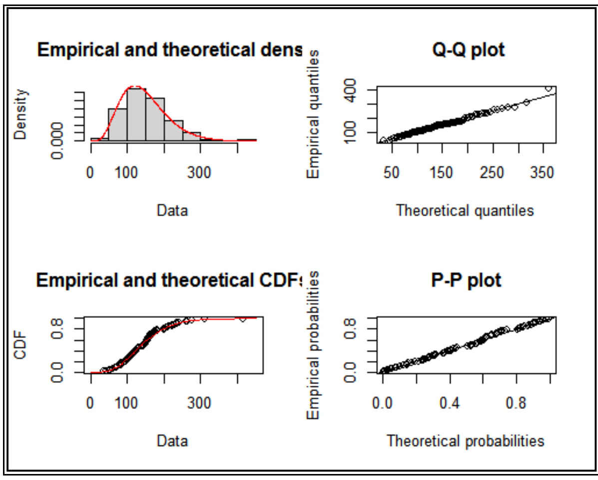


Figure 4(e): Fitted Gamma distribution of August

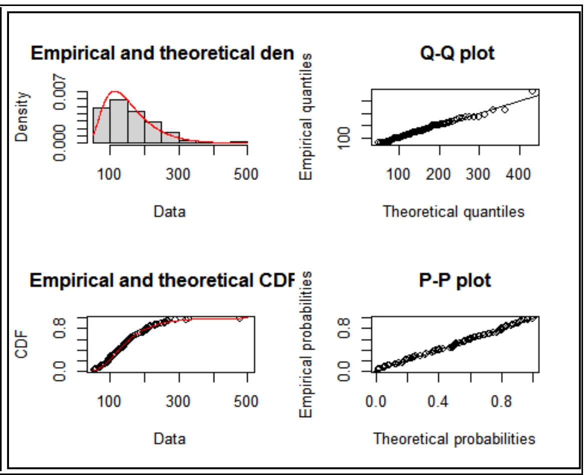


Figure 4(f): Fitted Inverse Gaussian distribution of September

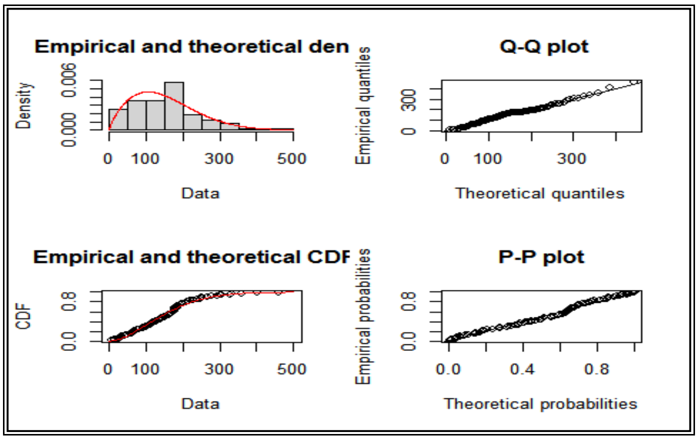


Figure 4(g): Fitted Weibull distribution of October

The graphs shown in Figures 3 and 4 indicate that the fitted distributions of Temperature and Precipitation of the Krishna district are very close to the observed data in each month.

D. Identification of bi-variate Copula

Using these fitted best distributions as Marginal distributions, we find a best fitted joint distribution, for each month, using Copula Analysis Technique, to estimate and forecast the Temperature. Maximum Likelihood Estimation is utilized to find parameter(s) estimates of fitted best Bi-variate Copula distribution. The best Copula distribution is determined using the minimum AIC and BIC criteria.

Table 4: Copula distribution parameter estimates

Month	Bivariate Copula						
	Distribution	**Parameter**	**AIC**	**BIC**	**MAPE**	**RMSE**	**NRMSE**
April	Gaussians	θ =–0.472	–18.019	–15.508	0.026	0.962	3.178
May	Rotated Gumbel 90 degrees	θ =–1.573	–30.998	–28.444	0.03	1.236	3.823
June	Rotated Tawn type 2 270 degrees	θ =–1.925 ω_1 = 0.539	–27.879	–22.75	0.035	1.392	4.477
July	Gaussian	θ = –0.532	–26.528	–23.963	0.026	0.955	3.317
August	Rotated Gumbel 270 degrees	θ =–1.570	–32.369	–29.805	0.024	0.830	2.924
September	Rotated Gumbel 90 degrees	θ =–1.626	–35.227	–32.663	0.024	0.873	3.097
October	Rotated Clayton 90 degrees	θ =–0.450	–7.529	–4.965	0.02	0.674	2.485

The error of estimation is calculated by using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Normalized Root Mean Square Error (NRMSE) has been presented in Table 4. The above table indicates that MAPE, RMSE and NRMSE are less than 5% for all the months. It is observed that in Krishna the average MAPE, RMSE and NRMSE for all months under consideration are 0.026, 0.989 and 3.329 respectively. It is implied that the variation in the observed Temperature data can be explained by these models with approximately 97.4% accuracy. Further, Table 5 shows that there is a similar relation between estimated values of Temperature and Precipitation as exhibited by the observed data.

Table 5: Relation between temperature and precipitation

Month	Apr	May	Jun	July	Aug	Sep	Oct
Observed data relation	– 0.305	– 0.342	– 0.285	– 0.232	– 0.173	– 0.289	– 0.178
Estimated data relation	– 0.318	– 0.445	– 0.317	– 0.261	– 0.215	– 0.371	– 0.213

4. Prediction

The MAPE, RMSE and NRMSE values being less than 5% indicates that the fitted best Bi-Variate Copula can be used for estimation and/or prediction of Temperature. Therefore, Temperature values for the Krishna district are estimated using Gaussian for the months of April and July, Rotated Gumbel 90 degrees for the months of May and September, Rotated Tawn type 2 270 degrees for the month of June, Rotated Gumbel 270 degrees for the month of August and Rotated Clayton 90 degree for the month of October Copula distributions for training as well as testing data sets.

Figure 5 presents the observed and predicted values for the testing period.

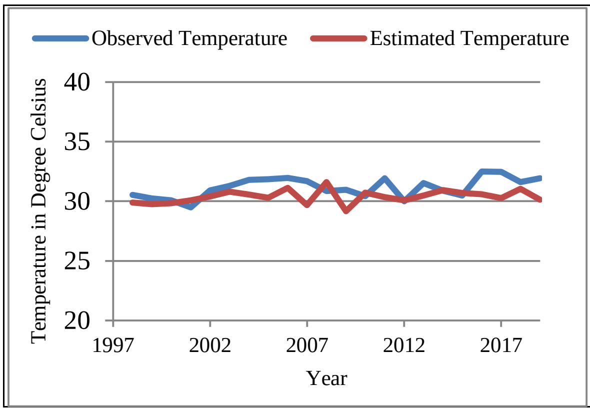


Figure 5(a): Krishna test data of April

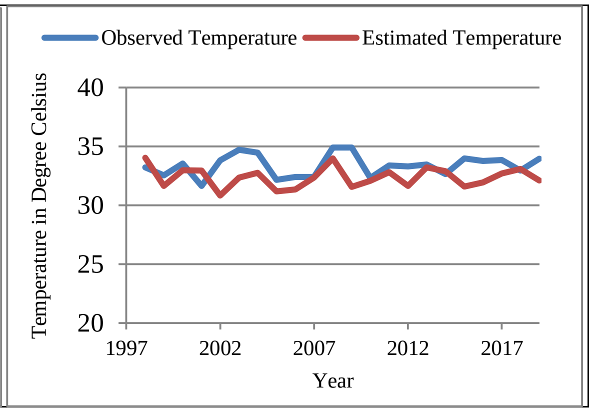


Figure 5(b): Krishna test data of May

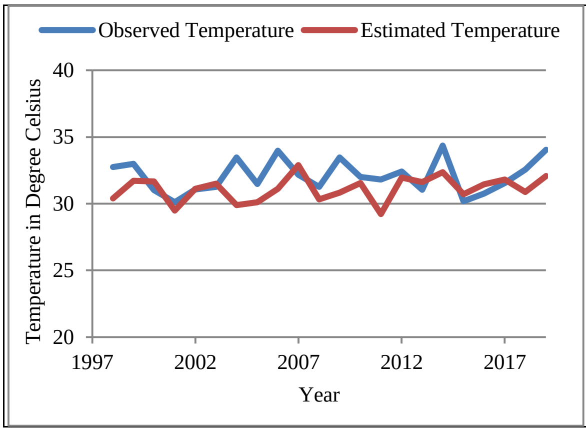


Figure 5(c): Krishna test data of June

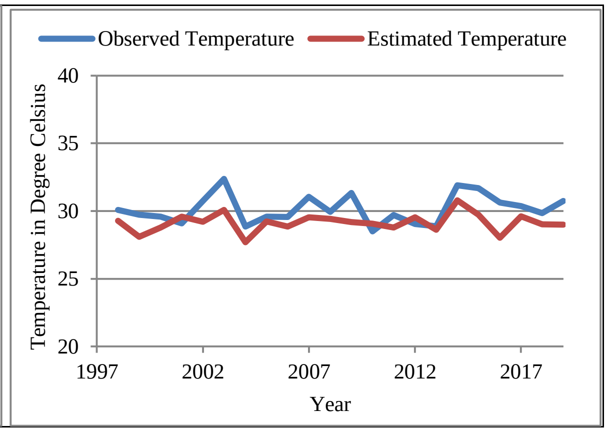
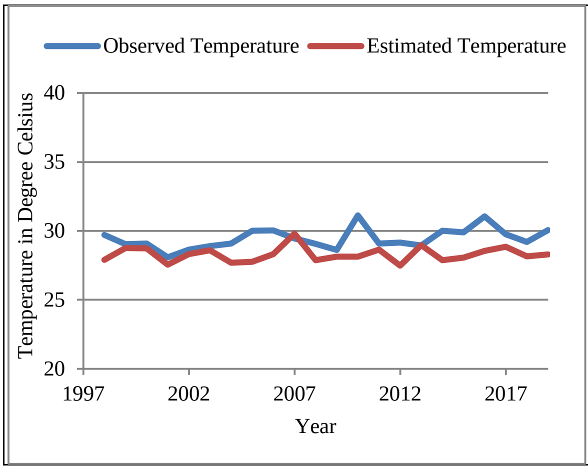


Figure 5(d): Krishna test data of July



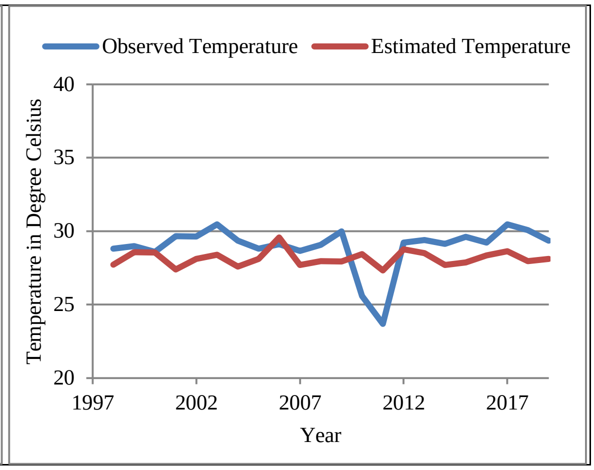


Figure 5(e): Krishna test data of August

Figure 5(f): Krishna test data of September

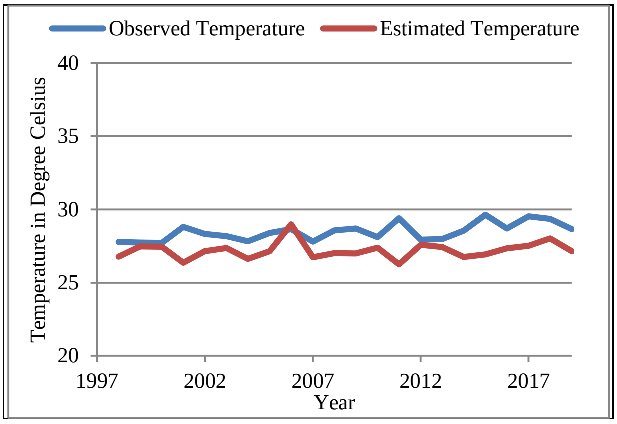


Figure 5(g): Krishna test data of October

The above graphs show that using the best fitted Copula distribution of April to October there is a reasonably good agreement in the patterns between observed and predicted Temperature values.

5. Conclusion

In this study, we have followed a novel approach to predict Temperature for Krishna district in Andhra Pradesh from April to October, by using Copula analysis. After the complete analysis, we could draw the following conclusions.

Using 80% of the collected data, initially, we identified the best fitted Probability distributions to the variables Temperature and Precipitation, separately. These distributions, in general, are different for different districts and months. The Probability density functions of these distributions are listed in Annexure 1.

Using the above identified best component distributions, we could identify the best fitted Joint Copula model that could predict month-wise Temperature for Krishna district in

Andhra Pradesh from April to October. As the climatic conditions change from region to region, the identified best Copula models are not the same for all the months in this district. The Probability density functions of best fitted Joint Copula distribution are listed in Annexure 2.

In order to establish the model adequacy of the best fitted Copulas, we computed AIC, BIC, MAPE, RMSE and NRMSE for the Krishna district from April to October, where ever the dependency is significant. It is observed that all the values of AIC and BIC are the least, all the values of MAPE are less than 5% and all the values of NRMSE are less than or around 5%. This establishes that the fitted models are adequate.

As the fitted models are adequate, using them we forecasted the values of Temperature for the time points in the testing data period. In the Krishna district, for all the months (where ever models are fitted), we could find a close agreement between observed and forecasted testing data.

Hence, it can be concluded that the identified best Copula models can be used for the prediction of future data points. Forecasted Temperature helps in planning agriculture and trading of commodities. This forecasting helps in deciding whether a crop has to be irrigated or not, should use fertilizers and whether it is a right time to complete harvesting *etc*.

References

Bezak, N., Zabret, K. and Sraj, M. (2018). Application of Copula functions for rainfall interception modelling. *Water*, **10**, 995.

Dzupire, N. C., Ngare, P. and Odongo, L. (2020). A Copula based bi-variate model for temperature and rainfall processes. *Scientific African*, **8**, 1-10.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. CRC press.

Lazoglou, G.and Anagnostopoulou, C. (2019). Joint distribution of temperature and precipitation in the Mediterranean, using the Copula method. *Theoretical and Applied Climatology*, **135**, 1399-1411.

Mesbahzadeh, T., Miglietta, M. M., Mirakbari, M., Soleimani Sardoo, F.and Abdolhoseini, M. (2019). Joint modeling of precipitation and temperature using Copula theory for current and future prediction under climate change scenarios in arid lands (Case Study, Kerman Province, Iran). *Advances in Meteorology*, 1-15.

Nelsen, R. B. (2007). *An Introduction to Copulas*. Springer Science and Business Media.

Olesen, J. E. and Bindi, M. (2002). Consequences of climate change for European agricultural productivity, land use and policy. *European Journal of Agronomy*, **16**, 239-262.

Pandey, P. K., Das, L., Jhajharia, D. and Pandey, V. (2018). Modelling of interdependence between rainfall and temperature using Copula. *Modeling Earth Systems and Environment*, **4**, 867-879.

Shaukat, M. H., Hussain, I., Faisal, M., Al-Dousari, A., Ismail, M., Shoukry, A. M., Elashkar, E.E. and Gani, S. (2020). Monthly drought prediction based on ensemble models. *PeerJ*, **8**, e9853.

Sklar, A. (1973). Random variables, joint distribution functions, and Copulas. *Kybernetika*, **9**, 449-460.

Vergara, O., Zuba, G., Doggett, T. and Seaquist, J. (2008). Modeling the potential impact of catastrophic weather on crop insurance industry portfolio losses. *American Journal of Agricultural Economics*, **90**, 1256-1262.

Yu, R., Yang, R., Zhang, C., Špoljar, M., Kuczyńska-Kippen, N. and Sang, G. (2020). A vine Copula-based modeling for identification of multivariate water pollution risk in an interconnected river system network. *Water*, **12**, 2741.

Zhang, L. and Singh, V. P. (2007). Bivariate rainfall frequency distributions using Archimedean Copulas. *Journal of Hydrology*, **332**, 93-109.

Zscheischler, J., Orth, R. and Seneviratne, S. I. (2017). Bivariate return periods of temperature and precipitation explain a large fraction of European crop yields. *Biogeosciences*, **14**, 3309-3320.

ANNEXURE 1

Density functions of the identified best fitted component (Temperature and Precipitation) distribution

Normal distribution

Normal distribution is a two-parameter distribution function and the parameterization of the normal distribution given in the function is

$$f(x/\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right); \text{ where } -\infty < x < \infty,\ -\infty < \mu < \infty \text{ and } \sigma > 0$$

Here μ and σ are mean and standard deviation of the distribution respectively.

Reverse Gumbel distribution

For positive skewed data the suitable distribution is Reverse Gumbel distribution. The probability density function of Reverse Gumbel distribution is

$$f(x/\mu,\sigma) = \frac{1}{\sigma} exp\left[-\left(\frac{x-\mu}{\sigma}\right) - exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right];$$

where $-\infty < x < \infty$, $-\infty < \mu < \infty$ and $\sigma > 0$

Here μ and σ are mean and standard deviation of the distribution respectively.

Logistic distribution

The Logistic distribution is suitable for moderate kurtosis data. The probability density function is given by

$$f(x/\mu,\sigma) = \frac{1}{\sigma}\left[exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right]\left[1 + exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right]^{-2}$$

where $-\infty < x < \infty$, $-\infty < \mu < \infty$ and $\sigma > 0$.

Weibull distribution

Weibull distribution is a two-parameter distribution function and the parameterization of the Weibull distribution given in the function is

$$f(x / \mu, \sigma) = \frac{\sigma x^{\sigma-1}}{\mu^{\sigma}} exp\left(-\frac{x}{\mu}\right)^{\sigma}; \text{ where } x > 0, \mu > 0 \text{ and } \sigma > 0$$

Here μ and σ are mean and standard deviation of the distribution respectively.

Exponential distribution

Exponential distribution is a one parameter distribution function and the parameterization of the Exponential distribution given in the function is

$$f(x/\mu) = \frac{1}{\mu} exp\left(-\frac{x}{\mu}\right); \text{ Where } x > 0, \mu > 0$$

Here μ is mean of the distribution respectively.

Gamma distribution

Gamma distribution is a one parameter distribution function and the parameterization of the Gamma distribution given in the function is

$$f(x/\mu, \sigma) = \frac{1}{(\sigma^2 \mu)^{1/\sigma^2}} \frac{x^{\left({}^{1}/_{\sigma^2}\right)-1} exp(\sigma^2 \mu)}{\gamma\left({}^{1}/_{\sigma^2}\right)};$$

where $x > 0$, $\mu > 0$ and $\sigma > 0$, μ and σ are mean and standard deviation of the distribution respectively.

Inverse Gaussian distribution

Inverse Gaussian distribution pdf is,

$$f(x/ \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2 y^3}} exp\left[-\frac{1}{2\mu^2\sigma^2 y}(x-\mu)^2\right]$$

for $x > 0, \mu > 0$ and $\sigma > 0$.

ANNEXURE 2

Density functions of identified best joint Copula distribution

Gaussian Copula

Gaussian Copula is defined as,

$$C(u, v) = \frac{1}{\sqrt{(1-\theta^2)}} e^{\left(\frac{\theta^2(u^2+v^2)-2\theta uv}{2(1-\theta^2)}\right)}$$

where, $-1 < \theta < 1$

Here, the Gaussian Copula parameter θ is given by $\theta = Sin\left[\frac{\pi}{2}\ \tau\right]$

Tawn Copula

The Tawn Copula is defined as

$$C(u,v) = (u,v)^{A(\alpha)} \text{ ; with } \alpha = \frac{ln(u)}{ln(uv)}$$

The Pickand function of Tawn Copula is given by

$$A(t) = (1-\omega_2)(1-t) + (1-\omega_1)t + \left[\left(\theta_1(1-t)\right)^{\alpha} + (\theta_2\, t)^{\alpha}\right]^{1/\theta}$$

where, $t \in [0,1]$, $0 \leq \omega_1$, $\omega_2 \leq 1$ and $\theta \in [0,\infty)$, the Tawn type 1 and type 2 refers to $\omega_1 = 1$ or $\omega_2 = 1$

Here, θ and ω_1 are the two parameters of Tawn Type 2 Copula, $\omega_2 = 1$, and the Parameter θ is given by $\tau = 1 - \theta^{-1}$.

$$\lambda_u = \omega_1 + 1 - \left(\omega_1^{\theta} + 1\right)^{\frac{1}{\theta}} \text{ or } \lambda_u = \omega_2 + 1 - \left(\omega_2^{\theta} + 1\right)^{\frac{1}{\theta}}$$

where Rotated 270 Copula means,

$$C^{270}(u,v) = C(v, 1-u)$$

Rotated Gumble Copula

Rotated Gumble Copula is defined as,

$$C(u,v) = e^{\left(-\left[(-\ln u)^{\theta} + (-\ln v)^{\theta}\right]^{\frac{1}{\theta}}\right)}$$

Here, the Gumbel parameter θ (≥ 1) is given by $\hat{\theta} = \frac{1}{1-\tau}$ and τ is the correlation between the variables. Here Rotated 270 Copula means,

$$C^{90}(u,v) = C(1-u, v)$$

Clayton Copula

Clayton Copula is defined as,

$$C(u,v) = max\left(\left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}, 0\right)$$

Here, the Clayton Copula parameter θ is given by $\tau = \frac{\theta}{\theta+2}$ and $\theta \in [-1, \infty)$

where Rotated 90 Copula means,

$$C^{90}(u,v) = C(1-u, v)$$