Multivariate modelling of meteorological droughts

E. Iliopoulos¹, V. Tsakiris², H. Vangelis¹, D. Tigkas¹ and G. Tsakiris¹*
¹ Centre for the Assessment of Natural Hazards and Proactive Planning, School of Rural and Surveying Engineering, National Technical University of Athens, Greece
² Department of Civil and Environmental Engineering, Imperial College London, United Kingdom
*e-mail: gtsakir@central.ntua.gr

Abstract: Drought is conventionally expressed as meteorological, hydrological and agricultural, dependent on the affected system and the variables used in the analysis. The main dimensions of the meteorological drought, apart from severity, are the duration of the event and the spatial extent of the affected area. The study presented in this paper attempts to perform frequency analysis for the pairs severity-duration and severity-areal extent, using 2D-Archimedean Copulas. The Gumbel-Hougaard, the Clayton and the Frank families of copulas have been tested. The proposed methodology is applied to the eastern part of the island of Crete - Greece. The area under study is divided in 1328 squares of 4 km² each. The sample of 30 years of precipitation and temperature data from 9 meteorological stations is reproduced 9 times. The Reconnaissance Drought Index is used for estimating the severity of each drought event. The duration is represented by the number of years under drought, whereas the areal extent is estimated by the number of squares under drought. The proposed methodology proves to be simple and practical and can contribute to more realistic decisions on measures against droughts by authorities and stakeholders.

Key words: Meteorological droughts, multivariate modelling, marginal probability distributions, drought severity, Reconnaissance drought index, Crete island

1. INTRODUCTION

Drought is a recurrent natural hazardous phenomenon affecting most parts of the world. Drought is conventionally expressed as meteorological, hydrological and agricultural, dependent on the affected system and the variables used in the analysis. The main dimensions of the meteorological drought are the severity and duration of the event, and the spatial extent of the affected area. Thus, for the frequency analysis it is appropriate to take into account all the three dimensions.

The study presented in this paper attempts to perform frequency analysis of the pairs “severity-duration” and “severity-areal extent” using 2D-Archimedean Copulas. Using meteorological data from the eastern part of the island of Crete, the 2D Gumbel-Hougaard, the Clayton and the Frank families of copulas have been tested.

2. BASIC NOTIONS

Droughts, as several other natural hazards, are multidimensional phenomena. Since drought episodes are stochastic in nature, numerous studies have been reported analysing droughts using probabilistic theories (Gupta and Duckstein 1975; Kendall and Dracup 1992; Rossi et al. 1992; Vangelis et al. 2011; Bonaccorso et al. 2015; and many others). There are several studies in the literature analysing the various dimensions of drought (e.g. Dalezios et al. 2000; Reddy and Ganguli 2012; Kim et al. 2011; Reddy and Singh 2014; Saghaian et al. 2003; Shiau and Modarres 2009; Todisco et al. 2013). Most of these studies have analysed droughts either as univariate phenomenon or they have analysed each drought characteristic separately. In this context, the double frequency analysis seems more appropriate, provided that the probability distributions (marginal distributions), describing each of the variables (dimensions) analysed, belong to the same family of
probability distributions. It is easily understood, that it is most unlikely that the both dimensions of drought (e.g. severity and duration or severity and areal extent) follow the same type of probability distribution in all possible applications. As has been shown in many occasions, the drought variables follow different probability distributions. Therefore, to overcome this problem, the copula approach can be implemented. As known, copulas are functions joining univariate distribution functions from any family into a single multivariate distribution function. From the first introduction of copulas by Sklar (1959), several important studies have been published with applications in various fields, among which is also hydrology (Salvadori et al. 2007 and Genest and Favre 2007). Interesting applications on droughts have been also made by Shiau (2006), Yusof et al. (2013), Chen et al. (2013), Bazrafshan et al. (2015), Huang et al. (2015), Vergni et al. (2015) and others. There are several types of copula functions, which are described in books (e.g. Nelsen 1999). The most popular copulas for hydrologic analysis are those of the Archimedean family, which are characterised by their simple structure and strong representativeness (Huang et al. 2015; Tsakiris et al. 2015).

3. APPLICATION

The application presented in this paragraph is partially based on the data used in a recent paper (Tsakiris et al. 2016). The area studied is the eastern part of the island of Crete which is one of the main drought prone areas of Greece. The area was divided into 1328 squares of 4 km² each. The monthly meteorological data, concerning precipitation and temperature, collected from 9 stations in the area, cover the period of 30 years (1962-1991). The meteorological data were transferred to the center of the squares using the method of inverse distance (Wei and McGuiness, 1973). As a result, annual series of data were produced, creating 2 matrices with dimensions 1328 x 30.

The available data were reproduced for additional 270 years using a simple stochastic process (AR 1). The synthetic time series for areal average precipitation and potential evapotranspiration are presented in Figures 1 and 2.

![Figure 1. The annual precipitation of the area generated from the 30-year sample](image)

Regarding the dimensions of drought (severity, areal extent and duration), they were calculated for all 300 years. Severity was estimated using the Reconnaissance Drought Index (RDI; Tsakiris and Vangelis 2005; Tsakiris et al. 2007). As known, RDI uses precipitation and potential evapotranspiration data. The areal extent is represented by the total number of squares under drought whilst duration by the number of consecutive years under drought. Figures 3 and 4 depict...
the severity and areal extent, whereas the duration dimension is presented in tables not included in this short version of the paper.

Figure 2. The annual potential evapotranspiration of the area generated from the 30-year sample

Figure 3. The severity of drought for the 300 years represented by the sum of absolute RDI values

Figure 4. The areal extent for the 300 years represented by the total number of squares under drought
The multivariate modelling of droughts was achieved using 2D Archimedean copulas for the pairs severity-duration and severity-area extent. The copula families tested were Gumbel-Hougaard, Clayton and Frank, with equations presented in Table 1.

<table>
<thead>
<tr>
<th>Archimedean Copula</th>
<th>Joint Probability Distribution Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumbel–Hougaard</td>
<td>[ \begin{aligned} F_{xy} &amp;= C_{\phi}(F(x), F(y)) = \varphi^{-1}\left( \varphi(F(x)) + \varphi(F(y)) \right) = \exp\left[ \left(-\ln F(x)\right)^\theta + \left(-\ln F(y)\right)^\theta \right] \end{aligned} ]</td>
</tr>
<tr>
<td>Clayton</td>
<td>[ \begin{aligned} F_{xy} &amp;= C_{\phi}(F(x), F(y)) = \varphi^{-1}\left( \varphi(F(x)) + \varphi(F(y)) \right) = \left[ (F(x))^{-\theta} + (F(y))^{-\theta} - 1 \right]^{\frac{1}{\theta}} \end{aligned} ]</td>
</tr>
<tr>
<td>Frank</td>
<td>[ \begin{aligned} F_{xy} &amp;= C_{\phi}(F(x), F(y)) = -\theta^{-1}\ln\left[ 1 + \frac{e^{\theta\varphi(x)} - 1}{e^{\theta\varphi(y)} - 1} \right] \end{aligned} ]</td>
</tr>
</tbody>
</table>

The parameter \( \theta \) was estimated through the Kendall’s \( \tau \) as shown in Table 2 for all three Archimedean copulas (Tsakiris et al. 2016). The following table gives the expressions of the generator \( \varphi \) and the Kendall’s \( \tau \), for the most popular Archimedean families, such as Gumbel-Hougaard, Clayton and Frank.

<table>
<thead>
<tr>
<th>Family</th>
<th>Generator ( \varphi )</th>
<th>Kendall’s ( \tau )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumbel-Hougaard</td>
<td>((-\ln(I))^\theta)</td>
<td>(1 - \frac{1}{\theta})</td>
</tr>
<tr>
<td>Clayton</td>
<td>((I^{-\theta} - 1)/\theta)</td>
<td>(\theta/\theta + 2)</td>
</tr>
<tr>
<td>Frank</td>
<td>(-\ln\left(\frac{e^{\theta - 1}}{e^{\theta} - 1}\right))</td>
<td>(1 - \frac{4}{\theta} + 4D_1(\theta))</td>
</tr>
</tbody>
</table>

In Table 2, \( D_1(\theta) = \int_0^\theta \left( \frac{x}{\theta} \right) \left( e^x - 1 \right) dx \), which is the first Debey function.

4. RESULTS AND DISCUSSION

For the construction of Archimedean copulas for the pairs ‘severity – areal extent’ and ‘severity – duration’, the following results were derived (Table 3).

<table>
<thead>
<tr>
<th>Pair</th>
<th>Kendall’s ( \tau )</th>
<th>( \theta_{\text{Gumbel-Hougaard}} )</th>
<th>( \theta_{\text{Clayton}} )</th>
<th>( \theta_{\text{Frank}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity-Areal Extent</td>
<td>0.598</td>
<td>2.488</td>
<td>2.975</td>
<td>1.420</td>
</tr>
<tr>
<td>Severity-Duration</td>
<td>0.191</td>
<td>1.236</td>
<td>0.472</td>
<td>1.380</td>
</tr>
</tbody>
</table>

Based on the results of Frechet-Hoeffding test, it was concluded that the Frank copula was unsuitable for modelling the set of data used, whereas Gumbel-Hougaard copula seemed to perform more successfully.

The 3D graphs presented on the Figures 5 and 6 illustrate the joint distribution function of Gumbel-Hougaard copulas for the pair ‘severity – areal extent’ and the ‘severity – duration’, respectively. Similarly, 3D graphs have been produced for the Clayton copulas for the same pairs, respectively.
Finally, Figures 7 and 8 show a convenient way for calculating the joint probability for both pairs severity-areal extent and severity-duration. In the graphs $S$ stands for severity (represented by absolute RDI value), $A$ for the areal extent (represented by the total number of squares under drought), and $D$ for the duration (in years).

5. CONCLUDING REMARKS

Using the Archimedean copulas, the joint distribution functions were derived for the pairs of drought events, ‘severity – areal extent’ and ‘severity – duration’. It is concluded that by this 2D frequency analysis, a more comprehensive assessment of droughts can be achieved. As a result, the proposed approach can assist authorities and stakeholders to base their preparedness planning on
more secure information by considering more than one dimension of drought.

**Figure 7. Joint Cumulative probability distribution function H for the pair ‘severity – duration’**

**Figure 8. Joint Cumulative probability distribution function H for the pair ‘severity – areal extent’**

**REFERENCES**


Genest C, Favre AC (2007) Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering 12(4):347-368


