

ON THE PROBABILISTIC ANALYSIS OF REGIONAL DROUGHTS

Antonino CANCELLIERE¹, Giuseppe ROSSI¹

(1) Department of Civil and Environmental Engineering, University of Catania, Italy, e-mail:
acance@dica.unict.it

ABSTRACT

Stochastic modelling of regional droughts is a topic of extreme interest due to its implications for a correct water resources management. The assessment of probabilities and return periods of regional drought characteristics represents a fundamental step toward a risk based water management in a region and can provide useful information for drafting drought management plans. In the paper, probabilistic features of regional droughts are analyzed in order to estimate the corresponding return periods. Exact and approximate analytical expressions are derived for the moments of areal coverage of deficit, of areal deficit and of accumulated deficit making use of realistic hypotheses about the distributions of the aforementioned characteristics as well as of copulas concepts. The derived moments, pdf and return period are applied for the characterization of regional droughts in Sicily island in Italy. The results of the application indicate that the derived expressions are able to model fairly well the occurrences and return period of droughts.

1 INTRODUCTION

Drought identification and characterization in a region is an important step in planning for drought preparedness and mitigation. While at-site analysis can provide useful information on drought occurrences in a limited area or in a river, regional analysis enables to identify droughts that affect significantly a large region by considering, besides length and accumulated deficit, also the areal extension of the drought (Tase, 1976; Santos, 1983; Rossi et al., 1992; Rossi and Cancelliere, 2003).

Tase (1976) investigated the relationship between area and deficit characteristics of drought events by simulating precipitation in a grid assuming a decaying exponential dependence function with distance. Probabilities of areal coverage of droughts were further investigated analyzing the effects of the size and the shape of the area. Sen (1980) derived the probability distributions of regional drought characteristics assuming independence in space for the precipitation field. Rossi (1983) investigated regional drought characteristics in Sicily by means of run method and concluded that due to the contemporaneous occurrence of droughts in different areas of the island, interbasin water transfer cannot be adopted as a mitigation measure. Santos (1983) analyzed the probabilistic features of regional drought characteristics, assuming precipitation distributed according to a multivariate normal. Bayazit and Onoz (2005) derived the pdf and return period of regional drought length assuming a Markovian spatial dependence

structure for streamflows in Turkey.

In the present paper, a probabilistic regional drought model is proposed that capitalizes on Santos (1983) formulation. More specifically, some of Santos (1983) assumptions are reviewed and alternatives are proposed that lead to an advanced model. Furthermore, return periods of alternative critical regional droughts are derived, using the derived distributions of drought characteristics as building blocks. As in Santos (1983) for the development of the model the derivation of the pdf of drought characteristics will be based on the statistics of the underlying precipitation.

The paper is organized as follows. First a review of regional drought characteristics is presented. Then moments and pdf of regional deficit and regional droughts characteristics are derived and the return period formulation originally proposed by Shiau and Shen (2001) is extended to the regional case. The methodology is then applied to precipitation series observed in Sicily region in the Mediterranean Sea and results obtained by the model are validated by simulation.

2 REGIONAL DROUGHT CHARACTERISTICS

The method of run, originally devised for at-site drought identification (Yevjevich, 1967), has been extended by Tase (1976) to the case of regional droughts, by considering series of the variable of interest (typically precipitation) observed at several stations and selecting, besides the truncation level at each site, an additional threshold, which represents the value of the area affected by deficit above which a regional drought is considered to occur. Thus, for a given interval t , two indices are computed, representative of the areal extension and amount of the deficits over the investigated region.

More specifically, let's assume that in a given area of interest there are m precipitation stations. Each station k can be characterized by the corresponding area of the Thiessen's polygon S_k expressed as a fraction a_k of the total area as:

$$a_k = \frac{S_k}{\sum_{j=1}^m S_j} \quad (1)$$

Throughout the paper, the following notation will be adopted: $X_{k,t}$ will indicate the precipitation at station k at time t , x_{ok} will indicate the threshold level at station k . Then the following indicator variable can be defined for station k :

$$\begin{aligned} I_{k,t} &= 1 \text{ if } X_{k,t} \leq x_{ok} \\ I_{k,t} &= 0 \text{ if } X_{k,t} > x_{ok} \end{aligned} \quad (2)$$

Obviously $I_{k,t}$ is a Bernoulli r.v. with probability $p_k = P[X_{k,t} \leq x_{ok}]$ of being 1. Furthermore, $I_{k,t} = 1$ indicates that a deficit occurs at time t at site k , i.e. $X_{k,t} \leq x_{ok}$.

The fraction of the area of the considered region affected by deficit in a given interval t will be termed here as areal coverage of deficit Ad_t and is defined as:

$$Ad_t = \sum_{k=1}^m a_k I_{k,t} \quad (3)$$

The regional extension of the run method assumes that in interval t a regional deficit occurs if the areal coverage of deficit exceeds a fixed threshold a_c (Tase, 1976). Accordingly, the regional deficit D_t can be computed as:

$$D_t = \sum_{k=1}^m a_k I_{k,t} (x_{ok} - X_{k,t}) \text{ if } Ad_t \geq a_c \quad (4)$$

The Ad_t index is a measure of the area affected by deficit, expressed as a fraction of the total area and it ranges between 0 and 1. The D_t index provides some insight on the total amount of the deficit in the area being basically a sum of the deficits at each sites, weighted by the corresponding fraction of the total area.

Then, regional drought is defined as an uninterrupted sequence of regional deficits preceded and succeeded by surpluses. A regional drought can be characterized by its duration L :

$$L = t_f - t_i + 1 \quad (5)$$

where t_i and t_f are such that $D_t > 0$ for $t_i \leq t \leq t_f$ and $D_{t_i-1} = 0$, $D_{t_f+1} = 0$.

Regional accumulated deficit is computed as:

$$DC = \sum_{t=t_i}^{t_f} D_t \quad (6)$$

Finally the mean areal coverage of drought can be computed by:

$$AD = \frac{1}{L} \sum_{t=t_i}^{t_f} Ad_t \quad (7)$$

3 MOMENTS AND DISTRIBUTIONS OF REGIONAL DEFICIT CHARACTERISTICS

It is of interest here to derive the moments and the pdf of the regional deficit characteristics based on the distribution of precipitation $X_{k,t}$.

More specifically we will assume that at a given time interval t the $X_{k,t}$ is jointly distributed according to a multivariate normal with vector mean $\hat{\mu}$ and variance-covariance matrix Σ whose generic element is $\gamma_{j,k} = \text{cov}[X_{j,t}, X_{k,t}]$. To simplify

notation, we introduce the r.v. $Y_{k,t} = a_k(x_{ok} - X_{k,t})$ which represents the deviation of areal precipitation at station k and time t from the threshold. Since $X_{k,t}$ is multivariate normal with variance σ_k^2 it follows $Y_{k,t} : N(a_k(x_{ok} - \mu_k), a_k^2 \sigma_k^2)$. Further, we can introduce the r.v.:

$$Z_t = \sum_{k=1}^m Y_{k,t} I_{k,t} = \sum_{k=1}^m a_k (x_{ok} - X_{k,t}) I_{k,t} \quad (8)$$

which represents the sum of precipitation deficits at time t , weighted by the areas of the Thiessen polygons a_k . The expected value of Z_t and variance have been derived by (Santos, 1983) as:

$$E[Z_t] = \sum_{k=1}^m \left[x_{ok} - \left[\mu_k - \frac{\sigma_k}{p_k} \phi\left(\frac{x_{ok} - \mu_k}{\sigma_k}\right) \right] \right] a_k p_k$$

$$\text{Var}[Z_t] = \sum_{k=1}^m \sum_{j=1}^m \text{Cov}[Y_{k,t}, Y_{j,t} | Y_{k,t} > 0, Y_{j,t} > 0] p_{k,j} \quad (9)$$

where $p_{k,j} = P[Y_{k,t} > 0, Y_{j,t} > 0]$ and as in usual notation $\phi()$ stands for the standard normal pdf. Also, given the normality of $X_{k,t}$, the probability p_k can be computed as

$$p_k = \Phi\left(\frac{x_{ok} - \mu_k}{\sigma_k}\right) \text{ where } \Phi(\cdot) \text{ stands for the standard normal cdf.}$$

The conditional covariance in Eq. (9) is the covariance of truncated bivariate normal r.v which can be computed numerically by the algorithm MVNTRC (Leppard and Tallis, 1989) while the joint bivariate normal probability can be computed numerically for example by means of the algorithm MULNOR (Shervish, 1984).

Since Z_t is the sum of truncated normal variables it will be approximated by means of a gamma distribution (Bonaccorso et al., 2003):

$$f_{Z_t}(z) = \frac{1}{\beta_z \Gamma(r_z)} e^{-\frac{z}{\beta_z}} \left(-\frac{z}{\beta_z}\right)^{r_z-1} \quad (10)$$

where the parameters r_z, β_z can be expressed in terms of the moments of Z_t given in eq. (9) as:

$$r_z = \frac{E[Z_t]^2}{\text{Var}[Z_t]} \quad \beta_z = \frac{\text{Var}[Z_t]}{E[Z_t]} \quad (11)$$

The first two moments of areal coverage of deficits can be computed as (Santos, 1983):

$$E[Ad_t] = \sum_{k=1}^m a_k p_k \quad \text{Var}[Ad_t] = \sum_{k=1}^m a_k p_k (1 - p_k) + 2 \sum_{k=1}^m \sum_{j=k+1}^m a_k a_j (p_{k,j} - p_k p_j) \quad (12)$$

Derivation of the exact distribution of the areal coverage of deficit is cumbersome except for simple cases. For example, Santos (1983) assumed Ad_t normally distributed

since it is the sum of r.v.'s therefore by virtue of the central limit theorem its asymptotic distribution will be normal. Such normality assumption however is not realistic, since the areal coverage of deficit Ad_t by definition is bounded between 0 and 1. Following Tase, (1976), here a beta distribution is adopted:

$$f_{Ad_t}(a) = \frac{1}{B(p,q)} a^{p-1} (1-a)^{q-1} \quad (0 \leq a \leq 1) \quad (13)$$

where $B(p,q)$ is the complete beta function $\int_0^1 w^{p-1} (1-w)^{q-1} dw$. In our case, the parameters p, q can be estimated as a function of the first two moments $\mu_A = E[Ad_t]$ and $\sigma_A^2 = Var[Ad_t]$ given by eqs. (12) as (Johnson et al., 1994):

$$p = \mu_A^2 - \frac{1 - \mu_A}{\sigma_A^2} - \mu_A \quad q = \mu_A - \frac{1 - \mu_A}{\sigma_A^2} - 1 - p \quad (14)$$

In order to derive the distribution of regional deficit D_t and its moments we will start by observing that by definition the Cumulative Distribution Function (cdf) of D_t is given by the following conditional cdf:

$$F_{D_t}(d) = F_{Z_t|Ad_t > a_c}(d) \quad (15)$$

Derivation of the conditional cdf on the l.h.s. of eq. (15) requires the definition of the bivariate cdf of Z_t and Ad_t and of the marginal of Ad_t . However, exact derivation of such bivariate distribution is quite complicated and therefore an approach based on copulas will be followed here. Copulas are mathematical tools that enable to "build" multivariate distributions starting from known marginals.

Several copulas have been proposed by different authors. Among the several copulas proposed, Frank's copula has been adopted here (Frank, 1979). Then, expressing by α a parameter that models the dependence between the two variables, a bivariate cdf of Z_t and Ad_t can be written as:

$$F_{Z_t, Ad_t}(z, a) = \alpha^{-1} \ln[1 + (e^{\alpha F_{Z_t}(z)} - 1)(e^{\alpha F_{Ad_t}(a)} - 1)(e^\alpha - 1)^{-1}] \quad (16)$$

Although in principle the dependence parameter α could be estimated from an available sample of Z_t and Ad_t , it may be preferable to estimate it by method of moments making use of the covariance between Z_t and Ad_t which in turn can be derived as a function of the underlying precipitation statistics. More specifically, the covariance between Z_t and Ad_t can be computed as:

$$\text{cov}[Z_t, Ad_t] = E[Y_{k,t} | I_{k,t} = 1, I_{j,t} = 1] p_{k,j} - E[Y_{k,t} | I_{k,t} = 1] p_k p_j \quad (17)$$

Once the the bivariate cdf of Z_t and Ad_t is known, the cdf of D_t follows as:

$$F_{D_t}(d) = F_{Z_t|Ad_t > a_c}(d) = \frac{F_{Z_t}(d) - F_{Z_t, Ad_t}(d, a_c)}{1 - F_{Ad_t}(a_c)} \quad (18)$$

where $F_{Ad_t}()$ is the marginal cdf of Ad_t given by eq. (13).

Then, the first two moments of regional deficit D_t can be computed as (Mood et al., 1974):

$$E[D_t] = \int_0^\infty [1 - F_{D_t}(z)] dz \quad \text{Var}[D_t] = \int_0^\infty 2z[1 - F_{D_t}(z)] dz \quad (19)$$

Although eq. (18) completely defines the distribution of D_t , for what follows it is convenient to approximate such distribution by a gamma distribution where the parameters r, β can be expressed in terms of the moments of D_t given in eq. (19) as:

$$r = \frac{E[D_t]^2}{\text{Var}[D_t]} \quad \beta = \frac{\text{Var}[D_t]}{E[D_t]} \quad (20)$$

Such gamma approximation enables to derive in a straightforward fashion the cdf of accumulated deficit as will be shown in the following paragraph.

4 MOMENTS AND PDF OF REGIONAL DROUGHT CHARACTERISTICS

From a formal standpoint, a regional drought differs from an at-site drought only in what concerns the definition of deficit. At-site deficit occurs whenever the variable of interest is below the threshold (Yevjevich, 1967), whereas, in the regional case, a regional deficit occurs whenever the areal coverage of deficit Ad_t exceeds a critical threshold a_c (Tase, 1976). Once a regional deficit occurs, the regional characteristics are computed in a very similar way as in the at site analysis. Therefore, probabilistic formulation of regional drought characteristics can inherit most of the results of the at-site case. By definition, a regional deficit occurs when the areal coverage of deficit Ad_t exceeds the areal threshold a_c . Thus, the occurrence probability p_d of a regional deficit is given by:

$$p_d = P[Ad_t > a_c] \quad (21)$$

The above probability can be computed by integrating the beta pdf of Ad_t , given in equation (13).

Under the hypothesis of non autocorrelated $X_{k,t}$ the regional drought duration will be distributed according to a geometric distribution with parameter $1 - p_d$ i.e.:

$$f_L(l) = P[L = l] = (1 - p_d) p_d^{l-1} \quad l = 1, 2, \dots \quad (22)$$

since the occurrence of regional deficits is a sequence of Bernoulli trials.

The expected value of L follows as:

$$E[L] = \frac{1}{1 - p_d} \quad (23)$$

The expected value and variance of the regional accumulated deficit DC can be derived from eq. (6) as:

$$E[DC] = \frac{E[D_t]}{1 - p_d} \quad \text{Var}[DC] = \frac{\text{Var}[D_t]}{1 - p_d} + \frac{p_d E[D_t]^2}{(1 - p_d)^2} \quad (24)$$

where the expected value and variance of D_t have already been derived in eq. (19).

Before tackling the marginal distribution of DC , it may be of some interest to derive an approximate expression for the joint pdf of regional drought length and accumulated deficit.

The regional accumulated deficit DC given by eq. (6) is the sum of L successive regional deficits D_t , where the regional drought duration L is a random variable. To derive its distribution, a scheme proposed by Salas et al. (2005) for the at site drought case will be extended to the regional case. More specifically, first, the derivation of the bivariate probability distribution function (pdf) of deficit and length can be pursued by considering the distribution of accumulated deficit conditioned on a fixed drought length $L = l$. Then the bivariate pdf of accumulated deficit D and length L can be expressed as:

$$f_{DC,L}(d, l) = f_{DC|L=l}(d) f_L(l) \quad (25)$$

where $f_{DC|L=l}$ is the conditional pdf of DC given $L = l$ and $f_L(l)$ is the marginal pmf of L . The r.v. $DC | L = l$ is the sum of exactly l regional deficits D_t , which have been assumed i.i.d gamma distributed with parameters r, β given by eq. (20). It follows that $DC | L = l$ will also be gamma distributed, with parameters (lr, b) (Mood et al., 1974; Guven, 1983; Shiau and Shen, 2001). On the other hand, it has already been shown that the pmf of regional drought length is geometric (see eq. (22)). Combining these results, the following expression for the bivariate distribution of DC and L is obtained:

$$f_{DC,L}(d, l) = \frac{1}{\beta \Gamma(lr)} e^{-\frac{d}{\beta}} \left(-\frac{d}{\beta} \right)^{lr-1} (1 - p_d) p_d^{l-1} \quad (26)$$

Therefore by integrating appropriately the bivariate pdf's, the occurrence probability of various drought events can be found. In particular, with reference to different types of critical drought events, the following expressions can be found (Salas et al., 2005):

(1) for drought event $E = DC > D_0$ and $L = l_0$ ($l_0 = 1, 2, \dots$):

$$P[DC > D_0, L = l_0] = \int_{D_0}^{\infty} f_{DC,L}(z, l_0) dz = \int_{D_0}^{\infty} \frac{1}{\beta \Gamma(l_0 r)} \left(\frac{z}{\beta} \right)^{l_0 r - 1} e^{-\frac{z}{\beta}} f_L(l) dz \quad (27)$$

(2) for drought event $E = DC > D_0$ and $L \geq l_0$ ($l_0 = 1, 2, \dots$):

$$P[DC > D_0, L \geq l_0] = \int_{D_0}^{\infty} \sum_{l=l_0}^{\infty} f_{DC,L}(z, l) dz = \int_{D_0}^{\infty} \sum_{l=l_0}^{\infty} \frac{1}{\beta \Gamma(lr)} \left(\frac{z}{\beta} \right)^{lr-1} e^{-\frac{z}{\beta}} f_L(l) dz \quad (28)$$

Furthermore, the marginal probability of droughts events $DC > D_0$ can be obtained from eq.(28), by letting $l_0 = 1$, thus yielding:

$$P[D > D_0, L \geq 1] = \int_{D_0}^{\infty} \sum_{l=1}^{\infty} f_{D,L}(z, l) dz = \int_{D_0}^{\infty} \sum_{l=1}^{\infty} \frac{1}{\beta \Gamma(lr)} \left(\frac{z}{\beta} \right)^{lr-1} e^{-\frac{z}{\beta}} f_L(l) dz \quad (29)$$

As already mentioned, the regional drought definition differs substantially from the corresponding at-site one only in what regards the definition of the regional deficit. It follows that, in order to derive an expression for the return period of critical regional droughts, the scheme similar to the one developed for the at-site case can be applied (Shiau and Shen, 2001; Bonaccorso et al., 2003; Gonzalez and Valdes, 2003; Salas et al., 2005).

For analyzing the return period of droughts we need to specify the regional drought event under consideration. Adopting the same scheme of Salas et al. (2005), let us denote by $\{E\}$ a critical drought event and by $\{\bar{E}\}$ the corresponding non critical drought events. For example, if $\{E\} = \{DC > DC_0, L \geq l_0\}$, the corresponding non drought events $\{\bar{E}\}$ will be $\{DC \leq D_0, L \geq l_0\}$, $\{DC > D_0, L < l_0\}$, and $\{DC \leq D_0, L < l_0\}$.

Assuming independence between drought events, the following expression for return period of droughts with characteristics E can be derived (Salas et al., 2005):

$$T(E) = \frac{E[L] + E[Ln]}{P[E]} \quad (30)$$

where L is the average duration of a critical drought event $\{E\}$, and Ln is the duration of a non-drought event $\{\bar{E}\}$.

Equation (30) allows one to compute the return period of any critical drought $\{E\}$ once the probability $P(E)$ is defined and the expected value of drought and non-drought durations are known. The former probability can be computed by making use of the joint pdf's of regional drought characteristics derived previously. The expected value of regional drought length L is given by Eq. ((23)) whereas it is easy to see that the non-drought length is distributed according to a geometric distribution with parameter p_d . Then, the return period of regional critical drought becomes:

$$T = \frac{1}{p_d(1 - p_d)P[E]} \quad (31)$$

5 APPLICATION OF THE REGIONAL MODEL

The regional model has been applied to Sicily island. Sicily is one of the largest islands in the Mediterranean sea and with a surface of approx. $25,000 \text{ km}^2$. The climate is semiarid, with mean annual precipitation around 700 mm and an high intra-annual variability as well as from year to year. Drought analysis in Sicily has been carried out with reference to 43 precipitation stations belonging to the former Hydrographic Service (now Observatory of Water). Available precipitation data include daily precipitation aggregated at annual scale covering the period 1921-2000 with occasional missing years that have been filled in by means of linear regression technique.

The location of the stations along with the related Thiessen's polygons is shown in Figure 1. As a first step, regional droughts that occurred in the island have been identified and characterized using the procedure outlined in previous paragraphs. In particular, Thiessen polygons areas a_k at each station have been computed by eq. (3). Also, for each station k a threshold corresponding to the observed sample mean has been assumed, i.e. $x_{ok} = \bar{X}_k$ for $k=1,2,\dots$, and fixed critical area thresholds $a_c = .5$ and $a_c = .7$. In Figure 2 the areal coverage of deficit Ad_t and the regional deficits D_t observed in Sicily during the period 1921-2000 are shown for $a_c = .5$. In addition, the plot indicates the various regional drought episodes that have occurred in Sicily throughout the historical record, corresponding to periods during which the areal coverage of deficit is above the threshold. For instance, the severe drought that occurred between the 1980's and 1990's is clearly shown. The figure confirms the intensification of droughts in the latter half of the century, with a tendency to longer and more severe droughts.

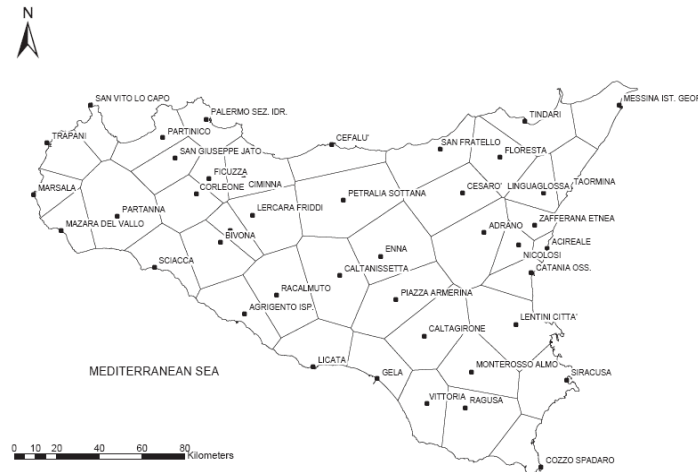


Figure 1: Localization of the investigated 43 precipitation stations in Sicily and related Thiessen's polygons

From Figure 2 it can be inferred that 19 regional droughts can be identified from the

available precipitation samples in Sicily for $a_c=.5$. Such number of droughts is not adequate to perform statistical inference on drought probabilities especially if a joint analysis of drought characteristics such as accumulated deficit and length is sought. To better illustrate the point, if the 19 regional droughts identified in Sicily are disaggregated according to the drought length, 10 droughts have length 1, 3 length 2, 5 length 3 and 1 length 9.

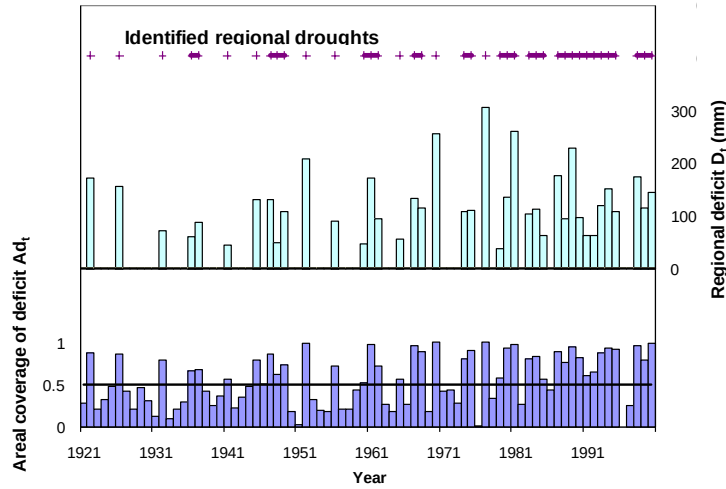


Figure 2: Areal coverage of deficit Ad_t , regional deficits D_t and identified regional droughts in Sicily during the 1921-2000 period ($a_c=.5$).

This means that it is not possible to obtain reliable estimates of the probability of droughts by fitting a parametric distribution to the observed droughts of fixed length, let alone estimates of return period.

Thus, in order to verify the analytical model developed previously, 200,000 years precipitation at the stations in Sicily have been generated assuming a normal distribution. Such normality assumption has not been rejected by means of a skewness test in 22 out of 43 stations at a significant level 5%. In particular, generation of normal vector $\underline{X}_t = \{X_{1,t}, X_{2,t}, \dots, X_{m,t}\}^T$ with $m = 43$ for Sicily has been carried out by first estimating the cross variance-covariance matrix of observed annual precipitation at the m stations $\hat{\Sigma}$.

The generic element $\hat{\gamma}_{i,j}$ of $\hat{\Sigma}$ represents the covariance between contemporaneous observations at station i and j and has been computed as:

$$\hat{\gamma}_{i,j} = \frac{\sum_{t=1}^m (x_{it} - \bar{x}_i)(x_{jt} - \bar{x}_j)}{\left(\sum_{t=1}^m (x_{it} - \bar{x}_i)^2 \sum_{t=1}^m (x_{jt} - \bar{x}_j)^2 \right)^{1/2}} \quad (32)$$

where $\bar{x}_i$ is the sample mean of annual precipitation at station i . Then, a random vector $\mathbf{X}_t$ has been generated as (Salas, 1993):

$$\underline{X}_t = \hat{\mathbf{B}}\underline{\varepsilon}_t + \hat{\underline{\mu}} \quad (33)$$

where $\hat{\mathbf{B}}$ is the Cholesky decomposition of $\hat{\hat{\Sigma}}$ (i.e a solution of the matrix equation $\hat{\mathbf{B}}\hat{\mathbf{B}}^T = \hat{\hat{\Sigma}}$), $\hat{\underline{\mu}} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m\}$ is the vector of the sample means of annual precipitation and $\underline{\varepsilon}_t$ is a vector of i.i.d. standard normal random variables.

After generation, regional drought identification in each investigated region has been carried out fixing for each station a threshold equal to the long term mean and two areal thresholds, namely $a_c = 0.5$ and $a_c = 0.7$. Then regional deficit and drought characteristics Ad_t , D_t , L and DC have been estimated on the generated samples. Corresponding moments have also been computed by means of the aforementioned expressions.

As a first step, the comparison between computed, generated and observed moments (i.e. computed from the observed samples) of Ad_t , D_t and DC has been carried out. The results of such comparison are reported in Table I, for the two fixed areal threshold values $a_c = .5$ and $a_c = .7$. It can be inferred from the table a general good agreement between the derived moments of deficit characteristics and corresponding moments computed on regional deficits identified on generated series, especially with respect to expected values. Also the moments computed from the observed sample generally match well the derived and generated ones for both regions.

Table I: Comparison of expected values and standard deviations of areal coverage of deficit Ad_t regional deficit D_t (mm) and regional cumulated deficit DC (mm) in Sicily for different areal thresholds a_c (M model, S simulated, O observed).

	$E[Ad_t]$		$\sqrt{\text{var}[Ad_t]}$		$E[D_t]$		$\sqrt{\text{var}[D_t]}$		$E[DC]$		$\sqrt{\text{var}[DC]}$	
a_c					.5	.7	.5	.7	.5	.7	.5	.7
M	0.50	0.29	120.1	155.5	71.5	72.5	241.7	222.4	199.2	149.7		
S	0.50	0.29	126.1	160.2	66.6	63.8	248.6	226.1	193.6	145.3		
O	0.54	0.30	123.7	145.9	62.7	56.3	260.4	243.2	232.3	136.5		

In order to verify the applicability of the derived pdf to compute return period of critical regional droughts, a comparison between return periods computed by means of eq. (27) or eq. (28) and corresponding sample estimates computed as the average interarrival time between drought events E identified on the generated and historical samples has been carried out. Figure 3a and 3b show the return periods of droughts defined jointly in terms of accumulated deficit and length as in case (1) above, i.e. $E = DC > D_0$ and $L = l_0$ ($l_0 = 1, 2, \dots$), which were obtained from the generated samples and from eq.(27) for $a_c = 0.5$ and $a_c = 0.7$ and for various values of the dimensionless accumulated deficit $\delta = D_0/\bar{\mu}$ where $\bar{\mu}$ is the average of the means of annual

precipitation. From the figure it can be inferred that, for a given drought duration l_0 , the return period increases rapidly with increasing δ , which means that estimating the return period T for large values of δ may require a very long sample. On the other hand, Fig. 3a and 3b indicate that using eq.(27) one can always find the results for any δ and any l_0 .

Figure 3a and 3b also show that results obtained from the generated samples and from eq. (27) are generally close. Furthermore, for all values of the deficit coefficient δ , as the drought duration l_0 increases, the return period T converges to the return period curve for $\delta = 0$. The comparison between the results obtained for the two critical areas considered reveals also some features. In particular, the curve related to $\delta = 0$, which represents critical droughts of fixed length, and of any accumulated deficit, is less steep in the case of $a_c = .5$ than in the other case. This is consistent with the fact that as the critical area increases, the occurrence probability of longer droughts will decrease, and therefore longer return period of longer droughts are expected. Figures 3a and 3b show also the return periods obtained from the historical sample for $\delta = 0$ and for some l_0 . In most cases, the return periods computed from the observed samples are lower than the generated or analytical counterpart. This can be explained by the fact that not many drought episodes can be observed from the historical records, and therefore the return periods estimated from the samples have a high variability.

The return periods of droughts defined as in case (2) above i.e. $E = DC > D_0$ and $L \geq l_0$ ($l_0 = 1, 2, \dots$), which were obtained from the generated samples and from eq. (29) for $a_c = 0.5$ and $a_c = 0.7$ and for various values of δ are shown in Figure 3c and 3d. The figure shows that as in case (1) above, as l_0 increases all return period curves converge to a single curve that is independent of δ . Furthermore the same figure confirms the close results that are obtained from the generated samples and from Eq.(29).

6 CONCLUSIONS

Probabilistic characterization of regional droughts is an important step toward a robust water resources planning, in particular when rainfed agriculture is predominant in the area, and therefore precipitation becomes the primary source of water. Drought assessment on a regional scale can also be useful to evaluate the appropriateness of interbasin water transfers as a drought mitigation measure.

Here run method has been adopted to derive analytically the distributional properties of regional drought characteristics, on the basis of the stochastic structure of the underlying precipitation series. The derived distributions enable to estimate return period of various types of critical droughts, also taking into account jointly different characteristics. The approach allows to overcome the difficulties associated with estimating probabilities and return periods of droughts identified on an historical sample, due to the limited number of droughts that are generally observed.

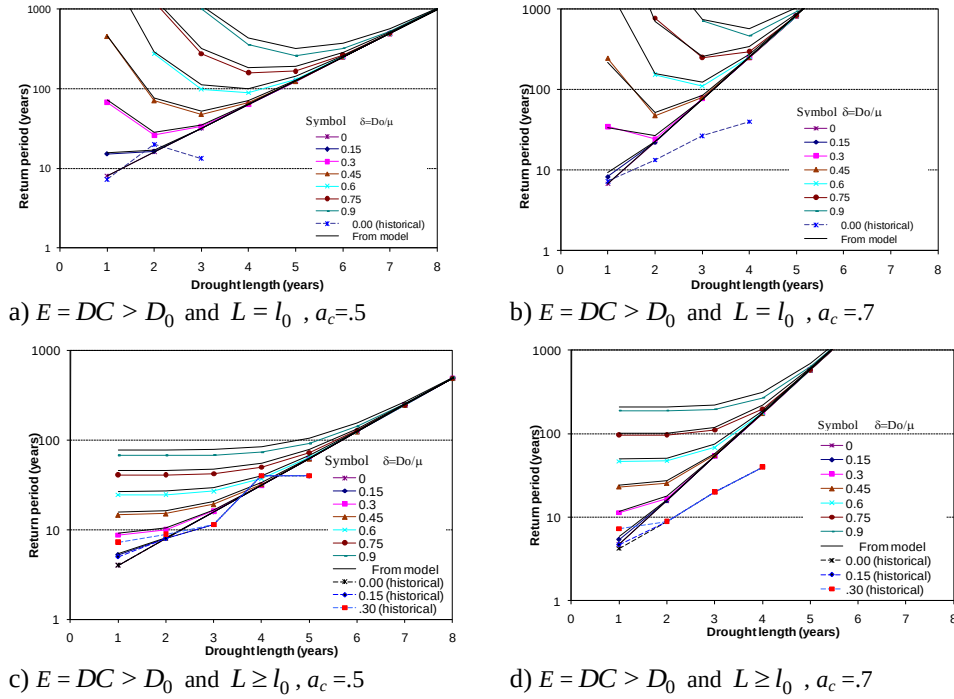


Figure 3: Return period of regional drought events in Sicily obtained from eq. (27) (a and b) or from eq. (28) (c and d) and from the generated sample for various values of the deficit coefficient δ ($D_0 = \delta\mu$) and for critical area $a_c = .5$ and $a_c = .7$. The return period curves for $\delta = 0$ obtained from the historical samples are also shown.

The analytical pdf's of drought characteristics are expressed as a function of the basic statistics of the underlying series (mean, cross-covariance) as well as of the demand thresholds and of the areal threshold. Since the basic statistics are computed from the whole available sample, their estimation can be considered reliable even with relatively short observation records.

Application of the methodology to Sicily island has revealed that it yields return period estimates that can generally be considered more reliable than those obtained by averaging the interarrival times between observed droughts in an historical sample. Also the comparison with results obtained from long synthetically generated samples indicates that the analytical expressions yield very similar return periods for a wide range of drought severities, which is indicative of the adequacy of the proposed methodology.

REFERENCES

- Bayazit, M. and Onoz, B. (2005) Probabilities and return periods of multisite droughts, *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, (50)4, 605
- Bonaccorso, B., Cancelliere, A., and Rossi, G. (2003). An analytical formulation of return period

- of drought severity. *Stoch. Environmental Research and Risk Assessment*, 17(1), 157–174.
- Frank, M. (1979), On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$, *Aequationes Math.*, (19), 194-226
- Gonzalez, J. and Valdes, J. (2003). Bivariate drought recurrence analysis using tree ring reconstructions. *Journal of Hydrologic Engineering*, 8(5), 247–258.
- Guen, O. (1983). A simplified semiempirical approach to probabilities of extreme hydrologic droughts. *Water Resources Research*, 19(2), 441–453.
- Johnson, N., Kotz, S., and Balakrishnan, N. (1994). *Continuous univariate distributions*, Vol. 1. Wiley & Sons, New York.
- Leppard, P. and Tallis, G. (1989). Algorithm AS249: Evaluation of the mean and covariance of the truncated multinormal distribution. *Applied Statistics*, 38(3), 543–553.
- Mood, A., Graybill, F., and Boes, D. (1974) *Introduction to the theory of statistics*. McGraw-Hill, New York
- Rossi, G. (1988) Droughts of sicily, in Yevjevich, V., Da Cunha, L. and Vlachos, E. (Eds), *Coping with drought*, Water Resources Publication, Littleton, CO, 244-258
- Rossi G., Benedini, M., Tsakiris, G., Giakoumakis, S. (1992). On regional drought estimation and analysis. *Water Resources Management*, vol. 6; p. 249-277
- Rossi, G. and Cancelliere, A. (2003) At site and regional identification by REDIM model, in Rossi, G. and Cancelliere, A. and Oweis, T. and Pereira, L.S. and Shatanawi, M. and Zairi, A. (Eds), *Tools for drought mitigation in Mediterranean regions*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 37-54
- Salas, J. (1993). Analysis and modeling of hydrologic time series (chap. 19). *Handbook of Hydrology*, D. Maidment, ed., McGraw Hill Book Co., N. York.
- Salas, J., Fu, C., Cancelliere, A., Dustin, D., Bode, D., Pineda, A., and Vincent, E. (2005). Characterizing the severity and risk of droughts of the Poudre river, Colorado. *Journal of Water Resources Planning and Management*, 131(5), 383–393.
- Santos, M.A.(1983) Regional droughts: a stochastic characterization, *J. of Hydrol.*, (66), 183-211
- Sen, Z. (1980) Statistical analysis of hydrologic critical droughts, *Journal of the Hydraulics Division*, (106) HY1, 99-104
- Shervish, M.J. (1984) Algorithm AS195: Multivariate normal probability with error bound, *Applied Statistics*, (33)3, 81-94
- Shiau, J. and Shen, H. (2001). Recurrence analysis of hydrologic droughts of differing severity. *Journal of Water Resources Planning and Management*, 127(1), 30–40.
- Tase, N., (1976) *Area-deficit-intensity characteristics of droughts*, Ph.D. Thesis, Colorado State University, Fort Collins, CO
- Yevjevich, V., (1967) An objective approach to definitions and investigations of continental hydrologic droughts, *Hydrology paper no. 23*, Colorado State University, Fort Collins, CO.