Hydrol. Earth Syst. Sci. Discuss., 10, 7469–7516, 2013
www.hydrol-earth-syst-sci-discuss.net/10/7469/2013/
doi:10.5194/hessd-10-7469-2013

Hydrology and Earth System Sciences
Open Access
Discussions

A copula-based assessment of Bartlett–Lewis type of rainfall models for preserving drought statistics

M. T. Pham[1], **W. J. Vanhaute**[1], **S. Vandenberghe**[1], **B. De Baets**[2], **and N. E. C. Verhoest**[1]

[1]Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, 9000 Ghent, Belgium
[2]Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, 9000 Ghent, Belgium

Received: 27 May 2013 – Accepted: 30 May 2013 – Published: 13 June 2013

Correspondence to: M. T. Pham (minhtu83@gmail.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Of all natural disasters, the economic and environmental consequences of droughts are among the highest because of their longevity and widespread spatial extent. Because of their extreme behaviour, studying droughts generally requires long time series of historical climate data. Rainfall is a very important variable for calculating drought statistics, for quantifying historical droughts or for assessing the impact on other hydrological (e.g. water stage in rivers) or agricultural (e.g. irrigation requirements) variables. Unfortunately, time series of historical observations are often too short for such assessments. To circumvent this, one may rely on the synthetic rainfall time series from stochastic point process rainfall models, such as Bartlett–Lewis models. The present study investigates whether drought statistics are preserved when simulating rainfall with Bartlett–Lewis models. Therefore, a 105 yr 10 min rainfall time series obtained at Uccle, Belgium is used as test case. First, drought events were identified on the basis of the Effective Drought Index (EDI), and each event was characterized by two variables, i.e. drought duration (D) and drought severity (S). As both parameters are interdependent, a multivariate distribution function, which makes use of a copula, was fitted. Based on the copula, four types of drought return periods are calculated for observed as well as simulated droughts and are used to evaluate the ability of the rainfall models to simulate drought events with the appropriate characteristics. Overall, all Bartlett–Lewis type of models studied fail in preserving extreme drought statistics, which is attributed to the model structure and to the model stationarity caused by maintaining the same parameter set during the whole simulation period.

1 Introduction

Drought impacts on the environment and on the economy are among the highest of all natural disasters due to their long-term and extensive scale (Wilby and Wigley, 2000). It is often stated that drought is one of the most complex natural hazards, and that it

HESSD
10, 7469–7516, 2013

affects more people than any other hazard (Wilhite et al., 2007). Understanding drought statistics, therefore, is essential for planning and management of water resources.

There are two main challenges with respect to the statistical analysis of droughts. Unlike extreme rainfall or flood problems, drought may last from several months to years; therefore, the first challenge consists of retrieving a historical climate data set which is sufficiently long for analysis. Precipitation data, being the most important variable for drought investigation, is not always available from observations for such a long period. In the latter case, one may consider the use of stochastic point process rainfall models, which allow for generating extreme long rainfall time series with similar statistics as what was observed (Verhoest et al., 2010).

The second challenge concerns the characterization of the dependence between the different variables that define a drought. Droughts are generally characterized by multiple attributes (Shiau and Modarres, 2009; Wong et al., 2010), of which drought duration and severity are the two most important variables in the majority of the reported drought research. Generally, traditional univariate analysis does not account for any correlation between these variables. However, the description of the extremity of an event depends upon the combination of duration as well as severity. A multivariate frequency analysis of droughts can be explored using copulas (Shiau and Shen, 2001; Shiau, 2003, 2006; Kim et al., 2006a,b; Shiau and Modarres, 2009; Wong et al., 2010). Copulas, proposed by Sklar (1959), are multivariate distribution functions that allow for the description of the dependence structure between random variables independently of their marginal behaviours (Genest and Favre, 2007; Salvadori and De Michele, 2010; Shiau, 2006; Shiau and Modarres, 2009; Vandenberghe et al., 2011). These functions have already been widely used to investigate dependence structures between hydrological variables, albeit mostly in the field of rainfall and flood problems (Shiau and Modarres, 2009). This study aims at investigating whether rainfall series simulated by the Bartlett–Lewis (BL) type of models can be used for drought analysis, as it can be questioned whether those models are able to preserve drought statistics. Therefore, a time series of 105 yr (1898–2002) of 10 min rainfall observed at Uccle, located near

Brussels, Belgium, is used for comparison. First, a drought index, i.e. the EDI index (Byun and Wilhite, 1999), is calculated for the observed time series. Then drought events are selected based on an investigation of the EDI, and are characterized by two variables, namely drought duration (D) and severity (S), whose dependence is modelled using a copula. A similar analysis is performed on simulated time series obtained from different types of Bartlett–Lewis models. Finally, four types of copula-based return periods for drought are calculated for both the observed and simulated time series. Through a comparative analysis of the results, the ability of the Bartlett–Lewis models for preserving drought statistics is assessed.

The following gives an overview of the structure of the paper. Section 2 introduces the observed rainfall time series and the five Bartlett–Lewis models used in this research. Section 3 describes the drought index used. Section 4 briefly introduces copulas and explains the calculation method of bivariate copula-based return period for drought. Section 5 presents results of marginal distribution fitting, copula selection and comparison of four types of drought return periods between observed and simulated data. Finally, conclusions and recommendations for further study are given in Sect. 6.

2 Observed and modelled data

In order to evaluate whether the selected BL rainfall models can preserve the drought statistics, this study conducts a comparison between an observed and several synthetic time series for Uccle, Belgium, simulated by the rectangular process rainfall models. The observed time series is a precipitation record with a time resolution of 10 min from 1 January 1898 to 31 December 2002 measured by a Hellmann–Fuess pluviograph in the climatological park of the Royal Meteorological Institute at Uccle, near Brussels, Belgium (Démarée, 2003). Analysing this series with the adaptive Kolmogorov–Zurbenko (KZA) filter, De Jongh et al. (2006) had detected droughts around 1920 and during the mid-1970s and drier-than-normal conditions at the beginning of the twentieth century. Five time series of 105 yr of 10 min rainfall were simulated by five versions

HESSD
10, 7469–7516, 2013

of the Bartlett–Lewis Rectangular Pulses Model. In these models, clusters of possibly overlapping rectangular pulses having a random length (duration) and height (intensity) are generated. Finally, a time series of rainfall is obtained through discretizing time into intervals of a given length (in this study, 10 min), and cumulating the rainfall volumes of the rectangles that fall within each interval. The applied versions of the BL models are the original Bartlett–Lewis (OBL) model (Rodriguez-Iturbe et al., 1987a), the modified Bartlett–Lewis (MBL) model (Rodriguez-Iturbe et al., 1988), the modified Bartlett–Lewis Gamma (MBLG) model (Onof and Wheater, 1994), the truncated Bartlett–Lewis (TBL) model (Onof et al., 2013) and the truncated Bartlett–Lewis Gamma (TBLG) model (Onof et al., 2013).

The OBL model was first proposed by Rodriguez-Iturbe et al. (1987a), in which storm events are generated randomly according to a Poisson process with parameter λ. Each storm origin is followed by a sequence of cell origins, modelled by a second Poisson process characterized by parameter β. Cells can be generated during a time interval having a duration which is exponentially distributed with parameter γ. For each cell origin, a rectangular cell is generated with random depth and duration, both exponentially distributed respectively with parameters $1/\mu_x$ and $1/\eta$. Rodriguez-Iturbe et al. (1987a) introduced dimensionless parameters $\kappa = \beta/\eta$ and $\phi = \gamma/\eta$ to ensure that the number of cell origins associated with one storm arrival follows a geometrical distribution with mean $\mu_x = 1 + \kappa/\phi$. As such, the OBL model has five parameters $(\lambda, \beta, \gamma, \mu_x, \eta)$. The OBL model, however, showed some shortcomings with respect to the preservation of the zero-depth probabilities (Rodriguez-Iturbe et al., 1987a,b). To solve this, the modified Bartlett–Lewis (MBL) model was introduced (Rodriguez-Iturbe et al., 1988) allowing the average cell duration to vary between storms by modifying the exponentially distributed cell duration through randomizing the parameter η according to a Gamma distribution with shape parameter α and scale parameter ν. The mean cell inter-arrival time, β^{-1}, and the mean storm duration time, γ^{-1}, can be varied randomly by keeping κ and ϕ constant while varying η. The MBL thus has six parameters $(\lambda, \mu_x, \alpha, \nu, \kappa, \phi)$. Since the MBL poorly reproduced the extreme behaviour of rainfall, Onof and Wheater

(1994) introduced the modified Bartlett–Lewis Gamma (MBLG) model as an updated version of MBL, in which the cell depth follows a two-parameter gamma distribution, with shape parameter p and scale parameter δ, resulting in a model with seven parameters $(\lambda, \alpha, \nu, \kappa, \phi, p, \delta)$.

A problem that remained unnoticed by many users of the aforementioned models is that very unrealistic events of excessively large cells are occasionally generated when sampling a large number of mean cell duration values during long simulations (typically, when simulation is much longer than the size of the data set) (Verhoest et al., 2010). To surpass this issue, the Gamma distribution for the sampling of η is truncated to inhibit the sampling of extremely large mean cell durations (Onof et al., 2013). The truncation parameter ε can be handled as an extra parameter during calibration. As with the MBL and MBLG, the truncated model can use either an exponential or gamma distribution to represent rainfall depth; these will be referred to as the TBL and TBLG model, respectively. TBL and TBLG respectively have seven $(\lambda, \mu_x, \alpha, \nu, \kappa, \phi, \varepsilon)$ and eight parameters $(\lambda, \alpha, \nu, \kappa, \phi, p, \delta, \varepsilon)$.

The models are calibrated using the Generalised Method of Moments (GMM) in which model parameters are chosen to minimize the difference between the model values calculated with the available analytical expressions and the empirical values of these statistics obtained from observed data. The fitting procedure is subject to a certain level of subjectivity. There are no general guidelines about the choice of moments or aggregation levels in the objective function, nor is there a general consensus about the weights used during fitting (Vanhaute et al., 2012). Several approaches exist, each exhibiting certain advantages and disadvantages, which make it hard to select one particular method for calibration. For example, attributing greater weight to the mean rainfall intensity during fitting will lead to a better reproduction of mean rainfall intensity, whereas other properties may be reproduced less accurately, as a result. Evidently, this extends to the choice of the included moments and their aggregation levels during fitting. Several authors have attempted to address some of these issues empirically which has led to differing conclusions, making it particularly hard to assess the merit of a certain method, since the inferred conclusions are obviously influenced by the chosen evaluation criteria (Burton et al., 2008; Cowpertwait et al., 2007; Vanhaute et al., 2012). The chosen fitting properties in the current work include the hourly mean and third order moment, and the variance, lag-1 auto-covariance and proportion of dry intervals or zero depth probability (ZDP) at time scales of 10 min, 1 h and 24 h. This is similar to the fitting properties chosen by Cowpertwait et al. (2007). The corresponding empirical variances (which can be obtained by treating yearly observed statistics as repetitions of a given rainfall statistic) are used as weights in the objective function (Jesus and Chandler, 2011). The aforementioned calibrations are realised using the Shuffled Complex Evolution algorithm (Duan et al., 1994). This method has shown to be a reliable and easy-to-use method, when compared to other heuristic optimization algorithms (Vanhaute et al., 2012).

3 Drought index

The objective of this study is to assess whether Bartlett–Lewis type of models are able to preserve drought statistics. These models have already extensively been validated for general statistics, such as mean, variance, auto-covariance and zero depth probability as well as for extreme precipitation events (Cameron et al., 2000, 2001; Kaczmarska, 2011; Vandenberghe et al., 2011; Verhoest et al., 1997; Wheater et al., 2005). However, an assessment on whether the statistics of the alternative extreme behaviour, i.e. low precipitation, to the authors' knowledge, has not been reported in literature.

Different forms of drought can be identified, including meteorological drought, which is defined on the basis of the severity and the duration of the dry spell, agricultural drought, which accounts for the agricultural impact of the drought event, and hydrological drought, in which the impact of reduced precipitation on surface or subsurface water supply is taken into account. Given the fact that agricultural and hydrological droughts also include effects of land use, soil management, hydrological characteristics

HESSD
10, 7469–7516, 2013

of a catchment and water management, we do not focus on these types of drought as comparing their drought statistics may not merely be attributed to shortcomings in the rainfall time series. As such, we focus on meteorological drought indices that are solely based on the rainfall time series.

Several types of meteorological drought indices have been proposed in literature, including the Rainfall Anomaly index (RAI) (Rooy, 1965), the Bhalme and Mooly Drought index (BMDI) (Bhalme and Mooley, 1980), the Standardized Precipitation index (SPI) (McKee et al., 1993), the National Rainfall index (NRI) (Gommes and Petrassi, 1994), the Effective Drought index (EDI) (Byun and Wilhite, 1999), and the Drought Frequency index (DFI) (González and Valdés, 2006). In this study, we opt for the EDI proposed by Byun and Wilhite (1999), as this index can be calculated on a daily time basis (Morid et al., 2006), whereas most of the other drought indices are calculated at a monthly scale, which renders them less interesting for assessing the temporal behaviour of BL models. For easier calculation of the drought index, all years are considered to have 365 days; for the leap years, the day 29 February is excluded and the rainfall on 28 February is recalculated as the average rainfall observed on 28 and 29 February.

The first step in the calculation procedure of EDI is to calculate the Effective Precipitation (EP). The EP refers to the cumulative daily precipitation with a time-dependent reduction function (Kim et al., 2009); in other words, the EP for any day j is a weighted sum of the precipitation of the I previous days with decreasing weights (Morid et al., 2006) where j is the number of the days since the beginning of the time series. For values $j > I$, EP is calculated as:

$$\mathrm{EP}(j) = \sum_{k=1}^{I}\left[\left(\sum_{m=1}^{k} P(j-m)\right)/k\right] \tag{1}$$

where $P(j-m)$ is the precipitation at m-th day before day j.

The duration I is usually chosen as 365 days (Dogan et al., 2012; Byun et al., 2008; Kim et al., 2009; Morid et al., 2006, 2007; Pandey et al., 2008; Yu-Won and Hi-Ryong, 2006), as a representative value of the total water resources stored for a long period

HESSD
10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

(Morid et al., 2006) or the most common precipitation cycle (Kim et al., 2009). The same value of I is also taken in this study. Since I is chosen as 365 days, the first year of all data sets is used to calculated the EP for the second year, therefore the EDI values are just available for the final 104 yr data.

Once the EP is obtained for each day, its deviation, DEP, with respect to a mean EP, i.e. MEP is calculated:

$$\mathrm{DEP}(j) = \mathrm{EP}(j) - \mathrm{MEP}(j) \tag{2}$$

where MEP(j) is the mean value of the EP values of the days $j' \equiv j(\mathrm{mod}\,365)$ (e.g. 15 May) of all years in a standard period of 30 yr or more (Byun and Wilhite, 1999; Kim et al., 2009; Morid et al., 2006; Pandey et al., 2008). In this study, the standard period of 30 yr from 1971 to 2000 is applied for Uccle observations and from year 71 to 100 for BL simulations.

Finally, the EDI for each day is calculated:

$$\mathrm{EDI}(j) = \frac{\mathrm{DEP}(j)}{\mathrm{SD}(\mathrm{DEP}(j))} \tag{3}$$

in which SD (DEP(j)) is the standard deviation of DEP of the days $j' \equiv j(\mathrm{mod}\,365)$ of all years over the standard period.

The classification of the drought severity by the EDI is presented in Table 1. For a more detailed explanation of the EDI calculation procedure, we refer to Byun and Wilhite (1999) and Kim et al. (2009).

We also use the Yearly Accumulated negative EDI (YAEDI_{365}) proposed by Kim et al. (2009), which represents annual average dryness. YAEDI_{365} for year t is calculated as:

$$\mathrm{YAEDI}_{365}(t) = \frac{\sum\limits_{j=(t-1)365+1}^{365t} \min(\mathrm{EDI}(j), 0)}{365} \tag{4}$$

HESSD
10, 7469–7516, 2013

where $t = 2, 3, \ldots, 105$.

In this research, a drought event is defined as an extremely dry to moderately dry period during which the EDI index is continuously less than −0.70 (see Table 1). Each drought is characterized by two dependent attributes: duration, D, and severity, S, where the latter is the cumulative value of the absolute value of the EDI within the drought event, i.e:

$$S = \sum_{j=T}^{D+T} |\mathrm{EDI}(j)| \tag{5}$$

where T is the value of j at the onset of a drought event, i.e. the day at which the EDI value becomes less than −0.70.

4 Bivariate copula-based drought analysis

The investigation and frequency analysis of droughts are of key importance in the planning and management of water resources. These studies, however, cannot be performed by traditional univariate parametric methods given the fact that droughts are characterised by a number of correlated variables. Univariate analyses, which investigate each variable separately, do not account for the relationship between those variables (Kim et al., 2003) and may lead to an over- or underestimation of return period calculation (Li et al., 2012). An analysis based on the multivariate distribution, for which copulas, as proposed by (Sklar, 1959), can be used, can overcome these problems (Shiau, 2006). Copulas are functions that couple univariate distribution functions into a multivariate distribution function. The merit of using copulas is that the dependence structure between variables can be modelled independently of their marginal distribution functions. Copula-based drought frequency analysis can be performed by first fitting a copula to the drought variables, D and S, and by then calculating bivariate drought return periods.

HESSD
10, 7469–7516, 2013

4.1 Fitting drought duration and severity with copulas

In order to perform a copula-based frequency analysis, a bivariate distribution function of the drought duration D and drought severity S needs to be characterized. According to the theorem of Sklar (Sklar, 1959), if $F_{DS}(d,s)$ is a two-dimensional distribution function of D and S with marginal distributions $F_D(d)$ and $F_S(s)$, then there exists a copula C such that

$$F_{DS}(d,s) = C(F_D(d), F_S(s)) = C(u,v) \tag{6}$$

Conversely, for any univariate distributions $F_D(d)$ and $F_S(s)$ and any copula C, the function $F_{D,S}(d,s)$ defined above is a two-dimensional distribution function with marginal distributions $F_D(d)$ and $F_S(s)$. The second equality in Eq. (6) describes a transformation based on the invariance property of copulas (Genest and Rivest, 2001), in which the marginal distribution functions $F_D(d)$ and $F_S(s)$ of the variables D and S are transformed into values U and V in the unit interval $I = [0,1]$:

$$\begin{cases} u = F_D(d) \\ v = F_S(s) \end{cases} \tag{7}$$

In order to obtain (u_i, v_i) for each couple (d_i, s_i) in the data set, theoretical or empirical cumulative distribution functions of D and S can be used (Vandenberghe et al., 2011). The later approach is preferred in this study as it allows for the use of an empirical distribution function for the random variables D and S. The values for U and V are calculated as follows:

$$\begin{cases} u_i = R_i/(n+1) \\ v_i = T_i/(n+1) \end{cases} \tag{8}$$

with n the number of drought events; and R_i and T_i are the ranks of d_i and s_i among drought events.

HESSD
10, 7469–7516, 2013

To model the dependence structure, we restricted the copulas to the most common one-parameter families, such as Clayton, Gumbel, Frank, AMH, A12 and A14 (Table 2).

Several techniques can be used for estimating the copula parameter θ, such as semi-parametrical rank-based methods, parametrical methods, and kernel techniques (Genest and Favre, 2007; Salvadori and De Michele, 2007). Here, we use a rank-based method based on Kendall's tau τ_K, in which the copula parameter θ is calculated as a function of Kendall's tau. We opted for this method as it has proven to be robust in describing variable correlation and outlier effects (Li et al., 2012). Furthermore, this methodology is easy to implement. Table 2 presents the copula functions, domain of the dependence parameter θ, relationship function between θ and Kendall's tau and the range of Kendall's tau for the different copulas tested.

4.2 Copula-based drought frequency analysis

Bivariate hydrologic events can be categorized as joint and conditional events (Shiau, 2003). The joint drought events can be defined in two cases: AND $\{D > d \text{ and } S > s\}$ and OR $\{D > d \text{ or } S > s\}$ (Vandenberghe et al., 2011). The return period definitions T_{AND} and T_{OR} respectively for AND and OR events are defined in term of u and v as:

$$T_{AND} = \frac{E(L)}{1 - F_D(d) - F_S(s) + F_{DS}(d,s)} = \frac{E(L)}{1 - u - v + C(u,v)} \tag{9}$$

$$T_{OR} = \frac{E(L)}{1 - F_{DS}(d,s)} = \frac{E(L)}{1 - C(u,v)} \tag{10}$$

where $E(L)$ is the expected drought inter-arrival time, which can be estimated from observed droughts which have been identified based on the EDI.

One may also be interested in two types of conditional drought situations which are referred as $\mathrm{COND}_1\ \{S|D > d\}$ and $\mathrm{COND}_2\ \{S > s|D \le d\}$ (Salvadori et al., 2007; Vandenberghe et al., 2011). The return periods T_{COND1} and T_{COND2} for respectively COND_1

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

10, 7469–7516, 2013

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

and COND_2 are defined as follows:

$$T_{\text{COND1}} = \frac{E(L)}{1 - F_D(d)} \times \frac{1}{1 - F_D(d) - F_S(s) + F_{DS}(d,s)} = \frac{E(L)}{1-u} \times \frac{1}{1 - u - v + C(u,v)} \quad (11)$$

$$T_{\text{COND2}} = \frac{E(L)}{1 - \frac{F_{DS}(d,s)}{F_D(d)}} = \frac{E(L)}{1 - \frac{C(u,v)}{u}} \quad (12)$$

Mathematical details of these return period calculation functions are provided by Salvadori (2004), Salvadori and De Michele (2004, 2007), Salvadori et al. (2007), and Vandenberghe et al. (2011).

5 Results and Discussion

5.1 General evaluation of BL models

To assess the general performance of the Bartlett–Lewis models, the models' ability to reproduce general historical rainfall characteristics is first considered. Table 3 compares general historical rainfall characteristics to simulated rainfall, at different levels of aggregation. For the purpose of comparison, the percentual deviation of the simulated values from the observations is also listed in Table 3. Several differences between the models can be discovered from the table. Generally, the mean is reproduced quite well by all models. However, the OBL and TBLG models show a slightly higher deviation from the observation than the other models. For the variance, the TBL and TBLG models are less accurate than the non-truncated models, especially at the 10 min level of aggregation. At the hourly and daily level, this difference is less apparent. The autocovariance at the subhourly level, however, is reproduced very well by the truncated models (TBL and TBLG), while at the hourly and subhourly level this is less the case. The reproduction of zero depths is comparable for all models. However, it can be seen that the OBL model fails to preserve this property at the daily level. The third moment, finally, is seriously underestimated at the hourly and subhourly level by the non-truncated models (OBL, MBL and MBLG), while the truncated models are able to limit this discrepancy, the latter having a definite positive effect on their extreme value behaviour (Onof et al., 2013). The errors are also comparable to those presented in other papers (e.g. Onof and Wheater, 1993; Verhoest et al., 1997, 2010).

The modest analysis above shows that, in general, certain differences exist between the models. However, it is not possible to conclude that one model performs better than the other, based on these general characteristics. It can be concluded that each of the parameterized BL models are well calibrated and the performances of the considered variants of the BL model are comparable.

5.2 Basic statistic of droughts by observed and modelled rainfall

5.2.1 Analysis of EDI values

In order to unveil any patterns in drought occurrence behaviour in the observed and simulated data, EDI values are plotted in function of the year (x-axis) and the day of the year (y-axis) (Fig. 1). This way, a quantitative assessment of any seasonal patterns as well as the inter-annual variability can be made. From the plots, it can be seen that droughts seem to occur more often and more severely in the Uccle observations than in the BL simulations, except for the MBL simulation. This figure reveals that drier (low EDI) and wetter (high EDI) periods seem to span over multiple years. The evidence of some very dry years and certain remarkable wet years can be found in the EDI figures of Uccle, MBL and TBL. No clear seasonal trends can be witnessed for both observed and simulated data. It can also be seen that the OBL, MBLG and TBLG simulations produce less extreme events than the other models; therefore, based solely on these plots, it seems that these three models did not represent well long-term dry or wet conditions in a realistic manner. Figure 2 shows the comparison between frequency distributions of EDI values of observed and simulated data. It is clear from the figure that all BL models simulate less lower EDI values and thus tend to reproduce more

HESSD
10, 7469–7516, 2013

higher EDI values than the observation; in other words, all the BL models seem to produce more wetter daily conditions than the Uccle observations.

To further investigate, Table 4 shows the intra- and inter-annual variance of the EDI for all data sets. This table confirms that except for MBL and TBL, all models produced more or less the same intra-annual variability of EDI in comparison with those calculated from the Uccle observations. The intra-annual variance for MBL and TBL is slightly higher than those for Uccle which means that, on average, the EDI values obtained from MBL- and TBL-modelled time series fluctuate more throughout the year than those derived from observed rainfall series. In case of the inter-annual variability, the values for TBL are more or less similar to those for Uccle, while there is an overestimation for MBL and an underestimation for the other BL models.

From a practical point of view, we only consider droughts with a duration of at least 7 days in the frequency analyses. Droughts with a duration smaller than 7 days are not easily detected in reality and may not cause any serious effects. The threshold duration of 7 days, which is much smaller than the minimum drought duration identified by other drought indices (usually a month), still allows for investigating the temporal behaviour of BL models with respect to predicting minor drought events. This choice also helps to remove a problem of "ties" in the data. This problem refers to the presence of events with identical values for both D and S which may cause difficulties in distribution fitting and copula-based statistical analysis. For more information about this problem, we refer to Salvadori and De Michele (2006, 2007). Table 5 gives some basic statistics for all observed and simulated drought events from which it can be seen that all models overestimate the numbers of drought events while generally they underestimate the drought duration.

5.2.2 Analysis of YAEDI_{365} values

In this analysis, we considered a year which has YAEDI_{365} less than −1.5 as a "seriously dry year" based on the classification in Table 1. Figure 3 presents YAEDI_{365} of all data during 104 yr; for observed data, the years are numbered from 1899 to 2002,

HESSD
10, 7469–7516, 2013
A copula-based assessment of Bartlett–Lewis type
M. T. Pham et al.

while they range between 2 and 105 for synthetic data. For Uccle, we may identify three seriously dry years, being 1921, 1949 and 1976; this agrees with the findings by De Jongh et al. (2006). From these data, one may infer that a seriously dry year occurs every 27 to 28 yr, however, this statistic should be treated with care given the limited length of the time series. All the models seem to underestimate dry conditions and fail to simulate the extreme events. There is not a single "seriously dry year" observed for any of the BL models. YAEDI_{365} values simulated by the OBL, MBL and MBLG models seemed to be smaller and less variable than those by the other models. The underestimation of BL models is also confirmed in Fig. 4 in which the empirical cumulative distribution functions for YAEDI_{365} are presented for all data sets.

5.3 Probability distributions for *D* and *S*

The marginal cumulative distribution functions of D and S need to be modelled separately as these are needed when conducting a copula-based frequency analysis in order to transform these values from R^2 to I^2 or vice versa. Different commonly used parametric models such as Generalized Pareto (GP), generalized extreme value (GEV), Exponential, Weibull and Gamma distribution functions, and a nonparametric Kernel model are considered in this fitting test. The reason for also conducting a non-parametric model fit is found in the fact that such model may avoid the typical problems of under- or overestimating extreme events when fitting a parametric model (Vandenberghe, 2012).

In order to assess the significance of the fit, the statistics of Anderson–Darling AD_n (Anderson and Darling, 1954) is calculated and used for identifying the appropriate distribution for D and S. Table 6 lists the values of the AD_n statistics for the five parametric and the one non-parametric distribution functions fitted to the duration D and the severity S of the Uccle observed, OBL-, MBL-, MBLG-, TBL-and TBLG-modelled droughts. Note that smaller AD_n values express a better distribution fit. Table 7 presents the p values for these AD_n tests. As can be seen from Table 6, the best fits for both D and S are provided by the GEV distribution. The second best fits for D and S are obtained

HESSD
10, 7469–7516, 2013

with the Kernel distribution. These results are different from some previous studies in which the distribution function for D is generally considered to be a geometric distribution (Kendall and Dracup, 1992; Mathier et al., 1992) or an exponential distribution (Shiau, 2006; Zelenhasi et al., 1987), and the distribution function for S is considered to be a gamma distribution (Mathier et al., 1992; Shiau and Shen, 2001; Shiau, 2006; Zelenhasi et al., 1987). p values in Table 7 indicate that only for the Kernel distribution, an appropriate fit for all data sets is obtained. In contrast, all the parametric distributions are clearly rejected. To avoid the problems of under- or overestimation of marginal distribution fitting for D and S, the distributions should be investigated by a graphical presentation. Figures 5 and 6 display the GEV and Kernel cumulative distribution functions of D and S for all data sets, respectively. As can be seen for all cases, the Kernel distribution (red line) is better in simulating the extreme values while the GEV (black line) overestimates the extremes.

Based on the above analysis, the Kernel distribution for both D and S is selected in further analysis. Figure 7 shows the comparison of the Kernel cumulative distribution functions of D and S for all data sets; it is clear that the cumulative probability of D or S calculated from the observed Uccle data (black line) is always smaller than what is found for the different BL models, which means that all BL models generally underestimate D and S.

5.4 Identification of the appropriate copula

The copula parameter estimation method for D and S makes use of the estimation of Kendall's tau (τ_K). As Kendall's tau values of all data sets (Table 8) are out of range for the AMH copula (Table 2), it is no longer considered in the study. For the copula selection, the root-mean-square error (RMSE) and two goodness-of-fit test statistics proposed by Genest et al. (2006), namely S_n and T_n, are calculated.

The RMSE is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C\left(u_i, v_i\right) - C_n\left(u_i, v_i\right)\right)^2} \tag{13}$$

in which n is the number of data points, C is the fitted copula based on parameters estimated via Kendall's tau, and C_n is the empirical copula. Table 9 presents the RMSE values obtained for different copulas tested for the observed and simulated data. As can be seen, similar results are obtained for all copulas, yet one could suggest the A12 and Frank copulas as best choice for all datasets if only the RMSE is used.

The two goodness-of-fit test statistics S_n and T_n are calculated based on Kendall's tau. The smaller their values the better the accuracy achieved. S_n and T_n are defined as follows:

$$S_n = \int_0^1 \left|K_n(w)\right|^2 k\left(w\right) \mathrm{d}w \tag{14}$$

$$T_n = \sup_{0 \le w \le 1} \left|K_n\left(w\right)\right| \tag{15}$$

where K_n is the Kendall process (Genest et al., 2006). Mathematical details of the calculation of S_n and T_n are provided by Genest et al. (2006). The p values associated with S_n and T_n are calculated by means of a bootstrap method (Genest et al., 2006).

The results of S_n, T_n and their respective p values are presented in Tables 10 and 11. In case of S_n, the p values indicate that only the Frank copula is found to be an appropriate copula at the 5 % significance level for all data sets. The p values for T_n from Table 11 also result in the same conclusion. Based on these statistics, the Frank copula is considered for the frequency analysis for all data sets. Comparison between empirical copulas (red dotted line) and fitted Frank copulas via Kendall's tau (full line) for all data sets is shown in Fig. 8. In this figure, the plotted variables U and V are the

HESSD
10, 7469–7516, 2013

normalized ranks of the variables D and S, respectively. It is clear that the copula fits for data of MBL, MBLG, TBL and TBLG models are worse than for the others; however the fits are still considered to be acceptable.

5.5 Copula-based frequency analysis

In this section, four types of return period will be investigated; we will focus on drought events with a return period of 5 yr (Fig. 9) and 10 yr (Fig. 10). In case of AND and OR return periods, for both 5 and 10 yr drought return periods, with an event with a given return period, it is clear that all models severely underestimate the magnitude of drought properties, or alternatively, an observed drought event having a return period of 5 yr or 10 yr will have a lower frequency of occurrence if it were modelled by the five BL models. The underestimations seem to become more pronounced for more extreme events. TBL model produces the closest 5 yr drought statistics compared to those by the Uccle observations for both AND and OR types, but it has very poor results in case of a 10 yr return period; for a 10 yr drought, MBLG and TBLG models, respectively, give the best result in the AND and OR cases. For COND_1 type, the TBL model simulates drought events that are statistically comparable to those in the observed Uccle data; the other models slightly underestimate the magnitude of drought events. For 10 yr drought, slight underestimations are witnessed for all models, however drought statistics from TBL simulation still remain closest those by the Uccle data, followed by MBLG and MBL simulations. The situation is different for the COND_2 type in both 5 and 10 yr drought return periods. All models seem able to simulate droughts with a duration smaller than 70 days; however, overall MBL model shows the best performance. It should be noted from the figure that lines presenting results for all Bartlett–Lewis models are shorter than those of observed data in case of OR and COND_2, which should be attributed to the lack of severe drought events (see Table 5 and Fig. 7); the interpolated results by copula are therefore limited to a certain scale.

Overall, it is difficult to conclude which model has the best performance. The TBL model seems to be the best in simulating droughts with a high frequency in AND, OR

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD
10, 7469–7516, 2013

and $COND_1$ types. The MBL model shows the best results in $COND_2$ type for both 5 and 10 yr return periods; the OBL model fails in almost cases. The shortcomings of all rainfall models in simulating extreme drought events can be partly explained by their overestimation of the cumulative value of marginal distribution functions for D and S. We can thus conclude that the BL models seem to simulate longer and more severe drought events at a too low pace. This can be attributed to the fact that the model itself does not foresee any non-stationarity and maintains the same parameters throughout the simulation period. The shortcomings also can be explained by the problems of inducing over-clustering by model structure (Vandenberghe et al., 2011) which may result in generating more and shorter dry periods.

6 Conclusion and recommendation

In this study, some basic statistics are compared and a copula-based bivariate frequency analysis is performed in order to assess whether rainfall series simulated by Bartlett–Lewis (BL) models are able to preserve drought statistics. A record of the 105 yr period of 1898–2002 of 10 min rainfall for Uccle in Belgium is used as observed data. For drought identification, the EDI is chosen and drought events were defined as an extremely dry to moderately dry period during which the EDI is continuously less than -0.7. Each drought event is characterized by two variables, i.e. drought duration (D) and severity (S). Through quantitatively analysing daily EDI time series, it was clear that droughts seem to occur more often and more severely in the observations and for the MBL simulation than for the other BL models. However, no clear seasonal trends can be witnessed for both observed and simulated data.

It was demonstrated that D and S could be modelled by a GEV distribution in contrast to what is generally considered that the distribution for D should be a geometric or an exponential distribution and for S should be a gamma distribution (Mathier et al., 1992; Shiau and Shen, 2001; Shiau, 2006; Zelenhasi et al., 1987). However, in this research context the non-parametric Kernel distribution was selected as it allowed to

HESSD
10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

better represent the upper tail of the distribution. The analysis of marginal distribution functions of D and S showed that all models overestimate the probability of extreme events. The application of the yearly accumulated negative EDI, $YAEDI_{365}$, also allows for identifying dry conditions in the time series for all data; for Uccle, three seriously dry years are witnessed within 105 yr time series. All BL models tested seem to underestimate these dry conditions and fail in simulating similar extreme events. $YAEDI_{365}$ values simulated by the MBL, MBLG and OBL models seemed to be smaller than those by the other models.

A frequency analysis was performed using bivariate copula-based return periods of droughts, expressed in term of D and S. The Frank copula was selected based on results of RMSE, S_n and T_n. Four types of copula-based drought return periods are conducted for all data sets. The comparison of four types of drought return periods indicated that all BL models seem to underestimate the drought severity compared with those observed in Uccle and it is therefore difficult to conclude which model best preserves drought statistics, although OBL shows disappointing results in most cases. TBL model produces 5 yr drought statistics that are closest to those of the Uccle observation in case of AND, OR and $COND_1$ types. MBL model performs very good in case of $COND_2$ type for both 5 and 10 yr droughts. OBL model shows disappointing results in almost cases. The shortcomings of all BL models in simulating extreme drought events can be partly explained by the fact that the BL models simulate longer and more severe drought events with a too low frequency which can be attributed to several reasons. First, as the models use the same parameter sets throughout the simulation period, these models thus cannot foresee any non-stationarity. This could explain why the variability in e.g. $YAEDI_{365}$ values calculated from BL simulations is too small compared to those by the observed time series. Furthermore, the temporal variability assured through the stochastic process within the BL models is insufficient to allow for generating the extreme drought events. The simulating process in all BL models also does not assume any temporal autocorrelation between successive storms which may be needed to model longer drought periods. Finally, the model problem of over-clustering

HESSD
10, 7469–7516, 2013

may have may have greater impacts during severe drought periods than during the remaining simulation period. One way to solve this problem is to investigate whether temporally changing parameter sets would allow to better preserve the droughts while still ensuring the other characteristics of rainfall (such as moments, extreme rainfall or zero depth probabilities). This could be obtained by including a dependence between models parameters or cluster variables through e.g. introducing copulas in the BL structure.

Acknowledgements. The authors gratefully acknowledge the Vietnamese Government Scholarship (VGS) and the project G.0013.11N of the Research Foundation Flanders (FWO) for their partial financial support for this work.

References

Anderson, T. W. and Darling, D. A.: A test of goodness of fit, J. Am. Stat. Assoc., 49, 765–769, 1954.

Bhalme, H. N. and Mooley, D. A.: Large-scale droughts/floods and monsoon circulation, Mon. Weather Rev., 108, 1197–1211, 1980.

Bogner, K., Konency, F., Nachtnebel, H., and Onof, C.: Simuation of rainfall with different temporal resolutions by stochastic pulse models, in: Scaling Problems in Hydrology, edited by: Gutknecht, D., Hantel, M., and Nachtnebel, H. P., National Committee of the International Hydrological Programme, Austrian Academy of Sciences, Vienna, 109–122, 2001.

Burton, A., Kilsby, C. G., Fowler, H. J., Cowpertwait, P. S. P., and O'Connell, P. E.: RainSim: a spatial–temporal stochastic rainfall modelling system, Environ. Modell. Softw., 23, 1356–1369, 2008.

Byun, H.-R. and Wilhite, D. A.: Objective quantification of drought severity and duration, J. Climate, 12, 2747–2756, 1999.

Byun, H.-R., Sun-Ju, L., Saeid, M., Ki-Seon, C., Sang-Min, L., and Do-Woo, K.: Study on the periodicities of droughts in Korea, Asia-Pac. J. Atmos. Sci., 44, 417–441, 2008.

Cameron, D., Beven, K., and Tawn, J.: An evaluation of three stochastic rainfall models, J. Hydrol., 228, 130–149, 2000.

Cameron, D., Beven, K., and Tawn, J.: Modelling extreme rainfalls using a modified random pulse Bartlett–Lewis stochastic rainfall model (with uncertainty), Adv. Water Resour., 24, 203–211, 2001.

Cowpertwait, P., Isham, V., and Onof, C.: Point process models of rainfall: developments for fine-scale structure, P. Roy. Soc. A-Math. Phy., 463, 2569–2587, 2007.

De Jongh, I. L. M., Verhoest, N. E. C., and De Troch, F. P.: Analysis of a 105-year time series of precipitation observed at Uccle, Belgium, Int. J. Climatol., 26, 2023–2039, 2006.

Démarée, G. R.: Le pluviographe centenaire du plateau d'Uccle: son histoire, ses données et ses applications (The centennial recording raingauge of the Uccle Plateau: its history, its data and its applications), La Houille Blanche, 4, 95–102, 2003.

Dogan, S., Berktay, A., and Singh, V. P.: Comparison of multi-monthly rainfall-based drought severity indices, with application to semi-arid Konya closed basin, Turkey, J. Hydrol., 470–471, 255–268, 2012.

Duan, Q., Sorooshian, S., and Gupta, V. K.: Optimal use of the SCE-UA global optimization method for calibrating watershed models, J. Hydrol., 158, 265–284, 1994.

Engida, A. N. and Esteves, M.: Characterization and disaggregation of daily rainfall in the Upper Blue Nile Basin in Ethiopia, J. Hydrol., 399, 226–234, 2011.

Genest, C. and Favre, A.: Everything you always wanted to know about copula modeling but were afraid to ask, J. Hydrol. Eng., 12, 347–368, 2007.

Genest, C. and Rivest, L.-P.: On the multivariate probability integral transformation, Stat. Probabil. Lett., 53, 391–399, 2001.

Genest, C., Quessy, J.-F., and Rémillard, B.: Goodness-of-fit procedures for copula models based on the probability integral transformation, Scand. J. Stat., 33, 337–366, 2006.

Glasbey, C. A., Cooper, G., and McGechan, M. B.: Disaggregation of daily rainfall by conditional simulation from a point-process model, J. Hydrol., 165, 1–9, 1995.

Gommes, R. A. and Petrassi, F.: Rainfall variability and drought in sub-Saharan Africa since 1960, Technical Report, Food and Agriculture Organization of the United Nations, Research and Technology Development Division, Agrometeorology Group, Rome, Italy, 1994.

González, J. and Valdés, J. B.: New drought frequency index: definition and comparative performance analysis, Water Resour. Res., 42, W11421, doi:10.1029/2005WR004308, 2006.

Gyasi-Agyei, Y.: Identification of regional parameters of a stochastic model for rainfall disaggregation, J. Hydrol., 223, 148–163, 1999.

HESSD
10, 7469–7516, 2013

Gyasi-Agyei, Y. and Willgoose, G. R.: Generalisation of a hybrid model for point rainfall, J. Hydrol., 219, 218–224, 1999.
Jesus, J. and Chandler, R. E.: Estimating functions and the generalized method of moments, Interface Focus, 1, 871–885, 2011.
Kaczmarska, J.: Further development of Bartlett–Lewis models for fine-resolution rainfall, Technical Report, Department of Statistical Science, University College London, 2011.
Kendall, D. and Dracup, J.: On the generation of drought events using an alternating renewal-reward model, Stoch. Hydrol. Hydraul., 6, 55–68, 1992.
Khaliq, M. N. and Cunnane, C.: Modelling point rainfall occurrences with the modified Bartlett–Lewis rectangular pulses model, J. Hydrol., 180, 109–138, 1996.
Kim, D.-W., Byun, H.-R., and Choi, K.-S.: Evaluation, modification, and application of the Effective Drought Index to 200-year drought climatology of Seoul, Korea, J. Hydrol., 378, 1–12, 2009.
Kim, T.-W., Valdes, J. B., and Yoo, C.: Nonparametric approach for estimating return periods of droughts in arid regions, J. Hydrol. Eng., 8, 237–246, 2003.
Kim, T.-W., Valdés, J. B., and Aparicio, J.: Spatial characterization of droughts in the Conchos river basin based on bivariate frequency analysis, Water Int., 31, 50–58, 2006a.
Kim, T.-W., Valdés, J. B., and Yoo, C.: Nonparametric approach for bivariate drought characterization using palmer drought index, J. Hydrol. Eng., 11, 134–143, 2006b.
Li, N., Liu, X., Xie, W., Wu, J., and Zhang, P.: The return period analysis of natural disasters with statistical modeling of bivariate joint probability distribution, Computat. Studies, 33, 134–145, 2012.
Marani, M. and Zanetti, S.: Downscaling rainfall temporal variability, Water Resour. Res., 43, W09415, doi:10.1029/2006WR005505, 2007.
Mathier, L., Perreault, L., Bobée, B., and Ashkar, F.: The use of geometric and gamma-related distributions for frequency analysis of water deficit, Stoch. Hydrol. Hydraul., 6, 239–254, 1992.
McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, Proceedings of the 8th Conference on Applied Climatology, American Meteorological Society, 179–184, 1993.
Morid, S., Smakhtin, V., and Moghaddasi, M.: Comparison of seven meteorological indices for drought monitoring in Iran, Int. J. Climatol., 26, 971–985, 2006.

Morid, S., Smakhtin, V., and Bagherzadeh, K.: Drought forecasting using artificial neural networks and time series of drought indices, Int. J. Climatol., 27, 2103–2111, 2007.

Onof, C. and Wheater, H. S.: Modelling of British rainfall using a random parameter Bartlett–Lewis rectangular pulse model, J. Hydrol., 149, 67–95, 1993.

Onof, C. and Wheater, H. S.: Improvements to the modelling of British rainfall using a modified random parameter Bartlett–Lewis rectangular pulse model, J. Hydrol., 157, 177–195, 1994.

Onof, C., Meca-Figueras, T., Kaczmarska, J., Chandler, R., and Hege, L.: Modelling rainfall with a Bartlett–Lewis process: third-order moments, proportion dry, and a truncated random parameter version, Technical Report, Department of Civil and Environmental Engineering, Imperial College London, 2012.

Pandey, R. P., Dash, B. B., Mishra, S. K., and Singh, R.: Study of indices for drought characterization in KBK districts in Orissa (India), Hydrol. Process., 22, 1895–1907, 2008.

Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., and Jeremiah, E.: A comparison of alternatives for daily to sub-daily rainfall disaggregation, J. Hydrol., 470, 138–157, 2012.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V.: Some models for rainfall based on stochastic point processes, P. Roy. Soc. Lond. A Mat., 410, 269–288, 1987a.

Rodriguez-Iturbe, I., De Power, B. F., and Valdes, J. B.: Rectangular pulses point process models for rainfall: analysis of empirical data, J. Geophys. Res., 92, 9645–9656, 1987b.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V.: A point process model for rainfall: further developments, P. Roy. Soc. Lond. A Mat., 417, 283–298, 1988.

Rooy, M. P. V.: A rainfall anomaly index (RAI) independent of time and space, Notos, 14, 43–48, 1965.

Salvadori, G.: Bivariate return periods via 2-copulas, Statistical Methodology, 1, 129–144, 2004.

Salvadori, G. and De Michele, C.: Frequency analysis via copulas: theoretical aspects and applications to hydrological events, Water Resour. Res., 40, W12511, doi:10.1029/2004WR003133, 2004.

Salvadori, G. and De Michele, C.: Statistical characterization of temporal structure of storms, Adv. Water Resour., 29, 827–842, 2006.

Salvadori, G. and De Michele, C.: On the use of copulas in hydrology: theory and practice, J. Hydrol. Eng., 12, 369–380, 2007.

HESSD
10, 7469–7516, 2013

Salvadori, G. and De Michele, C.: Multivariate multiparameter extreme value models and return periods: a copula approach, Water Resour. Res., 46, W10501, doi:10.1029/2009WR009040, 2010.

Salvadori, G., Michele, C. D., Kottegoda, N., and Rosso, R.: Extremes in Nature: an Approach Using Copulas, Springer, New York, 2007.

Shiau, J. T.: Return period of bivariate distributed extreme hydrological events, Stoch. Env. Res. Risk A., 17, 42–57, 2003.

Shiau, J. T.: Fitting drought duration and severity with two-dimensional copulas, Int. Ser. Prog. Wat. Res., 20, 795–815, 2006.

Shiau, J. T. and Modarres, R.: Copula-based drought severity-duration-frequency analysis in Iran, Meteorol. Appl., 16, 481–489, 2009.

Shiau, J. T. and Shen, H. W.: Recurrence analysis of hydrologic droughts of differing severity, J. Water Res. Pl.-ASCE, 127, 30–40, 2001.

Sklar, A.: Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris, 8, 229–231, 1959.

Smithers, J. C., Pegram, G. G. S., and Schulze, R. E.: Design rainfall estimation in South Africa using Bartlett–Lewis rectangular pulse rainfall models, J. Hydrol., 258, 83–99, 2002.

Vandenberghe, S.: Copula-based Models for Generating Design Rainfall, PhD dissertation, Ghent University, Faculty of Bioscience Engineering, Ghent, Belgium, 2012.

Vandenberghe, S., Verhoest, N. E. C., Onof, C., and De Baets, B.: A comparative copula-based bivariate frequency analysis of observed and simulated storm events: a case study on Bartlett–Lewis modeled rainfall, Water Resour. Res., 47, W07529, doi:10.1029/2009WR008388, 2011.

Vanhaute, W. J., Vandenberghe, S., Scheerlinck, K., De Baets, B., and Verhoest, N. E. C.: Calibration of the modified Bartlett-Lewis model using global optimization techniques and alternative objective functions, Hydrol. Earth Syst. Sci., 16, 873–891, doi:10.5194/hess-16-873-2012, 2012.

Velghe, T., Troch, P. A., De Troch, F. P., and Van de Velde, J.: Evaluation of cluster-based rectangular pulses point process models for rainfall, Water Resour. Res., 30, 2847–2857, 1994.

Verhoest, N., Troch, P. A., and De Troch, F. P.: On the applicability of Bartlett–Lewis rectangular pulses models in the modeling of design storms at a point, J. Hydrol., 202, 108–120, 1997.

HESSD
10, 7469–7516, 2013

Verhoest, N. E. C., Vandenberghe, S., Cabus, P., Onof, C., Meca-Figueras, T., and Jameleddine, S.: Are stochastic point rainfall models able to preserve extreme flood statistics?, Hydrol. Process., 24, 3439–3445, 2010.

Wheater, H. S., Chandler, R. E., Onof, C. J., Isham, V. S., Bellone, E., Yang, C., Lekkas, D., Lourmas, G., and Segond, M. L.: Spatial-temporal rainfall modelling for flood risk estimation, Stoch. Env. Res. Risk A., 19, 403–416, 2005.

Wilby, R. L. and Wigley, T. M. L.: Precipitation predictors for downscaling: observed and general circulation model relationships, Int. J. Climatol., 20, 641–661, 2000.

Wilhite, D., Svoboda, M., and Hayes, M.: Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness, Int. Ser. Prog. Wat. Res., 21, 763–774, 2007.

Wong, G., Lambert, M. F., Leonard, M., and Metcalfe, A. V.: Drought analysis using trivariate copulas conditional on climatic states, J. Hydrol. Eng., 15, 129–141, 2010.

Yu-Won, K. I. M. and Hi-Ryong, B.: On the cause of summer droughts in Korea and their return to normal, Asia-Pac. J. Atmos. Sci., 42, 237–251, 2006.

Zelenhasi, E. and Salvai, A.: A method of streamflow drought analysis, Water Resour. Res., 23, 156–168, 1987.

HESSD
10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 1. Drought severity classification by the EDI index (Morid et al., 2006).

Extremely wet	≥ 2.50
Very wet	1.50 to 2.49
Moderately wet	0.70 to 1.49
Normal	−0.69 to 0.69
Moderately dry	−0.70 to −1.49
Severely dry	−1.50 to −2.49
Extremely dry	≤ -2.50

HESSD
10, 7469–7516, 2013

Table 2. Selected copulas and their domain of dependence parameter θ and Kendall's tau.

Copula	$C_\theta\ (u,v)$	Parameter θ	$\tau_K = g(\theta)$	Range of τ_K
Clayton	$\max\left(\left[u^{-\theta}+v^{-\theta}-1\right]^{-1/\theta},0\right)$	$[-1,\infty[\setminus\{0\}$	$1-\frac{2}{2+\theta}$	$(0,1]$
Gumbel–Hougaard	$\exp\left(-\left[(-\ln u)^{\theta}+(-\ln v)^{\theta}\right]^{1/\theta}\right)$	$[-1,\infty[$	$1-\theta^{-1}$	$[0,1]$
Frank	$-\frac{1}{\theta}\ln\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right)$	$]-\infty,\infty[\setminus\{0\}$	$1-\frac{4}{\theta}\left(1-\frac{1}{\theta}\int_0^{\theta}\frac{t}{e^t-1}dt\right)$	$[-1,1]\setminus\{0\}$
AMH	$\frac{uv}{1-\theta(1-u)(1-v)}$	$[-1,1[$	$1-\frac{2}{3}\frac{\theta^2\ln(1-\theta)-2\theta\ln(1-\theta)+\theta+\ln(1-\theta)}{\theta^2}$	$\left[-0.181726,\frac{1}{3}\right]$
A12	$\left(1+\left[\left(u^{-1}-1\right)^{\theta}+\left(v^{-1}-1\right)^{\theta}\right]^{1/\theta}\right)^{-1}$	$[-1,\infty[$	$1-\frac{2}{3\theta}$	$\left[\frac{1}{3},1\right]$
A14	$\left(1+\left[\left(u^{-1/\theta}-1\right)^{\theta}+\left(v^{-1/\theta}-1\right)^{\theta}\right]^{1/\theta}\right)^{-\theta}$	$[-1,\infty[$	$1-\frac{2}{1+2\theta}$	$\left[\frac{1}{3},1\right]$

HESSD
10, 7469–7516, 2013

Table 3. Comparison of general historical rainfall characteristics with simulation results. Values between brackets are percentual deviations of the simulated characteristic with respect to the observation.

	Mean (mm)		Variance (mm^2)		Autocovariance (mm^2)		ZDP (–)		3rd moment (mm^3)	
					level of aggregation: 10 min					
Observed	0.015		0.013		0.007		0.940		0.046	
OBL	0.016	(6.1 %)	0.012	(−9.3 %)	0.009	(21.1 %)	0.938	(−0.2 %)	0.016	(−66.3 %)
MBL	0.015	(−1.1 %)	0.012	(−4.2 %)	0.007	(2.4 %)	0.939	(−0.1 %)	0.023	(−50.5 %)
MBLG	0.015	(−1.1 %)	0.013	(−0.8 %)	0.008	(20.2 %)	0.942	(0.2 %)	0.024	(−47.5 %)
TBL	0.015	(−0.6 %)	0.017	(29.2 %)	0.007	(1.2 %)	0.955	(1.6 %)	0.047	(2.7 %)
TBLG	0.017	(8.3 %)	0.018	(43.7 %)	0.007	(1.3 %)	0.939	(−0.1 %)	0.056	(21.0 %)
					level of aggregation: 1 h					
Observed	0.092		0.222		0.092		0.874		1.584	
OBL	0.097	(6.1 %)	0.243	(9.5 %)	0.088	(−4.8 %)	0.878	(0.4 %)	1.213	(−23.4 %)
MBL	0.091	(−1.1 %)	0.216	(−2.7 %)	0.082	(−11.5 %)	0.870	(−0.5 %)	1.282	(−19.0 %)
MBLG	0.091	(−1.1 %)	0.239	(7.5 %)	0.084	(−9.2 %)	0.871	(−0.4 %)	1.679	(6.0 %)
TBL	0.091	(−0.6 %)	0.237	(6.8 %)	0.093	(0.4 %)	0.918	(5.0 %)	1.421	(−10.3 %)
TBLG	0.100	(8.3 %)	0.234	(5.2 %)	0.075	(−18.3 %)	0.878	(0.4 %)	1.518	(−4.1 %)
					level of aggregation: 24 h					
Observed	2.206		18.819		3.885		0.450		304.396	
OBL	2.340	(6.1 %)	18.107	(−3.8 %)	2.990	(−23.0 %)	0.475	(5.6 %)	233.084	(−23.4 %)
MBL	2.181	(−1.1 %)	18.910	(0.5 %)	4.261	(9.7 %)	0.441	(−1.9 %)	419.491	(37.8 %)
MBLG	2.183	(−1.1 %)	19.069	(1.3 %)	4.126	(6.2 %)	0.446	(−0.7 %)	437.200	(43.6 %)
TBL	2.194	(−0.6 %)	19.182	(1.9 %)	1.997	(−48.6 %)	0.446	(−0.8 %)	359.080	(18.0 %)
TBLG	2.389	(8.3 %)	19.876	(5.6 %)	2.950	(−24.1 %)	0.458	(1.9 %)	311.815	(2.4 %)

HESSD
10, 7469–7516, 2013

A copyula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 4. Intra- and inter-annual variance of EDI for Uccle and BL models.

Data	Intra-annual variance	Inter-annual variance
Uccle	0.58	1.06
OBL	0.58	0.87
MBL	0.69	1.13
MBLG	0.56	0.92
TBL	0.73	1.08
TBLG	0.62	0.98

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 5. Basic drought statistics of observed and simulated data.

Data	Numbers of droughts*	Total drought duration (days)	Average drought duration (days)	Longest drought event (days)	$E(L)$ (yr)
Uccle	211	10 168	48.19	498	0.49
OBL	243	6690	27.53	352	0.43
MBL	249	7232	29.04	196	0.42
MBLG	201	7392	36.78	321	0.52
TBL	317	9739	30.72	281	0.33
TBLG	266	8288	31.16	323	0.39

* Drought with duration of at least 7 days.

HESSD
10, 7469–7516, 2013

Table 6. Values of the AD_n statistic test for some distribution functions fitted to drought duration D and drought severity S.

Variable		Duration (D)						Severity (S)					
Distribution		GP	GEV	WBL	EXP	Gamma	Kernel	GP	GEV	WBL	EXP	Gamma	Kernel
AD_n	Uccle	7.556	1.541	Inf	15.368	12.845	4.590	5.958	0.646	Inf	33.370	15.456	3.336
	OBL	12.060	4.559	Inf	11.351	11.114	8.155	9.146	2.183	Inf	14.300	13.488	5.822
	MBL	11.858	1.336	119.733	11.537	11.731	8.884	9.884	0.776	Inf	12.984	13.584	6.893
	MBLG	7.669	1.421	108.128	8.462	9.073	4.858	6.048	0.836	Inf	12.892	10.791	3.574
	TBL	14.036	1.594	150.663	13.433	13.255	9.856	10.846	0.830	Inf	17.500	16.537	6.852
	TBLG	14.725	1.391	115.727	15.550	16.601	9.832	10.974	0.702	Inf	21.111	17.857	6.410

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 7. p values for the AD_n statistic test (p values larger than 0.05 indicating an appropriate fit of the distribution are displayed in bold).

Variable		Duration (D)						Severity (S)					
Distribution		GP	GEV	WBL	EXP	Gamma	Kernel	GP	GEV	WBL	EXP	Gamma	Kernel
p_{AD_n}	Uccle	0.000	0.018	0.000	0.000	0.000	**0.402**	0.000	**0.058**	0.000	0.000	0.000	**0.424**
	OBL	0.000	0.014	0.000	0.000	0.000	**0.322**	0.000	0.016	0.000	0.000	0.000	**0.256**
	MBL	0.000	0.004	0.000	0.000	0.000	**0.392**	0.000	0.016	0.000	0.000	0.000	**0.288**
	MBLG	0.000	0.000	0.000	0.000	0.000	**0.446**	0.000	0.014	0.000	0.000	0.000	**0.376**
	TBL	0.000	0.000	0.000	0.000	0.000	**0.440**	0.000	0.016	0.000	0.000	0.000	**0.318**
	TBLG	0.000	0.002	0.000	0.000	0.000	**0.358**	0.000	0.026	0.000	0.000	0.000	**0.342**

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 8. Kendall's tau τ_K for couple of D and S.

Dataset	τ_K
Uccle	0.960
OBL	0.930
MBL	0.940
MBLG	0.953
TBL	0.937
TBLG	0.940

HESSD
10, 7469–7516, 2013

HESSD

10, 7469–7516, 2013

A copy-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 9. RMSE of different fitted copulas.

	A12	A14	Frank	Clayton	Gumbel
Uccle	0.010	0.010	0.010	0.011	0.010
OBL	0.015	0.016	0.015	0.016	0.016
MBL	0.011	0.011	0.011	0.013	0.011
MBLG	0.011	0.011	0.011	0.011	0.011
TBL	0.010	0.010	0.010	0.011	0.010
TBLG	0.012	0.012	0.012	0.013	0.012

HESSD
10, 7469–7516, 2013

A copy-based assessment of Bartlett–Lewis type

M. T. Pham et al.

Table 10. S_n and p values for D and S for different copulas. The best fit is indicated in bold.

Data	A12		A14		Frank		Clayton		Gumbel	
	S_n	p	S_n	p	S_n	p	S_n	p	S_n	p
Uccle	**0.014**	0.024	0.016	0.006	0.018	0.340	0.016	0.005	0.016	0.003
OBL	0.031	0.002	0.035	0.001	**0.030**	0.563	0.041	0.000	0.035	0.000
MBL	**0.018**	0.026	**0.018**	0.032	0.019	0.993	0.033	0.000	0.019	0.035
MBLG	**0.016**	0.029	0.018	0.007	0.022	0.243	0.018	0.006	0.018	0.003
TBL	**0.021**	0.007	0.025	0.002	0.033	0.698	0.038	0.001	0.025	0.002
TBLG	**0.017**	0.054	0.021	0.011	0.025	0.859	0.038	0.000	0.021	0.004

Table 11. T_n and p values for D and S for different copulas. The best fit is indicated in bold.

Data	A12		A14		Frank		Clayton		Gumbel	
	T_n	p	T_n	p	T_n	p	T_n	p	T_n	p
Uccle	0.289	0.399	**0.284**	0.428	0.413	0.308	0.345	0.128	**0.284**	0.438
OBL	0.466	0.040	0.457	0.041	**0.406**	0.899	0.560	0.003	0.455	0.040
MBL	0.332	0.504	**0.284**	0.893	0.380	0.995	0.499	0.009	**0.284**	0.876
MBLG	**0.315**	0.358	0.319	0.337	0.380	0.606	0.379	0.094	0.323	0.259
TBL	0.410	0.102	**0.367**	0.297	0.562	0.396	0.566	0.002	0.376	0.218
TBLG	**0.361**	0.307	0.434	0.052	0.374	0.938	0.514	0.006	0.440	0.046

HESSD
10, 7469–7516, 2013

HESSD
10, 7469–7516, 2013

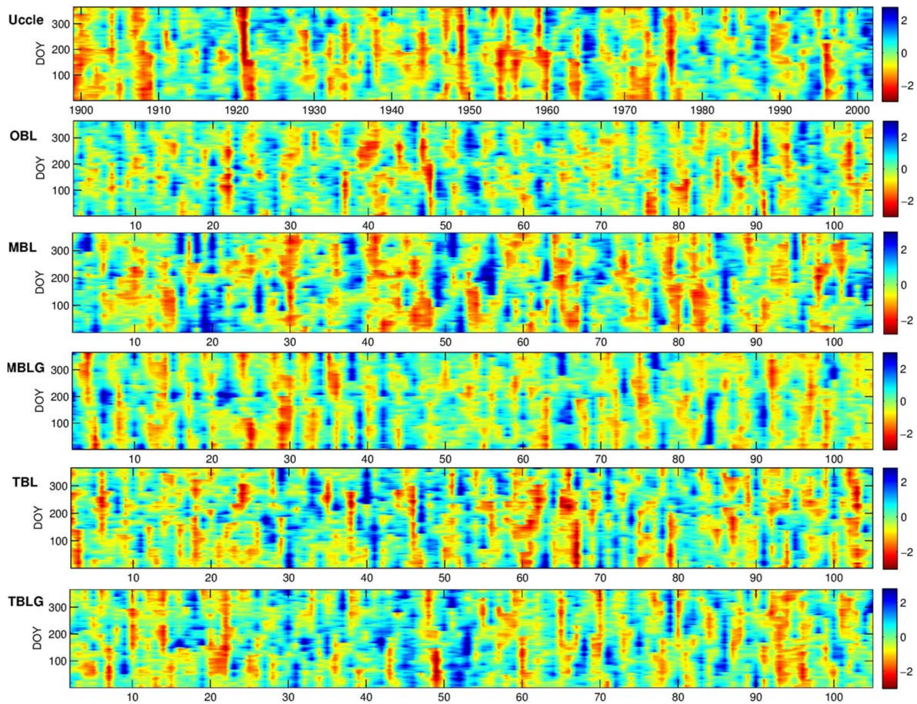

Fig. 1. Evolution of the EDI index of observed and simulated rainfall records for 104 yr. The y-axis corresponds to the day of the year (DOY), while the x-axis displays the year.

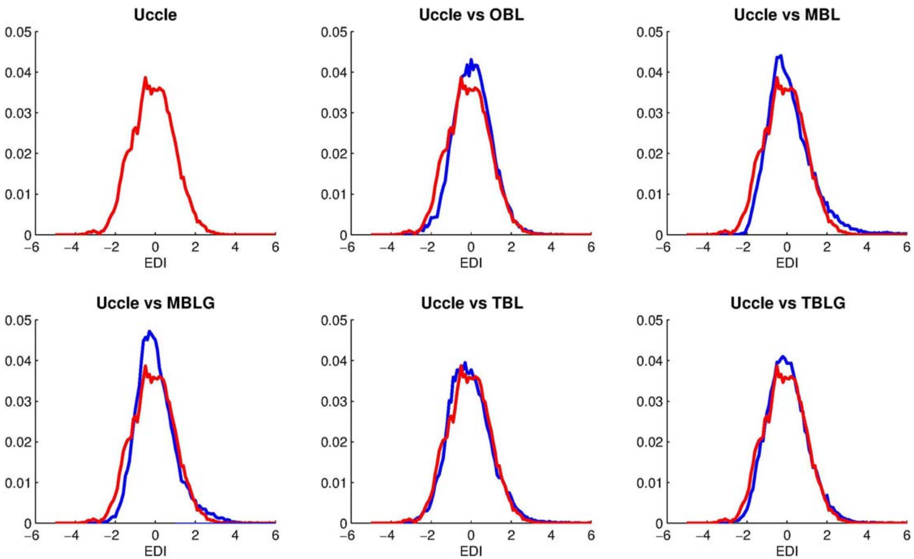

Fig. 2. Comparison between frequency distributions of EDI values of observed and simulated data: Uccle (red), BL models (blue).

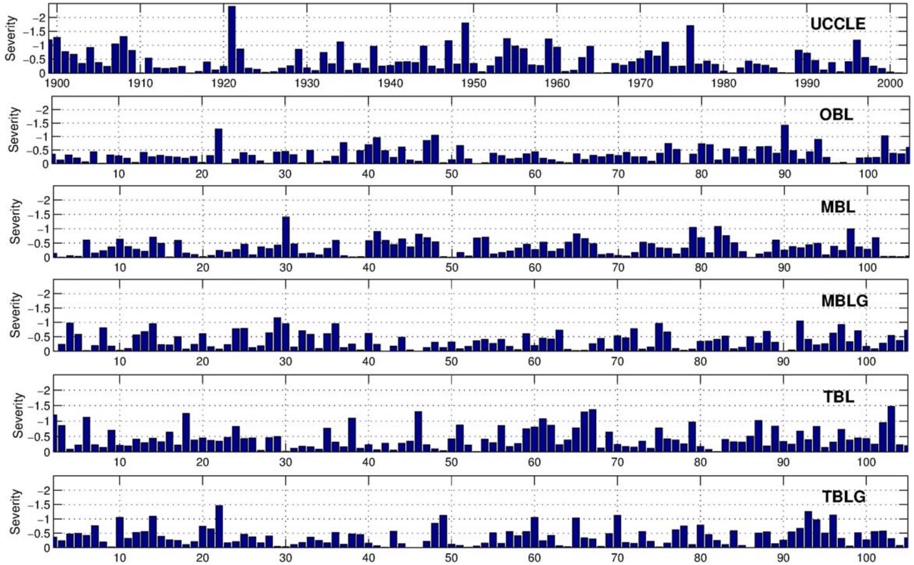

Fig. 3. Annual dryness of observed and modelled data represented by $YAEDI_{365}$.

HESSD
10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

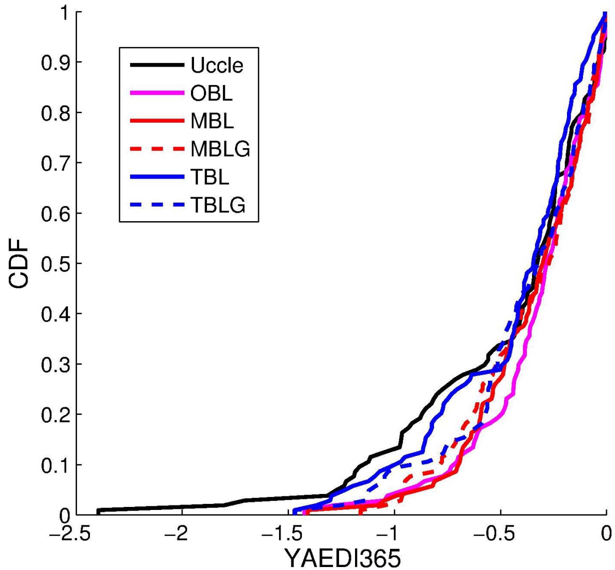

Fig. 4. Empirical cumulative distribution functions for $YAEDI_{365}$.

HESSD
10, 7469–7516, 2013

A copula-based assessment of Bartlett–Lewis type

M. T. Pham et al.

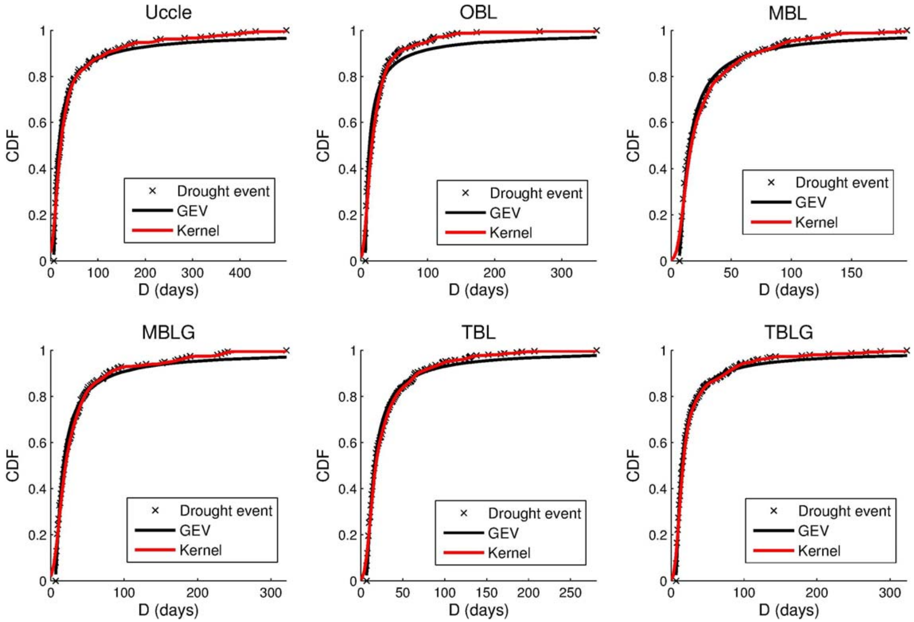

Fig. 5. GEV and Kernel cumulative distribution functions of drought duration D.

HESSD
10, 7469–7516, 2013

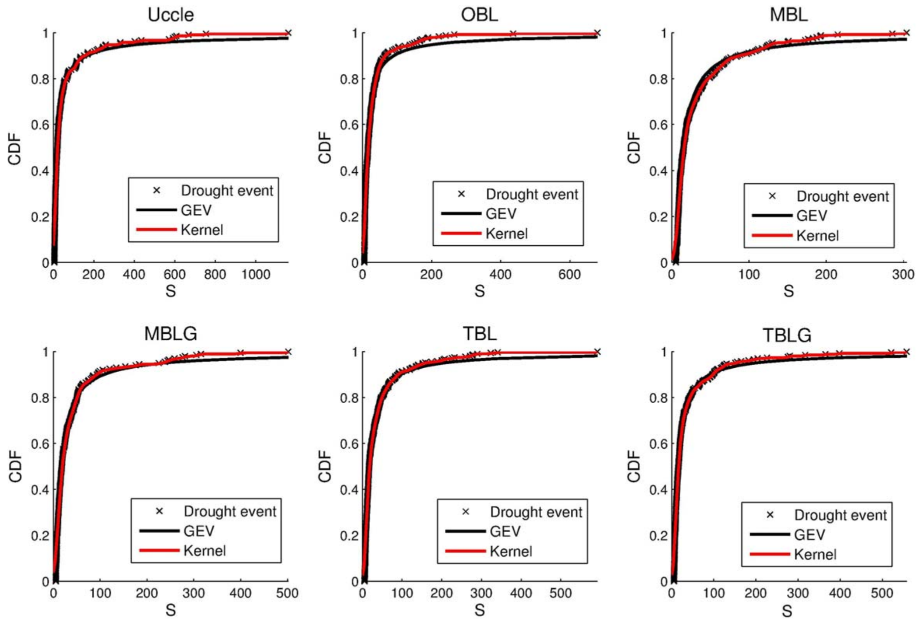

Fig. 6. GEV and Kernel cumulative distribution functions of drought severity S.

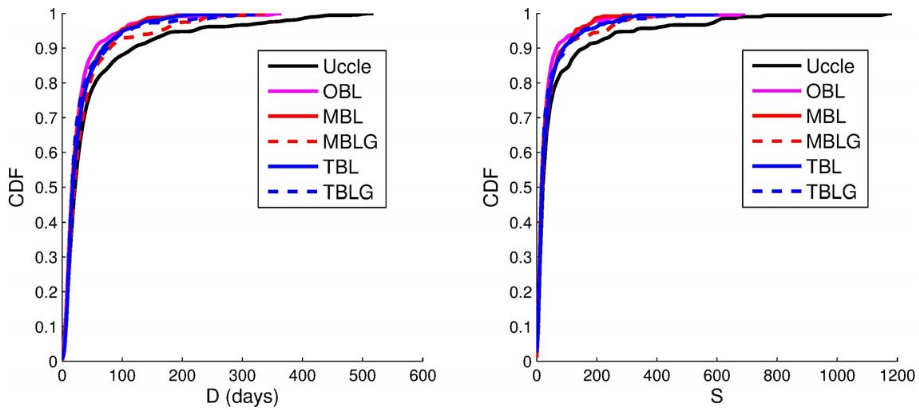

Fig. 7. Comparison of Kernel cumulative distribution functions of D and S of observed and simulated data.

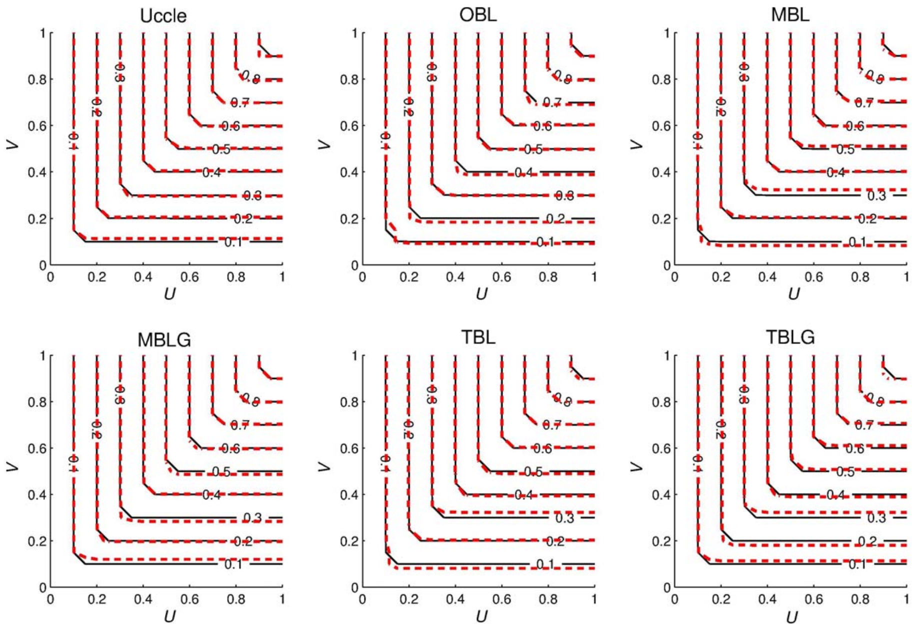

Fig. 8. Fitted Frank copulas (full line) and empirical copulas (dotted red line) for all data set.

HESSD
10, 7469–7516, 2013

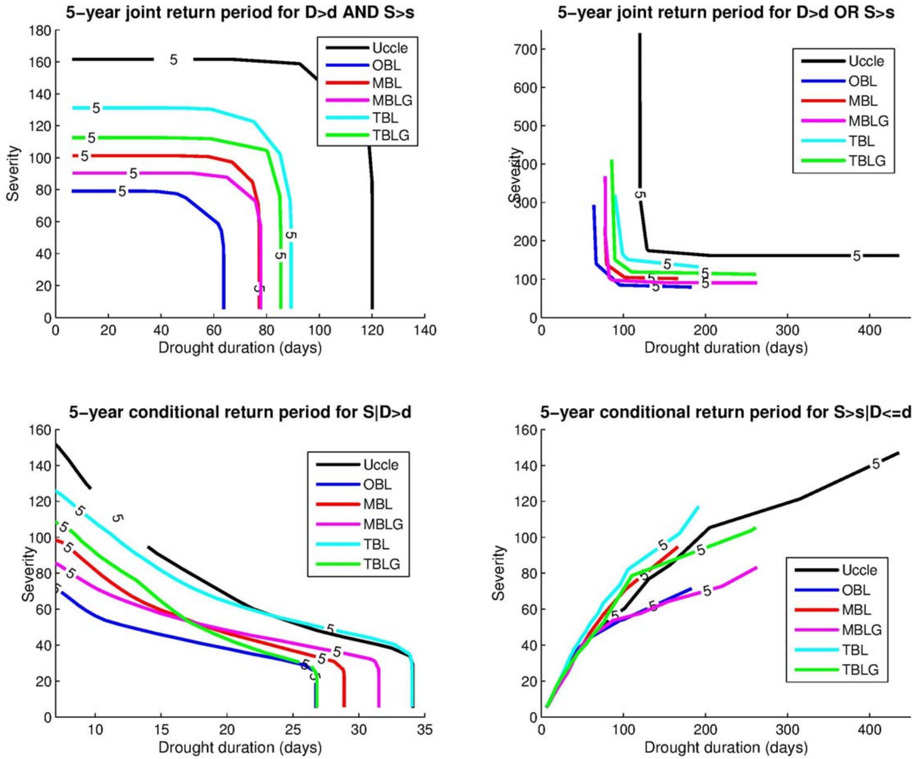

Fig. 9. 5 yr return period for droughts simulated by Uccle (black), OBL (blue), MBL (red), MBLG (pink), TBL (cyan) and TBLG (green) with $\{D \geq d \text{ AND } S \geq s\}$ (left, above); $\{D \geq d \text{ OR } S \geq s\}$ (right, above); $\{S|D > d\}$ (left, below); $\{S > s|D \leq d\}$ (right, below).

HESSD
10, 7469–7516, 2013

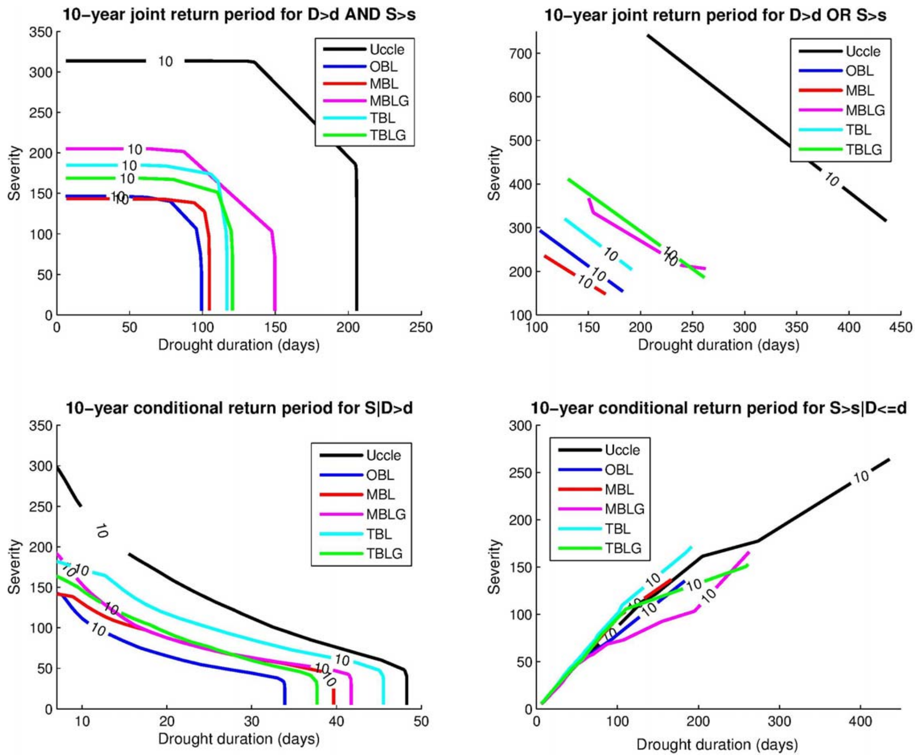

Fig. 10. 10 yr return period for droughts simulated by Uccle (black), OBL (blue), MBL (red), MBLG (pink), TBL (cyan) and TBLG (green) with $\{D \geq d$ AND $S \geq s\}$ (left, above); $\{D \geq d$ OR $S \geq s\}$ (right, above); $\{S|D > d\}$ (left, below); $\{S > s|D \leq d\}$ (right, below).