International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 5, Issue 07, July-2019

Analysis of Hydrological Variables using COPULA Models

AKSHAY KUMAR[1]

[1] *Department of Water Resources Engineering, Punjab Engineering College, Chandigarh, India*

Abstract— For some last decades, copula functions are being widely used in many areas of human interest including environment, energy, risk management, physical and natural sciences. Copula functions are considered to be the strongest methods and one of the most advanced methods than create a model between multivariate dependent variables that held dependency between each other. These functions combined number of different variables possessing different types of cumulative distribution functions. To establish more effective results for joint distribution best fit copula function for two dependent variables can be computed. In country like India, usage of copula functions are still not popular in the field of hydrology. In this review paper, we mainly focused on the growth of copula functions and their advancement in the areas of hydrological field. This review paper based on the applications of bivariate copula function by reviewing some of very important literatures based on copula function.

Keywords— Copula, Dependent structures, Hydrology, Marginal and joint distributions, Bivariate copula

I. INTRODUCTION

Copula functions are developed by Sklar in 1973.The need of these functions observed when the events start depending on multivariable and the need to establish results that are more robust observed. In many hydrological studies, many random variables contribute a significant role and the variables may or may not dependent on each other. However, it has observed that in many studies such variables are generally nonindependent. The combination of random variables (for ex: rainfall intensity and the duration of storm) generates flood showing different characteristics. The management of hydraulic structures or open channel management strongly depends on the combination of flood volume and its peak (Salvadori and De Michele, 2007). In case of drought analysis, the analysis considered by the combination of duration, magnitude and intensity of rainfall. Therefore, it is important to correlate the marginal distributions of dependent variables to obtain the joint effect of hydrological events.

In past studies, univariate cases was often considered to determine the return periods of flood events. The problem with that approach was, Sometimes this may lead to an over or under estimation of the risk of flood events. In reality, many events related to hydrology classified by the joint combination of many arbitrary variables, most of the time they are independent. A few examples include: the flood peak and volume (De Michele *et al.*, 2004), the storm duration and intensity(Salvadori and De Michele, 2004), drought events classified by intensity–magnitude–duration (Kim, Valdés and Yoo, 2003), and so on. Further applications in the copula world concern with the joint probability distributions of extreme hydrological events (Yue *et al.*, 1999; Yue, 2000; Yue and Rasmussen, 2002).

Many studies has been concern with the frequency of hydrological events. The high concern shows towards the hydrologic frequency analysis. There are number of researches that has dealt with the frequency of flood water peaks. In river management and hydraulic design, consideration of joint analysis of flood volume, flood peak and its duration plays a major role. For this, number of attempts had taken to implement a multivariate analysis that considered the significant dependence among above-mentioned variables (volume, duration, flood peak) but with some restrictive assumptions (Sackl and Bergmann, 1987; Singh and Singh, 1991; Yue, 2000). In a majority of the above studies, two basic assumptions were made. First, the variables follow the normal distribution. Second, the flood variables have the same type of the cumulative probability distribution. However, in practical application such assumptions does not hold true, flood variables are dependent, they do not have the same type of marginal distributions and do not follow the normal distribution (unless transformed to normal) (Zhang and Singh, 2006).Many studies has done for determining bivariate distributions for computing frequency of flood peaks using the copula function without using the above-mentioned assumptions. Copulas are the most considerable tools till now that can model dependency between various complex correlated variables.

II. METHODOLOGY

This paper focussed on use and growth of applications of copula functions has discussed using some important literatures. We will discuss the development in the field of bivariate probability distribution using standard and copula approach. Various studies in which such approaches have been used are discussed in the next section and using these studies, the development in the joint probability will be discussed. The various method used for obtaining the objectives so defined are discussed in detail with previous literature.

III. RESEARCH AND DEVELOPMENT IN THE FIELD OF BIVARIATE PROBABILITY DISTRIBUTION

A bivariate model by using two characteristics variables of direct runoff was presented by **Sackl and Bergmann, (1987)**. Flood peak and flood volume were the two variables and the bivariate normal distribution had selected for both the variables. The marginal distributions of both the samples transformed into normal distributions. The theoretical bi-normal distribution fitted and tested by using the equilines of probability density function. Conditional distributions were computed for one variable for a constant value of the other variable.

Exponential distributions play a central role in reliability, hydrology, life testing, and other fields of application**. Singh and Singh, (1991)** derived bivariate probability distributions with exponential marginal. The parameters of exponential distribution were estimated from the sample data. Rainfall data for 22 major storms that occurred during January 1986 and October 1987 in Singapore were used. The data included rainfall intensity and rainfall depth for various durations. Rainfall intensity and the corresponding depths were considered dependent random variables. Observed and computed probability distributions were compared. The correlation coefficient between the computed and observed conditional quantiles was 0.976.

Hydrological phenomena like flood flows always appear as multivariate events that are characterized by various components such as volume, duration, and flood peak. **Goel, N.K. *et al*., (1998)**developed a systematic methodology for multivariate modelling of flood flows using partial duration series approach and used bivariate normal distribution as distribution function. The Two Step Power transformation method was employed for normalization of the variables. The two step power transformation approach came after box cox transformation approach. The result showed that Box-Cox transformation reduced the coefficient of skewness to zero but it was unable to make the coefficient of kurtosis equal to 3. Because of this reason two step power transformation approach was used in the study. The application and validation of the methodology had done by using daily flows (1949 – 1979) of the Narmada River in Garudeshwar (India). Return period was obtained associated with flood peaks and flood volumes.

In many applied statistical fields, such as hydrology, the analysis of multivariate events is of particular interest. **Favre *et al.*, (2004)**presented an article for the modelling of multivariate extreme values using copulas. The methodology was applied on two different problems in hydrology. The first application concerned with the combined risk in the framework of frequency analysis. Four copulas were tested on peak flows from the watershed of Peribonka in Québec, Canada. The second application related to the joint modelling of peak flows and volumes. Three copulas were applied to the watershed of the Rimouski River in Québec, Canada. The present approach using copulas was found to be promising since choice of marginal distribution was not restricted.

Bivariate distributions of flood peak and volume, and flood volume and duration was derived by **Zhang *et al*., (2006)** using copula method. The copula method was applied to obtain the conditional return periods that were needed for hydrologic design. The derived distributions were tested using flood data from Amite River at Denham Springs, La., and the Ashuapmushuan River at Saguenay, Quebec, Canada. The derived distributions were also compared with the Gumbel mixed and the bivariate Box–Cox transformed normal distributions. The copula-based distributions were found to be in better agreement with the observed data than other distributions.

Applications of copula functions in hydrology is at initial stage. In many hydrological problems, several dependent variables are used in designing of hydrological structures or for calculating return periods of flood. One such study of presenting some recent advances in modeling that use copulas. **Salvadori and De Michele, (2007)** calculated the conditional probabilities, level curves of joint distribution and the return periods of bivariate events in both conditional and unconditional cases for the rainfall duration and average intensity of Bisagno river basin, Italy was obtained. The variables taken in the studies was nonindependent. To check the interdependency between the variables two nonparametric methods were used: Kendall's tau and Spearman's correlation coefficient. Several applications to hydrological data were shown.

The hydroclimatic teleconnection between hydrologic variables and large-scale atmospheric circulation is being investigated across the world. **Maity and Nagesh Kumar, (2008)**proposed a semiparametric copula-based approach to capture the dependency between teleconnected hydroclimatic variables for the prediction of response variable using the information of climate precursors namely, El Nino– Southern Oscillation and Equatorial Indian Ocean Oscillation. Usefulness of the proposed method was recognized in three distinct aspects and it was concluded that the proposed method could be applied to capture the relationship between teleconnected hydroclimatic variables having some linear or nonlinear cause-effect relationship. The dependence between them was captured and investigated for its potential use to predict the monthly variation of Indian summer monsoon rainfall using the proposed method. Predicted rainfall corresponded well with the observed rainfall with a correlation coefficient of 0.81 for the summer monsoon months, i.e., June to September. It was concluded that the method could be applied to similar analysis to assess the dependence between teleconnected hydroclimatic variables for other regions of the world and for different temporal scales such as seasonal.

Understanding of precipitation changes is the first step towards the consideration of hydrological responses to climate change and for the effective water resources management **(Zhang, Li and Singh, 2012)**.These changes are occurring due to high growth of human activities which resulting in hydrometerological changes. One such study held in china in which joint probability distribution of total rainfall and its duration with rainfall exceeding the 75^{th} percentile and falling below 25^{th} percentile was obtained. In the first case of rainfall exceeding 75^{th} percentile no variation was observed in 54 stations of Xinjiang, China. But in the second case(rainfall falling below 25^{th} percentile) kendall's coefficient varied a lot for different stations. In the evaluation of the data from 1955 - 2008 three change points were observed in both the cases. Copula functions got change for each subsequent series. Therefore, it was suggested that researchers should aware with the variation between subsequences and entire series, and a copula should be selected after the detection of change points to avoid possible bias in results or conclusions.

In Fuzhou City, a study was carried out to determine the impact of storm time and extreme precipitation and its change using a joint probabilistic approach by **KuiXu *et al.*, (2014**). The change point at the year of 1984 detected by Mann-

Kendall and Pettit's tests divides the extreme precipitation into two subsequence. Joint probabilities was calculated for each of the joint behaviour of extreme precipitation and storm tide. Results indicated that the joint probability had increased by more than 300% on an average after 1984 with a confidence limit of 95%. This study concluded that the installation of drainage systems should be required, considering both the parameters simultaneously in order to be prepared for the likely floods with minimum risk to loss of life and property and environment.

The problem of drought has been become a concerning issue in some parts of world. Several authors have investigated many researches on drought. Various drought indices (SPI, BMI etc.) have used to analyse droughts. A similar study on drought is done by **Cisty, Becova and Celar, (2016)** for southwestern Slovakia. In this study, importance of irrigation had been estimated by joint analysis of the duration and severity of the most demanding potential annual irrigation periods by bivariate copula method. The results using joint analysis approach indicated that, the suitability of the proposed method for calculation of frequency of irrigation needs, with greater benefits than the typical one-dimensional analysis of individual climatic variables. The results indicated that after every second year, moisture deficit of 30mm could expected which can significantly damage agricultural yields. The need of installation of irrigation structures was suggested in the study area.

IV. CONCLUSIONS

In some previous studies aforementioned above it was seen that most of studies was carried out using two variables whereas, most hydrological events are intrinsically multivariate. Additionally, in standard classical methods used in previous studies variables were assumed to follow normal marginal distribution, as the choice of distribution function was restricted for such methods. Thus to overcome these research gaps copula-based joint distribution function using multivariate hydrological variables should be used in the hydrological studies related to calculation of joint probability or return periods of flood flows. In hydrological studies, the use of copula have been applied to vast studies as discussed but in India, it is still at its initial stage. Furthermore copula models have been applied widely in studies for example, irrigation studies, earthquakes, hydrometeorology studies and rainfall forecasting. After going through all the studies discussed above it can be seen that the versatility of copula function. Copula functions are best for determining joint dependence among arbitrary dependent variables.

REFERENCES

[1] Cisty, M., Becova, A. and Celar, L. (2016) 'Analysis of Irrigation Needs Using an Approach Based on a Bivariate Copula Methodology', *Water Resources Management*. Water Resources Management, 30(1), pp. 167–182. doi: 10.1007/s11269-015-1153-1.

[2] Favre, A. C. *et al.* (2004) 'Multivariate hydrological frequency analysis using copulas', *Water Resources Research*, 40(1), pp. 1–12. doi: 10.1029/2003WR002456.

[3] Goel, N. K., Seth, S. M. and Chandra, S. (1998) 'MULTIVARIATE MODELING OF FLOOD FLOWS', *Journal of Hydraulic Engineering*, 124(2), pp. 146–155.

[4] Kim, T.-W., Valdés, J. B. and Yoo, C. (2003) 'Nonparametric Approach for Estimating Return Periods of Droughts in Arid Regions', *Journal of Hydrologic Engineering*, 8(5), pp. 237–246. doi: 10.1061/(ASCE)1084-0699(2003)8:5(237).

[5] Maity, R. and Nagesh Kumar, D. (2008) 'Probabilistic prediction of hydroclimatic variables with nonparametric quantification of uncertainty', *Journal of Geophysical Research*, 113(D14), p. D14105. doi: 10.1029/2008JD009856.

[6] De Michele, C. *et al.* (2004) 'Bivariate Statistical Approach to Check Adequacy of Dam Spillway', *Journal of Hydrologic Engineering*, 10(1), pp. 50–57. doi: 10.1061/(asce)1084-0699(2005)10:1(50).

[7] Sackl, B. and Bergmann, H. (1987) 'A Bivariate Flood Model and Its Application', in *Hydrologic Frequency Modeling*. Dordrecht: Springer Netherlands, pp. 571–582. doi: 10.1007/978-94-009-3953-0_40.

[8] Salvadori, G. and De Michele, C. (2004) 'Analytical calculation of storm volume statistics involving Pareto-like intensity-duration marginals', *Geophysical Research Letters*, 31(4), pp. 2–5. doi: 10.1029/2003GL018767.

[9] Salvadori, G. and De Michele, C. (2007) 'On the Use of Copulas in Hydrology: Theory and Practice', *Journal of Hydrologic Engineering*, 12(4), pp. 369–380. doi: 10.1061/(ASCE)1084-0699(2007)12:4(369).

[10] Singh, K. and Singh, V. P. (1991) 'Derivation of bivariate probability density functions with exponential marginals', *Stochastic Hydrology and Hydraulics*, 5(1), pp. 55–68. doi: 10.1007/BF01544178.

[11] Sklar, A. (1973) 'Random Variables, Joint Distribution Functions, and Copulas*', *Kybernetika*, 9(6), pp. 449–460. Available at: http://dml.cz/dmlcz/125838%0Ahttp://project.dml.cz.

[12] Xu, K. *et al.* (2014) 'Joint probability analysis of extreme precipitation and storm tide in a coastal city under changing environment', *PLoS ONE*, 9(10). doi: 10.1371/journal.pone.0109341.

[13] Yue, S. *et al.* (1999) 'The Gumbel mixed model for flood frequency analysis', *Journal of Hydrology*, 226(1–2), pp. 88–100. doi: 10.1016/S0022-1694(99)00168-7.

[14] Yue, S. (2000) 'The gumbel logistic model for representing a multivariate storm event', *Advances in Water Resources*, 24(2), pp. 179–185. doi: 10.1016/S0309-1708(00)00039-7.

[15] Yue, S. and Rasmussen, P. (2002) 'Bivariate frequency analysis: Discussion of some useful concepts in hydrological application', *Hydrological Processes*, 16(14), pp. 2881–2898. doi: 10.1002/hyp.1185.

[16] Zhang, L. and Singh, V. P. (2006) 'Bivariate Flood Frequency Analysis Using the Copula Method', *Journal of Hydrologic Engineering*, 11(2), pp. 150–164. doi: 10.1061/(asce)1084-0699(2006)11:2(150).

[17] Zhang, Q., Li, J. and Singh, V. P. (2012) 'Application of Archimedean copulas in the analysis of the precipitation extremes: Effects of precipitation changes', *Theoretical and Applied Climatology*, 107(1–2), pp. 255–264. doi: 10.1007/s00704-011-0476-y.