

# Large Scale Probabilistic Drought Characterization Over Europe

Brunella Bonaccorso · David J. Peres ·  
Antonino Cancelliere · Giuseppe Rossi

Received: 31 December 2011 / Accepted: 12 October 2012 /  
Published online: 23 October 2012  
© Springer Science+Business Media Dordrecht 2012

**Abstract** A reliable assessment of drought return periods is essential to help decision makers in setting effective drought preparedness and mitigation measures. However, often an inferential approach is unsuitable to model the marginal or joint probability distributions of drought characteristics, such as drought duration and accumulated deficit, due to the relatively limited number of drought events that can be observed in the historical records of the hydrological variables of interest. As an alternative, the marginal and multivariate probability cdf's of drought characteristics can be derived as functions of the parameters of the cdf of the underlying variable (e.g. precipitation), whose sample series is usually long enough to obtain trustworthy estimates in a statistical sense. In this study, the latter methodology is applied to investigate space-time variability of drought occurrences over Europe by using the CRU TS3.10.01 precipitation dataset for the period 1901–2009. In particular, a methodology able to take into account autocorrelation in the underlying precipitation series is adopted. First, a spatial analysis of historical droughts at European level is carried out. Then, the joint probability distributions of drought duration and accumulated deficit are derived for each cell, with reference to both historical and design drought events. Finally, the corresponding bivariate drought return periods are computed, as the expected values of the interarrival time between consecutive critical droughts. Results show that several heavy drought episodes have widely affected the continent. Among the

---

B. Bonaccorso (✉)

Department of Civil Engineering, University of Messina, C.da di Dio, S. Agata, 98166 Messina, Italy  
e-mail: bbonaccorso@unime.it

D. J. Peres · A. Cancelliere · G. Rossi

Department of Civil and Environmental Engineering, University of Catania, V.le A. Doria 6, 95125 Catania, Italy

D. J. Peres

e-mail: djperes@dica.unict.it

A. Cancelliere

e-mail: acance@dica.unict.it

G. Rossi

e-mail: grossi@dica.unict.it

most recent events, drought occurred during the period 1985–1995 was the worst in terms of extent of the regions characterized by return periods greater than 250 years. Besides Euro-Mediterranean regions, North Western and Central Eastern regions appear more drought prone than the rest of Europe, in terms of low values of return periods.

**Keywords** Drought · Probabilistic characterization · Return period · Large scale analysis

## 1 Introduction

An in depth knowledge of drought phenomena plays an important role for an appropriate planning and management of water resources (Yevjevich et al. 1983). Indeed, severe and prolonged drought events have affected a large part of Europe during the last decades (Zaidman et al. 2001; Lloyd-Hughes and Saunders 2002; Fink et al. 2004; Hannaford et al. 2011), with harmful impacts on public water supply, industrial and agricultural production, as well as on the environment, with special reference to the degradation of aquatic ecosystems.

According to the last report of the Intergovernmental Panel on Climate Change (IPCC) released in 2007, “drought is likely to intensify in both duration and severity”, with special reference to Southern Europe, as a consequence of global warming. As a matter of fact, a clear trend towards drier conditions during the 20th century has been documented in several studies carried out on drought indices time series computed on reanalysis datasets (Bordi et al. 2009), gridded data (Lloyd-Hughes and Saunders 2002; Sousa et al. 2011) and observations from meteorological stations (Van der Schrier et al. 2006) covering Europe and a large part of the Mediterranean basin.

In light of lessons learnt from past experiences of coping with severe drought events, and of the potential intensification of such phenomena in the next future, a common awareness has risen about the need to develop and implement advanced drought risk management strategies. This requires on the one hand a better understanding and modelling of the natural phenomenon and of its impacts on economic activities, society and environment, and on the other hand the definition and implementation of adequate long-term measures, oriented to significantly reduce vulnerability of water supply systems to drought events, as well as of short term measures, oriented to minimize drought impacts (Wilhite et al. 1987; Rossi 2000).

Probabilistic characterization of historical drought events, properly identified on hydro-meteorological series of interest, is necessary to achieve the former objective, as it can help to assess drought hazard over a region, which combined with drought vulnerability yields the corresponding drought risk. In particular, estimation of drought return periods can provide useful information for an appropriate water use planning under drought conditions.

Over the years, many approaches have been suggested for characterizing droughts. Yevjevich (1967) used the theory of runs to characterize droughts as a sequence of consecutive intervals where the water supply variable remains below a threshold level (somehow representative of water demand), preceded and succeeded by values above the threshold. Thus, each drought event can be characterized by two main properties, namely drought duration, and accumulated deficit, defined as the sum of single deficits, i.e. the deviations of the variable from the threshold, over drought duration. Such characteristics are statistically dependent and therefore a multivariate approach should be employed for their probabilistic analysis. However, due to the relatively limited number of drought events that can be observed from the historical records, fitting parametric distributions to observed

drought characteristics is unsuitable to model the marginal or joint probability distributions of drought characteristics, such as drought duration and accumulated deficit.

A traditional solution involve data generation by applying stochastic models (Millan and Yevjevich 1971; Kendall and Dracup 1992; Shiau and Shen 2001).

An alternative approach (Downer et al. 1967; Llamas and Siddiqui 1969; Sen 1976; Sharma 1995; Bonaccorso et al. 2003; Cancelliere and Salas 2004, 2010) consists in deriving the marginal and multivariate probability cdf's of drought characteristics as functions of the parameters of the cdf of the underlying variable (e.g. precipitation), whose sample series is usually long enough to obtain reliable estimates in a statistical sense.

Once that such probability distributions are derived, return period of different type of drought events, with respect to one or more features (i.e. drought duration and/or accumulated deficit) can be determined as well.

Since drought can span several years, it is not possible to identify a unique time unit (or trial) with respect to which, the exceedance probability  $P[X_t > x_t]$  can be expressed, as one can usually make in flood frequency analysis, where the return period can be evaluated by the well known formula  $T = 1/P[X_t > x_t]$  (Fernandez and Salas 1999).

In general drought return period can be defined as the expected value of the interarrival time between consecutive critical droughts (Loaiciga and Mariño 1991; Fernandez and Salas 1999; Shiau and Shen 2001; Bonaccorso et al. 2003; Cancelliere and Salas 2010).

In this regard, Shiau and Shen (2001) developed a procedure for deriving the return period of accumulated deficit, assuming independent and identically distributed events. Return period is defined as the expected value of the average interarrival time between two successive events with accumulated deficit greater than or equal to a fixed value. This procedure has been further extended to the case of drought events characterized by both drought duration and accumulated deficit or drought duration and intensity (Bonaccorso et al. 2003; Gonzalez and Valdes 2003) as well as to the case of periodic series such as monthly or seasonal variables (Cancelliere and Salas 2004). Recently, Cancelliere and Salas (2010) have modified the foregoing methodology in order to deal with autocorrelated series.

In the present study, the latter procedure is applied to investigate space-time variability of meteorological drought occurrences over Europe, by using the annual precipitation series retrieved by the CRU TS3.10.01 gridded dataset for the period 1901–2009. Such dataset covers uniformly the globe and is freely accessible through the British Atmospheric Data Centre (BADC) website <http://badc.nerc.ac.uk>.

Firstly, the time dependence structure of the annual precipitation series under consideration is analyzed by verifying the significance of the lag-1 autocorrelation values through the Anderson's correlogram test (Anderson 1941).

Then, the goodness of fit of several probability distributions (normal, lognormal and gamma) to the considered precipitation dataset is verified cell by cell through the Lilliefors test (Lilliefors 1967, 1969; 1973; Crutcher 1975), and the best distribution is chosen for each cell based on the lowest value of test statistic. Next, historical droughts at European level are identified through the theory of runs and classified based on the probabilities of occurrence of annual deficit.

Finally, once that the marginal and bivariate cdf's of drought duration and accumulated deficit for each cell are analytically derived, based on the parameters of the probability distribution of annual precipitation series, the return periods of selected historical droughts, which have heavily affected a large part of Europe, are analyzed. In addition, return periods of design critical droughts, with respect to fixed drought duration and accumulated deficit, are computed and the corresponding spatial distributions are analyzed.

A brief description of the methodology for the analytical derivation of the probability distributions of drought characteristics and for the assessment of drought return period is reported in Section 2. In Section 3, the results of the application of the proposed methodology are illustrated. Conclusions are drawn in Section 4.

## 2 Methodology

### 2.1 Probability Distributions of Drought Characteristics

Let  $X_t, t=1, 2, \dots$ , be a time series of the hydrological variable of interest and  $x_0$  the threshold level. The drought duration  $L$  is defined as the number of consecutive intervals where  $X_t \leq x_0$ , followed and preceded by at least one interval where  $X_t > x_0$ , whereas the accumulated deficit  $D$  is defined as the sum of single deficits  $S_t = x_0 - X_t$  over the duration  $L$ . It follows that the accumulated deficit can be expressed as:

$$D = \sum_{t=1}^L S_t = \sum_{t=1}^L (x_0 - X_t) \quad \text{for } X_t \leq x_0 \quad (1)$$

In what follows the analytical derivations of the marginal and joint pdf's and cdf's of drought duration  $L$  and accumulated deficit  $D$  for the general case of autocorrelated  $X_t$  series are briefly described.

#### 2.1.1 Probability Distribution of Drought Duration

Statistical properties of run length for stationary series have been far back derived in literature (Downer et al. 1967; Feller 1968; Llamas and Siddiqui 1969) and then applied to drought duration (Sen 1976; Guven 1983; Sharma 1995).

Assuming that the sequence of deficits and surpluses can be modelled by a stationary lag-1 Markov chain, it can be shown (Sen 1976) that the probability mass function (*pmf*) of drought duration  $L$  is geometric, with parameter  $p_{01}$ :

$$f_L(\ell) = P[L = \ell] = (1 - p_{01})^{\ell-1} p_{01} \quad (2)$$

The parameter  $p_{01}$  represents the transition probability from a deficit to a surplus, namely  $p_{01} = P[X_t > x_0 | X_{t-1} \leq x_0]$ . Equation (2) enables to compute the probability that a drought will last exactly  $\ell$  time-steps. The expected value and the variance of drought duration  $L$  follow from Eq. (2) as:

$$E[L] = \frac{1}{p_{01}} \quad (3)$$

$$\text{Var}[L] = \frac{1 - p_{01}}{p_{01}^2} \quad (4)$$

In Appendix A, details about the estimation of the transition probability  $p_{01}$  are provided.

The adequacy of the lag-1 Markov hypothesis for modeling the sequence of deficits and surpluses has been investigated, among others, by Cancelliere and Salas (2010) and Akyuz et al. (2012). In particular, Cancelliere and Salas (2010) have shown that for low to moderate

autocorrelations in the underlying hydrological series the Markov hypothesis holds with a fair approximation, while for stronger autocorrelations, models with a stronger time dependence should be employed, as confirmed by Akyuz et al. (2012).

2.1.2 Joint Probability Distribution of Drought Duration and Accumulated Deficit

The joint probability distribution of drought accumulated deficit  $D$  and duration  $L$  can be expressed as follows (Salas et al. 2005):

$$f_{D,L}(d, \ell) = f_{D|L=\ell}(d) \cdot f_L(\ell) \tag{5}$$

where  $f_{D|L=\ell}(d)$  is the conditional probability distribution of accumulated deficit conditioned by drought duration, namely  $D|L$ .

Exact analytical derivation of the conditional probability distribution of  $D|L$  is still an unsolved problem (except for a few simple cases), due to the mathematical difficulties that generally prevent closed form solutions (Millan and Yevjevich 1971; Sen 1976; Chung and Salas 2000).

In order to overcome analytical difficulties, some authors have assumed a parametric distribution for  $D|L$ , and have estimated the parameters from observed droughts (e.g. Guven 1983; Sharma 1995; Shiau and Shen 2001). In some cases, due to the limited number of droughts that can be observed from the available records, synthetic generation (Shiau and Shen 2001) or long series reconstructed from tree rings records (Gonzalez and Valdes 2003; Biondi et al. 2005) have been utilized.

An alternative approach consists in evaluating the parameters of the distributions of  $D|L$ , a priori selected, based on the parameters of the distribution of  $X_t$ , which, in turn, can be assessed with good approximation through the traditional inferential approach applied to common hydrometeorological sample series with sufficient length (typically more than 30 observations).

For instance, under the assumption of serially independent series, the first two moments of  $D|L$  are given by:

$$E[D|L = \ell] = E\left[\sum_{t=1}^L S_t \mid L = \ell\right] = \ell E[S_t] \tag{6}$$

$$\text{Var}[D|L = \ell] = \text{Var}\left[\sum_{t=1}^L S_t \mid L = \ell\right] = \ell \text{Var}[S_t] \tag{7}$$

The pdf of  $S_t$  is equal to the truncated distribution of  $X(t)$  (Bonaccorso et al. 2003), i.e.:

$$f_{S_t}(s_t) = \frac{1}{p_0} \cdot f_{X(t)}(x_0 - s_t) \cdot I(s_t)_{(0,\infty)} \tag{8}$$

where  $p_0 = P[X_t \leq x_0]$ , and  $I(s_t)$  is an indicator equal to 1 for  $0 < s_t < \infty$  and 0 otherwise.

Thus, the  $k$ -th moment of  $S_t$  is given by:

$$E[S_t^k] = \frac{1}{p_0} \int_0^\infty s_t^k \cdot f_{X_t}(x_0 - s_t) \cdot ds_t \tag{9}$$

Therefore, once that the moments of  $D|L$  are computed based on Eqs. (6) and (7), the parameters of the distribution of  $D|L$  can be evaluated through the method of moments.

Clearly, if the underlying series exhibit a strong autocorrelation, Eqs. (6) and (7) are no longer valid. In this case, computation of the conditional moments of accumulated deficit is more difficult, and deriving a closed form solution becomes cumbersome. Empirical approximations have been provided by Cancelliere and Salas (2010), that enable to compute the conditional moments as a function of the skewness coefficient of the underlying variable  $X_t$  of the lag-1 autocorrelation  $\rho_1$ , and of the threshold parameterized as  $x_0 = \mu_x - \alpha \sigma_x$  where  $\mu_x$  and  $\sigma_x$  are the mean and standard deviation of  $X_t$  and  $\alpha$  is a parameter ranging between 0 and 1.

In particular, the proposed approximate expressions for the mean and the variance of accumulated deficit of fixed duration  $\ell$ , are (Cancelliere and Salas 2010)

$$E[D|L = \ell] = \mu_S a_m \ell^{b_m} \tag{10}$$

$$\text{Var}[D|L = \ell] = \sigma_S^2 a_v \ell^{b_v} \tag{11}$$

where  $\mu_S$  and  $\sigma_S^2$  are respectively the expected value and variance of a single year deficit  $S_t$  assuming that the series are skewed but uncorrelated, which are functions of the underlying marginal distribution and the threshold  $x_0$  (as from Eq. (9)), and the parameters  $a_m, b_m, a_v$  and  $b_v$  are related to  $\rho_1$  and  $x_0$  through the following expressions:

$$a_m = (-0.6634\alpha - .3418)\rho_1^2 + (0.6983\alpha - .5592)\rho_1 + 1.0 \tag{12a}$$

$$b_m = (0.1865\alpha + 0.0839)\rho_1^2 + (-0.1840\alpha + 0.5903)\rho_1 + 1.0 \tag{12b}$$

$$a_v = (-0.7969\alpha - 0.0928)\rho_1^2 + (0.7415\alpha - 1.0325)\rho_1 + 1.0 \tag{13a}$$

$$b_v = (0.4175\alpha + 0.5707)\rho_1^2 + (-0.4414\alpha + 1.078)\rho_1 + 1.0 \tag{13b}$$

Note that for  $\rho_1=0$ ,  $E [D|L=\ell]=\ell\mu_S$  and  $\text{Var}[D|L = \ell] = \ell\sigma_S^2$  since  $a_m=b_m=a_v=b_v=1$ . By assuming  $D|L=\ell$  beta distributed, the pdf takes the form (Johnson et al. 1994):

$$f_{D|L=\ell}(d) = \frac{1}{B(p, q)} \frac{(d - a)^{p-1}(b - d)^{q-1}}{(b - a)^{p+q-1}} \quad (a \leq d \leq b) \tag{14}$$

where  $B(p, q) = \int_0^\infty y^{p-1}(1 - y)^{q-1} dy$  is the complete beta function, and  $a$  and  $b$  are the lower and upper bounds, respectively. In our case,  $a=0$  and  $b=\ell x_o$ , since a drought of length  $\ell$  cannot have accumulated deficit greater than  $\ell x_o$ , and the parameters  $p, q$  can be estimated as a function of the first two moments of accumulated deficit  $\mu_D=E[D|L=\ell]$  and  $\sigma_D^2 = \text{Var}[D|L = \ell]$  as (Johnson et al. 1994):

$$p = \left(\frac{\mu_D}{\sigma_D}\right)^2 \left(1 - \frac{\mu_D}{\ell x_o}\right) - \frac{\mu_D}{\ell x_o} \tag{15}$$

$$q = \frac{\mu_D(\ell x_o - \mu_D)}{\sigma_D^2} - (1 + p) \tag{16}$$

where  $\mu_D$  and  $\sigma_D^2$  are determined from Eqs. (10) and (11) respectively. Then, the bivariate pdf of drought accumulated deficit and duration takes the following form:

$$f_{D,L}(d, \ell) = \frac{1}{B(p, q)} \frac{(d)^{p-1} (\ell x_0 - d)^{q-1}}{(\ell x_0)^{p+q-1}} f_L(\ell) \quad , \quad (0 \leq d \leq \ell x_0) \quad (17)$$

where  $f_L(\ell)$  is the pdf of drought duration (see Section 2.1.1).

By integrating appropriately the bivariate pdf's, the occurrence probability of various drought events can be found. In particular, with reference to specific drought events, it follows that:

- (1) for drought event  $E = \{D > d_0 \text{ and } L = \ell_0 (\ell_0 = 1, 2, \dots)\}$ :

$$P[D > d_0, L = \ell_0] = \int_{d_0}^{\ell_0 x_0} f_{D,L}(z, \ell_0) dz = f_L(\ell_0) \int_{d_0}^{\ell_0 x_0} \frac{1}{B(p, q)} \frac{(z)^{p-1} (\ell_0 x_0 - z)^{q-1}}{(\ell_0 x_0)^{p+q-1}} dz \quad (18)$$

- (2) for drought event  $E = \{D > d_0 \text{ and } L \geq \ell_0 (\ell_0 = 1, 2, \dots)\}$ :

$$P[D > d_0, L \geq \ell_0] = \sum_{\ell=\ell_0}^{\infty} \int_{d_0}^{\ell x_0} f_{D,L}(z, \ell) dz = \sum_{\ell=\ell_0}^{\infty} \left[ f_L(\ell) \int_{d_0}^{\ell x_0} \frac{1}{B(p, q)} \frac{(z)^{p-1} (\ell x_0 - z)^{q-1}}{(\ell x_0)^{p+q-1}} dz \right] \quad (19)$$

where  $z$  is a dummy variable of integration.

Furthermore, the marginal probability of droughts events  $E = \{D > d_0\}$ , namely  $P[D > d_0]$ , can be obtained from Eq. (19), by letting  $\ell_0 = 1$ .

### 2.2 Assessment of Drought Return Period

In general, the return period of a drought can be defined as the expected value of the average interarrival time  $T_E$  between two successive droughts, recognized “critical” with respect to one or more drought characteristics (Lloyd 1970; Loaiciga and Mariño 1991; Fernandez and Salas 1999; Shiau and Shen 2001).

Let  $E$  be a critical drought, and  $\bar{E}$  a non critical drought, by definition the return period of drought event  $E$  can be expressed as:

$$E[T_E] = E \left[ L | \{E\} + L_n + \sum_{j=2}^{N_l} (L_j | \{\bar{E}\} + L_{n_j}) \right] \quad (20)$$

where  $L | \{E\}$  is the duration of the first critical drought  $E$ ,  $L_n$  is the duration of the next non drought period,  $L_j | \{\bar{E}\}$  and  $L_{n_j}$  are the duration of  $j$ -th non critical drought event and the duration of the next  $j$ -th non drought period, and  $N_l$  is the number of drought and non drought events preceding the second critical drought.

Assuming independence between consecutive drought events, the above expectation can be simplified into (Gonzalez and Valdes 2003):

$$E[T_E] = \frac{E[L] + E[L_n]}{P[E]} \quad (21)$$

where  $P[E]$  is the probability of occurrence of a critical drought  $E$ , which can be determined once that the pdf or cdf of event  $E$  are known.

For the case of a lag-1 Markov process, both  $E[L]$  and  $E[L_n]$  can be obtained from Eq. (3) (in the latter case by replacing  $p_{01}$  with  $p_{10}$ ). Also,  $P[E]$  can be computed from the joint cdf's given in Eqs. (18)–(19) depending on the case.

Cancelliere and Salas (2010) have shown by Monte Carlo simulation that, although for autocorrelated process the independence assumption between drought events is not exactly met, yet Eq. (21) still provides an excellent approximation of drought return period.

### 3 Applications

#### 3.1 Data

Data used for large scale drought analysis are annual precipitation series derived from the CRU TS3.10.01 dataset for the period 1901–2009 (<http://badc.nerc.ac.uk>). This dataset is produced by the British Atmospheric Data Centre through a model provided by the Climate Research Unit (CRU) at the University of East Anglia. It includes monthly values of several climate variables (e.g. cloud cover, daily mean temperature and wet day frequency, among others) computed on high-resolution grid (0.5x0.5 degree).

The gridded dataset is based on an archive of monthly climate observations from more than 4000 weather stations distributed around the world. The database is checked for heterogeneities in the stations records using an automated method which includes the development of reference stations using neighbouring stations. For further details on the interpolation of the observed records onto the 0.5 latitude-longitude grid, readers may refer to Mitchell and Jones (2005).

The study area is the European region 31.75–71.25°N and 33.75°E–15.25°W.

#### 3.2 Autocorrelation Analysis and Goodness of Fit of Selected Probability Distributions to Precipitation Data

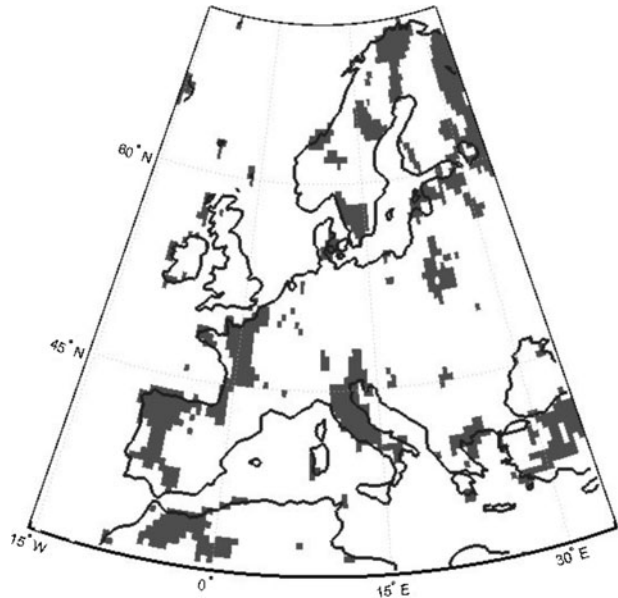
Although annual precipitation series are generally characterized by low values of autocorrelation, the Anderson's test (Anderson 1941) has been carried out to test the hypothesis that the sample lag-1 serial correlation (or autocorrelation) coefficient,  $r_1$ , is not significantly different from zero. If the hypothesis cannot be rejected, the series can be considered uncorrelated, otherwise it is to be assumed autocorrelated.

Figure 1 shows the results of the Anderson's test at 5 % significance level. White cells correspond to series for which the lag-1 autocorrelation is not significantly different from zero, whereas grey ones correspond to series that should be considered autocorrelated. Grey cells amount to about 23 % of the total study area, thus implying that a time dependence structure must be taken into account in the following drought analysis when dealing with the corresponding series.

Then, the goodness of fit of normal, lognormal and gamma probability distributions to the considered annual precipitation dataset has been checked cell by cell through the Lilliefors statistic test (Lilliefors 1967, 1969; 1973; Crutcher 1975). Such test is a modified version of the Kolmogorov Smirnov test, and is valid when the parameters of the underlying distribution are estimated from a sample.



**Fig. 1** Results of the Anderson’s test for the significance of the lag-1 autocorrelation coefficient (*grey* cells indicate that the hypothesis of uncorrelated series for the corresponding observed samples is rejected at 5 % significance level)



First, the goodness of fit of each distribution was verified at the 5 % significance level. Since test results did not clearly indicate a preference towards one distribution over all the investigated area, for each cell the distribution with the lowest Lilliefors test statistic value was selected. Table 1 shows, for each distribution, the percentage of cells for which the null hypothesis  $H_0$  (i.e. the probability distribution fits the data) is rejected at the 5 % significance level, as well as the percentage of cells with the lowest test statistic value. Although the gamma distribution shows a good fit for the most part of the cells, the normal distribution is the one to which corresponds the greater number of cells characterized by the lowest test statistic. Figure 2 illustrates the spatial coverage of probability distribution identified for each cell. Such figure also includes a few scattered cells (less than 2 % out of the total) for which  $H_0$  is rejected at the 5 % significance level for all the considered distributions.

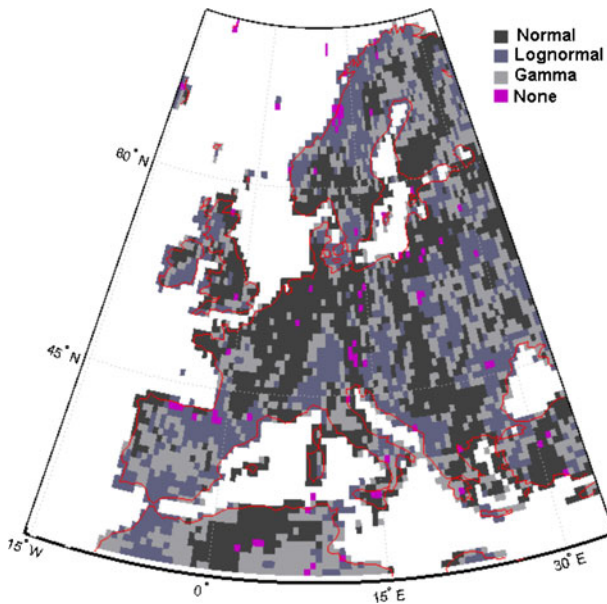
### 3.3 Drought Identification and Characterization

Drought conditions over Europe were identified by application of the theory of runs on annual precipitation gridded data, by considering a threshold level  $x_0$  equal to the median computed from the fitted distribution.

Dry and wet conditions with respect to the  $k$ -th cell were classified based on the probabilities of occurrence of annual deficit and surplus. In particular, the time series of

**Table 1** Comparison of the Lilliefors test results for the normal, lognormal and gamma probability distributions (Percentages are referred to a total amount of 4507 cells)

Probability distribution	Cells for which $H_0$ is rejected (%)	Cells with the lowest test statistic value (%)
Normal	14.55	40.12
Lognormal	12.40	27.67
Gamma	6.72	30.40



**Fig. 2** Spatial coverage of probability distributions, selected on the basis of the lowest Lilliefors test statistic value for each cell of CRU TS3.10.01 grid

non exceedance probability of annual precipitation data were computed by making use of probability distributions previously determined for each cell. Then dry and wet conditions were identified according to the classification reported in Table 2.

It is worth highlighting that, since the present work focuses on drought analysis, in what follows normal and wet conditions have been grouped into one class.

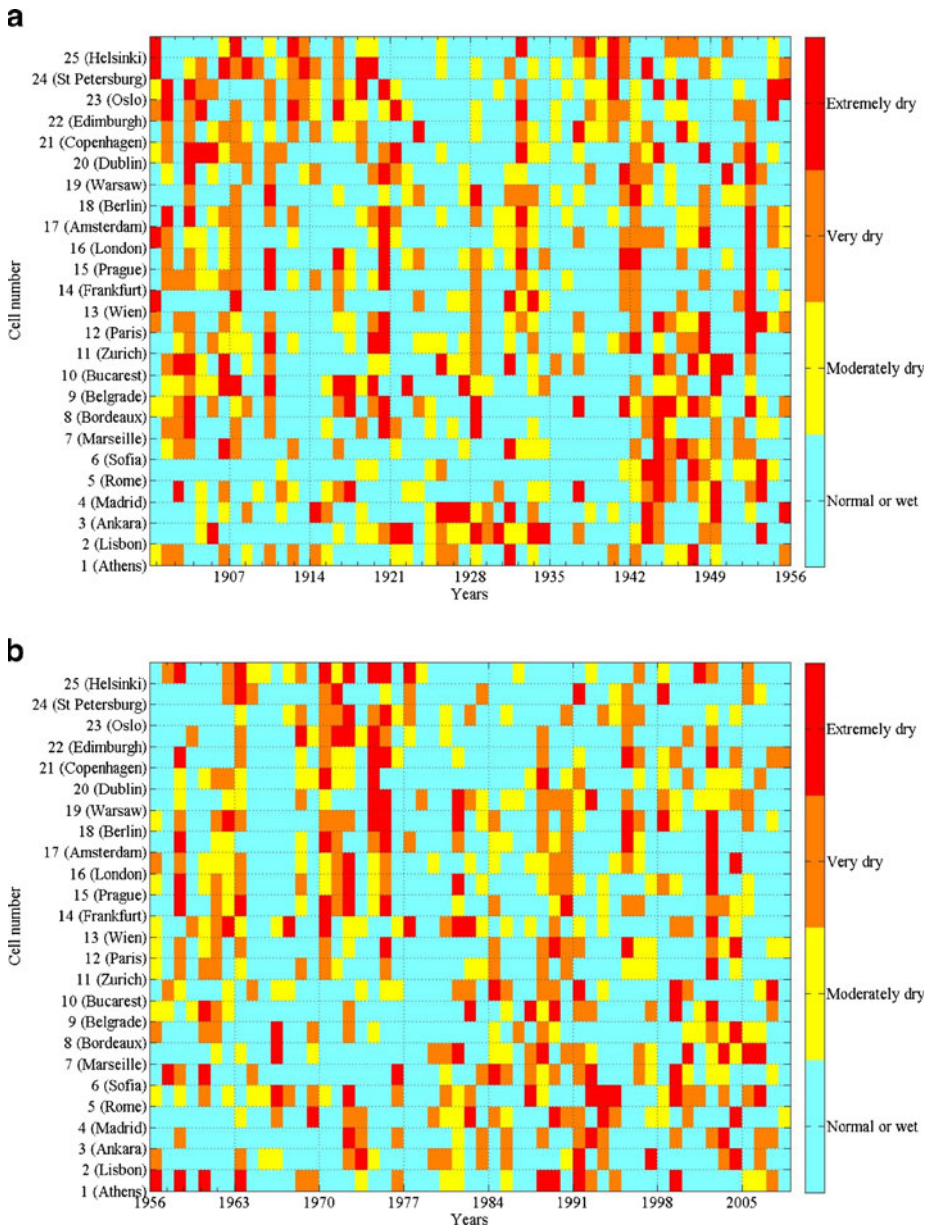
Figure 3 illustrates the temporal distribution of dry and wet periods from 1901 to 1956 (a) and from 1957 to 2009 (b), related to selected cells of the CRU TS3.10.01 grid including main cities in Europe. Although, the two images refer to a few cells, they enable to identify at a glance the most severe dry periods, as well as the corresponding duration and affected areas.

In particular, with reference to Fig. 3a, 1921 looks like an extremely dry year for many regions in North Western and Central Europe, also affected by longer dry periods during the first decade of the last century. During the '40s, long and severe dry periods have affected most part of Europe, with worse conditions in Southern Central Europe. Also, extremely dry conditions can be observed in 1953/1954 spread across the investigated area.

From the '60s onward (see Fig. 3b) the worst dry conditions seem to affect smaller area, with respect to the periods previously observed. For instance, one of the main dry periods, which has particularly affected North Western Europe, has been recorded during mid '70s.

**Table 2** Wet and dry period classification based on the probability of occurrence of annual deficit

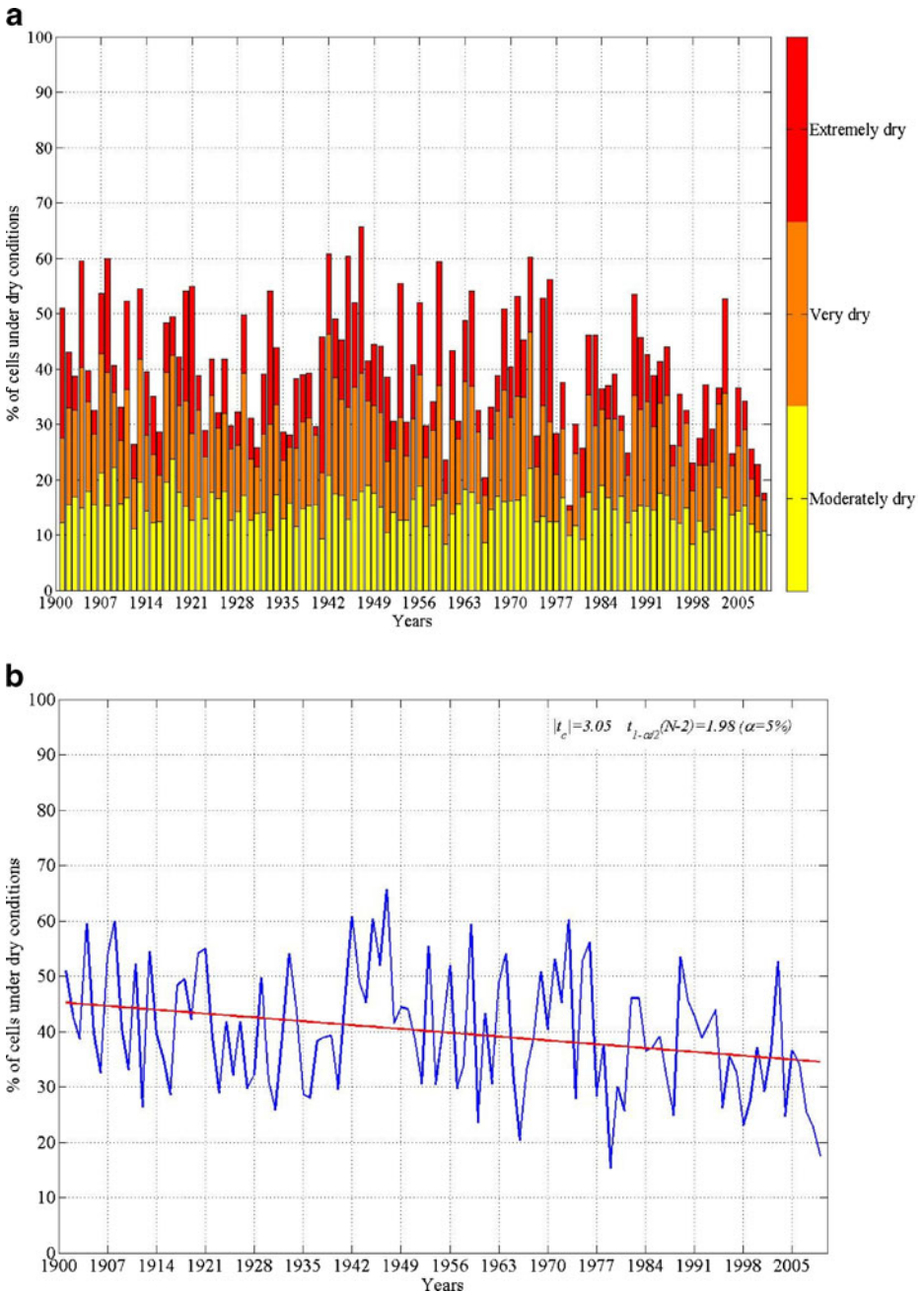
$P(x_i)$	Class
>0.40	Normal or wet
0.25–0.40	Moderately dry
0.10–0.25	Very dry
<0.10	Extremely dry



**Fig. 3** Dry and wet conditions over Europe from 1901 to 1956 (a) and from 1957 to 2009 (b)

Between the end of the '80s and mid '90s, severe dry periods have occurred mainly in the South side. Finally, it is worth mentioning the 2003 event characterizing Central Europe, and those occurred in 2004 and 2005 in Southern Europe, which extended into 2007/2008 in South Eastern regions.

In order to complement such an analysis, Fig. 4 illustrates the time series of the spatial coverage of dry conditions for the whole study area. In particular, Fig. 4a represents the time



**Fig. 4** Time series of the percentage of cells under different dry conditions from 1901 to 2009 (a) and trend pattern in the percentage of cells under dry conditions (b)

variation of spatial extent of different dry conditions. Results show that in 24 years out of 109, more than 50 % of the investigated area have been affected by droughts at the same time. Among these dry years is worth mentioning the periods: 1907–1908, 1920–1921,

1945–1947, 1975–1976, and more recently 1989 and 2003. Figure 4b illustrates the trend pattern of the time series of cells under dry conditions. From the figure, it can be observed a decreasing trend, which is significant at  $\alpha=5\%$  of the Student's t test. This result, which apparently contradicts the common perception, is actually in agreement with other previous studies. For instance in their study on drought and wetness trends in Europe, based on Standardized Precipitation Index (SPI) values computed on NCEP/NCAR reanalysis dataset of monthly precipitation from January 1948 to February 2009, Bordi et al. (2009) observe that the percentage of grid points in dry conditions noticeably decreases in the latest 15 year or so, in contrast with a positive trend detected in the previous 45 years. Also, Briffa et al. (2009), analyzing moisture availability using the Palmer Drought Severity Index (PDSI), derived from observed precipitation and temperatures records in 22 stations across Western and Central Europe, covering the period 1750–2003, conclude that recent widespread drying is apparent in the latter part of the 20<sup>th</sup> century, showing that anomalously high temperatures can be seen to be a major cause for the large areal extent of summer drought in the last two decades.

### 3.4 Return Periods of Historical Critical Droughts

Return periods of severe historical drought events have been computed, based on the methodology described in section 2. In particular, for each cell, precipitation series have been assumed either autocorrelated or not based on the results of the autocorrelation analysis illustrated in par. 3.2. Among the most severe drought events occurred in Europe during the latest 40 years, it is worth reminding those ones in 1972–1973 and 1975–1976, in 1988–1990 and 1992, and in 2000 and 2003–2005.

In order to take into account the different durations of concurrent droughts, a time span of 11 years approximately centered around the above mentioned events was considered. To this end, the following periods have been analyzed: 1968–1978, 1985–1995 and 1999–2009. Figure 5 illustrates the spatial distribution of the bivariate return periods of drought events characterized by  $\{D > D_o, L \geq L_o\}$ , where  $D_o$  and  $L_o$  are respectively the accumulated deficit and duration of historical drought identified in each cell for each period. In case that more than one event is identified in a cell in a given period, drought characterized by the largest accumulated deficit was selected for that cell.

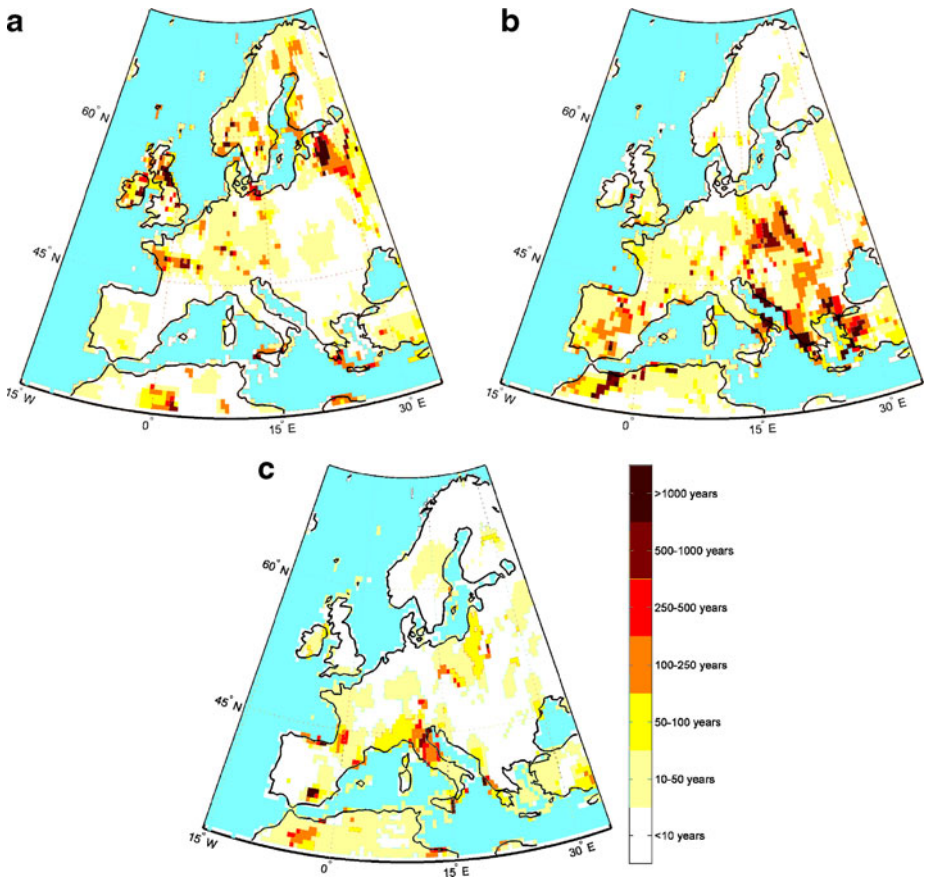
According to the adopted chromatic scale, drought return period increases as cell colour ranges from white to dark colors, identifying from common ( $T < 10$  years) to extremely rare events ( $T > 1000$  years). As it can be observed, almost the whole case study area is under drought condition during the three considered periods. Nonetheless, the areas affected by extreme drought events ( $T > 100$  years), change significantly from one period to another. In particular, the period 1985–1995 seems characterized by the broadest spread of extreme drought events, mainly localized on the Iberian peninsula, Central Europe, the Balkans and Western Turkey.

For the sake of completeness, Table 3 reports the percentage of cells with corresponding return periods greater than 100 and 250 years and mean drought characteristics. Results confirm that the period 1985–1995 has been the worst in terms of areal extent of extreme drought events occurred, as well as with respect to mean drought duration and accumulated deficit.

### 3.5 Return Periods of Design Critical Droughts

Previous results show that regions affected by extreme droughts (in terms of higher values of corresponding return periods) may significantly differ from one period to another. Thus, it





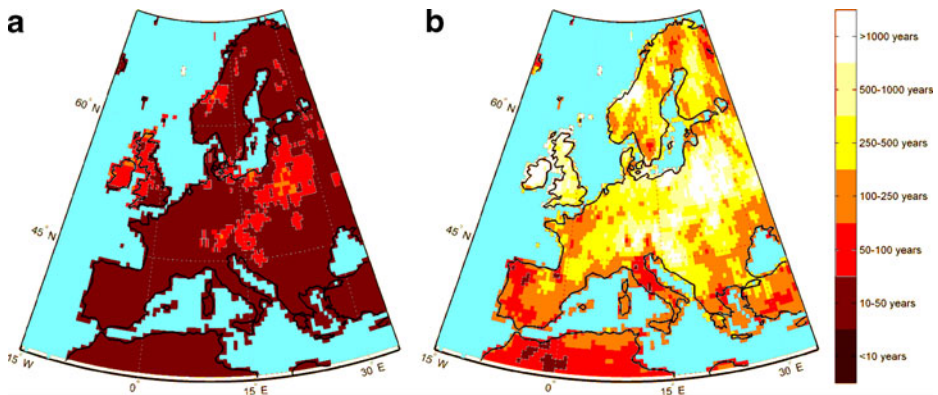
**Fig. 5** Spatial distributions of bivariate return periods of historical droughts occurred during 1968–1978 (a), 1985–1995 (b) and 1999–2009 (c)

could be interesting to analyze how the European regions respond to equal drought conditions, by fixing drought accumulated deficit and duration.

To this end, Fig. 6 illustrate return periods of two design droughts: the first (Fig. 6a) characterized by dimensionless accumulated deficit (i.e. divided by the drought threshold level) greater than 0.5 and duration greater than or equal to 3 years (a), the second (Fig. 6b) with dimensionless accumulated deficit greater than 1 and duration greater than or equal to

**Table 3** Percentage of cells with corresponding drought return periods greater than 100 and 250 years and mean drought characteristics for each selected period

Period	Cells with T>100 years (%)	Cells with T>250 years (%)	Mean characteristics	
			L (years)	D/x <sub>0</sub>
1968–1978	9.39	3.11	2.97	0.49
1985–1995	9.90	4.04	2.95	0.46
1999–2009	3.42	1.40	2.47	0.40



**Fig. 6** Spatial distributions of bivariate return periods of droughts with dimensionless accumulated deficit greater than 0.5 and duration greater than or equal to 3 years (a) and dimensionless accumulated deficit greater than 1 and duration greater than or equal to 5 years (b)

5 years (b). For Fig. 6, the chromatic scale has been inverted in order to better highlight drought prone areas, characterized by the lowest values of return periods.

From both figures, it clearly appears that Euro-Mediterranean regions are more drought prone than the rest of the continent. However, with special reference to drought with moderate accumulated deficit and limited duration (Fig. 6a) some North Western and Central Eastern regions also exhibit a marked trend towards frequent events (i.e. return periods less than 50 years).

#### 4 Conclusions

In this paper a methodology for drought characterization, proposed in previous studies (Bonaccorso et al. 2003; Cancelliere and Salas 2004, 2010), is applied for deriving the probability distributions of drought events, considering both drought duration and accumulated deficit, and for estimating the ensuing return periods.

The proposed approach has been used for a large scale drought analysis at European level, based on annual precipitation series derived from the CRU TS3.10.01 gridded dataset for the period 1901–2009. Different probability distributions, namely normal, lognormal and gamma, were fitted cell by cell to the data, based on the criteria of the lowest value of the Lilliefors test statistic. Historical dry and wet conditions have been identified through the theory of runs with a threshold equal to the theoretical median, and then classified based on the probabilities of occurrence of annual deficit and surplus.

The analysis of spatial and temporal distribution of dry and wet periods, although limited to selected cells of the CRU TS3.10.01 grid, has revealed that quite a few severe dry periods, affecting a large part of the continent, have occurred during the period of observation, such as at the beginning of the '20s, during the '40s, in the mid '70s, at the end of the '80s and in 2002/2003.

The time behaviour of the spatial coverage of dry conditions for the whole study area has shown a general significant decreasing trend, which leads one to believe that recent severe droughts have a reduced extent with respect to those occurred in the first mid of the past century. On the other hand, the comparison of the spatial distribution of return periods of droughts occurred during 1968–1978, 1985–1995 and 1999–2009, has highlighted worst

conditions during the second period, in terms of broader extent of the regions with a corresponding drought return periods greater than 250 years. Finally, the spatial distributions of return periods of droughts with fixed accumulated deficit and duration have revealed that, in addition to Euro-Mediterranean regions, some North Western regions (e.g. from Southern England to Germany) and Central Eastern regions (e.g. countries close to the Black Sea) are more drought prone.

It's worth pointing out that, although the joint analysis of drought characteristics presented in this study is built on the theory of runs applied on observed hydrometeorological series (i.e. annual precipitation), nevertheless it can be easily extended to drought characteristics identified on drought indices, such as the Standardized Precipitation Index, for instance by means of copula functions. Interested readers may refer to Shiau (2006) and Serinaldi et al. (2009) and reference therein for further details.

Further investigations are ongoing to identify homogeneous regions for characterizing meteorological droughts at European level.

**Acknowledgments** The financial support of the national project MIUR-PRIN 2008, "Water resources assessment and management under climate change scenarios" is gratefully acknowledged.

**Appendix A: Estimation of Transition Probabilities**

The transition probability  $p_{01}$  can be defined as:

$$p_{01} = 1 - p_{00} \tag{A.1}$$

where  $p_{00} = P[X_t \leq x_0 | X_{t-1} \leq x_0] = \frac{P[X_t \leq x_0, X_{t-1} \leq x_0]}{P[X_{t-1} \leq x_0]}$

Estimation of  $p_{01}$  can be carried out either through a frequency approach or by a parametric approach, assuming an underlying bivariate distribution for  $(X_t, X_{t-1})$ . Using the latter approach, assuming a bivariate normal distribution for  $(X_t, X_{t-1})$ , the transition probability can be estimated as:

$$\hat{p}_{01} = 1 - \frac{P[X_t \leq x_0, X_{t-1} \leq x_0]}{P[X_{t-1} \leq x_0]} = 1 - \frac{\Phi(x_0; \underline{\mu}, \underline{\Sigma})}{\Phi(x_0; \mu, \sigma^2)} \tag{A.2}$$

where in usual notation  $\Phi(x_0; \underline{\mu}, \underline{\Sigma})$  is the bivariate normal cdf evaluated at the point  $\underline{x}_0 = [x_0, x_0]'$  and  $\underline{\mu}$  and  $\underline{\Sigma}$  are the vector of the means and the variance-covariance matrix of  $(X_t, X_{t-1})$  respectively, while  $\Phi(x_0; \mu, \sigma^2)$  is the normal cdf with mean and variance  $\mu$  and  $\sigma^2$  respectively evaluated at  $x_0$ .

The above approach can also be employed when  $X_t$  is not normally distributed but it can be assumed that the transformed variable  $Y_t=f(X_t)$  is. For instance, if  $X_t$  is log normal distributed, then  $Y_t=\log(X_t)$  is normally distributed. Also, if  $X_t$  is gamma distributed, then  $Y_t=X_t^{1/3}$  is normal distributed as well. Then the transition probability in Eq. A.1 can be expressed as:

$$\hat{p}_{01} = 1 - \frac{P[X_t \leq x_0, X_{t-1} \leq x_0]}{P[X_{t-1} \leq x_0]} = 1 - \frac{P[f(X_t) \leq f(x_0), f(X_{t-1}) \leq f(x_0)]}{P[f(X_{t-1}) \leq f(x_0)]} \tag{A.3}$$

Note that having assumed an appropriate transformation for  $f(X_t)$  both probabilities in A.3 are normal and therefore can be computed as:



$$\hat{p}_{01} = 1 - \frac{P[f(X_t) \leq f(x_0), f(X_{t-1}) \leq f(x_0)]}{P[f(X_{t-1}) \leq f(x_0)]} = 1 - \frac{\Phi\left(\frac{y_0; \mu_y, \Sigma_y}{\sigma_y}\right)}{\Phi\left(y_0; \mu_y, \sigma_y^2\right)} \quad (\text{A.4})$$

where  $y_0 = f(x_0)$  and the means, variances and covariances are now related to the transformed variable  $Y_t$ .

## References

- Akyuz D, Bayazit M, Onoz B (2012) Markov chain models for hydrological drought characteristics. *J Hydrometeorol* 13:298–309. doi:10.1175/JHM-D-11-019.1
- Anderson RL (1941) Distribution of the serial correlation coefficient. *Ann Math Stat* 8(1):1–13
- Biondi F, Kozubowski TJ, Panorska AK (2005) A new model for quantifying climate episodes. *Int J Climatol* 25:1253–1264
- Bonaccorso B, Cancelliere A, Rossi G (2003) An analytical formulation of return period of drought severity. *Stoch Environ Res Risk Assess* 17:157–174
- Bordi I, Fraedrich K, Sutera A (2009) Observed drought and wetness trends in Europe: an update. *Hydrol Earth Syst Sci* 13:1519–1530
- Briffa KR, van der Schrier G, And Jones PD (2009) Wet and dry summers in Europe since 1750: evidence of increasing drought. *Int J Climatol* 29:1894–1905
- Cancelliere A, Salas JD (2004) Drought length properties for periodic-stochastic hydrological data. *Water Resour Res* 10(2):1–13
- Cancelliere A, Salas JD (2010) Drought probabilities and return period for annual streamflows series. *J Hydrol* 391:77–89
- Chung C, Salas JD (2000) Drought occurrence probabilities and risks of dependent hydrological processes. *J Hydrol Eng* 5(3):259–268
- Crutcher HL (1975) A note on the possible misuse of the Kolmogorov-Smirnov Test. *J Appl Meteorol* 14:1600–1603
- Downer R, Siddiqui M, Yevjevich V (1967) Application of runs to hydrologic droughts, Proc. of the International Hydrology Symposium. Colorado State University, Fort Collins
- Feller W (1968) An introduction to the probability theory and its application. Wiley, New York
- Fernandez B, Salas JD (1999) Return period and risk of hydrologic events. I: mathematical formulation. *J Hydrolog Eng* 4(4):297–307
- Fink AH, Brücher T, Krüger A, Leckebush GC, Pinto JG, Ulbrich U (2004) The 2003 European summer heatwaves and drought-synoptic diagnosis and impacts. *Weather* 59:209–216
- Gonzalez J, Valdes JB (2003) Bivariate drought recurrence analysis using tree ring reconstructions. *J Hydrol Eng* 8(5):247–258
- Guven O (1983) A simplified semiempirical approach to probabilities of extreme hydrologic droughts. *Water Resour Res* 19(2):441–453
- Hannaford J, Lloyd-Hughes B, Keef C, Parry S, Prudhomme C (2011) Examining the large-scale spatial coherence of European drought using regional indicators of precipitation and streamflow deficit. *Hydrol Process* 25:1146–1162
- IPCC (2007) Climate change 2007: synthesis report. In: Pachauri RK, Reisinger A (eds) Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change, core writing team. IPCC, Geneva, p 104
- Johnson N, Kotz S, Balakrishnan N (1994) Continuous Univariate distributions, vol. 1. Wiley, New York
- Kendall DR, Dracup JA (1992) On the generation of drought events using an alternating renewal-reward model. *Stoch Hydrol Hydraul* 6(1):55–68
- Lilliefors HW (1967) On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *J Am Stat Assoc* 62:399–402
- Lilliefors HW (1969) On the Kolmogorov–Smirnov test or the exponential distribution with mean unknown. *J Am Stat Assoc* 64:387–389
- Lilliefors, H.W. (1973). The Kolmogorov–Smirnov and other distance tests for the Gamma distribution and for the extreme-value distribution when parameters must be estimated. George Washington University, supported in part by U. S. Department of Commerce Contract 1-35214.
- Llamas J, Siddiqui M (1969) Runs of precipitation series. Hydrology paper 33. Colorado State University, Fort Collins

- Lloyd EH (1970) Return period in the presence of persistence. *J Hydrol* 10(3):202–215
- Lloyd-Hughes B, Saunders MA (2002) A drought climatology for Europe. *Int J Climatol* 22:1571–1592
- Loaiciga M, Mariño MA (1991) Recurrence interval of geophysical events. *J Water Resour Plann Manag (ASCE)* 117(3):367–382
- Millan J, Yevjevich V (1971) Probabilities of observed droughts. Hydrology Paper 50. Colorado State University, Fort Collins
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol* 25(6):693–712
- Rossi G (2000) Drought mitigation measures: a comprehensive framework. In: Vogt JV, Somma F (eds) *Drought and drought mitigation in Europe*. Kluwer Academic Publishers, The Netherlands, pp 233–246
- Salas JD, Fu C, Cancelliere A, Dustin D, Bode D, Pineda A, Vincent E (2005) Characterizing the severity and risk of droughts of the Poudre river. *J Water Resour Plann Manag* 131(5)
- Sen Z (1976) Wet and dry periods of annual flow series. *J Hydraul Div* 102(HY10):1503–1514
- Serinaldi F, Bonaccorso B, Cancelliere A, Grimaldi S (2009) Probabilistic characterization of drought properties through copulas. *Phys Chem Earth* 34:596–605. doi:10.1016/j.pce.2008.09.004
- Sharma T (1995) Estimation of drought severity on independent and dependent hydrologic series. *Water Resour Manag* 11:35–49
- Shiau JT (2006) Fitting drought duration and severity with two-dimensional copulas. *Water Resour Manag* 20:795–815
- Shiau J, Shen HW (2001) Recurrence analysis of hydrologic droughts of differing severity. *J Water Resour Plann Manag (ASCE)* 127(1):30–40
- Sousa PM, Trigo RM, Aizpurua P, Nieto R, Gimeno L, Garcia-Herrera R (2011) Trends and extremes of drought indices throughout the 20th century in the Mediterranean. *Nat Hazard Earth Syst Sci* 11:33–51
- van der Schrier G, Briffa KR, Jones PD, Osborn TJ (2006) Summer moisture variability across Europe. *J Clim* 19:2818–2834
- Wilhite DA, Easterling WE, Wood DA (1987) *Planning for drought*. Westview Press, Boulder
- Yevjevich V (1967) An objective approach to definitions and investigations of continental hydrologic droughts. Hydrology Paper 23. Colorado State University, Fort Collins
- Yevjevich V, Da Cunha L, Vlachos E (1983) *Drought research needs*. Water Resources Publication, Fort Collins
- Zaidman MD, Rees HG, Young AR (2001) Spatio-temporal development of stream flow droughts in north-west Europe. *Hydrol Earth Syst Sci* 5:733–751