CHAPTER 5

STOCHASTIC FORECASTING OF DROUGHT INDICES

A. CANCELLIERE, G. DI MAURO, B. BONACCORSO AND G. ROSSI
Department of Civil and Environmental Engineering, University of Catania, Italy

Abstract: Unlike other natural disasters, drought events evolve slowly in time and their impacts generally span a long period of time. Such features do make possible a more effective mitigation of the most adverse effects, provided a timely monitoring of an incoming drought is available

Among the several proposed drought monitoring indices, the Standardized Precipitation Index (SPI) has found widespread application for describing and comparing droughts among different time periods and regions with different climatic conditions. However, limited efforts have been made to analyze the role of the SPI for drought forecasting

The aim of the chapter is to provide two methodologies for the seasonal forecasting of SPI. In the first methodology, the transition probabilities from a current drought condition to another in the future, and from a single value of current SPI to a drought class are derived as functions of the statistics of the underlying monthly precipitation process. The proposed analytical approach appears particularly valuable from a practical stand point in light of the difficulties of applying a frequency approach or a Markov chain approach, due to the limited number of transitions generally observed even on relatively long SPI records. In the second methodology, SPI forecasts at a generic time horizon *M* are analytically determined, in terms of conditional expectation, as a function of a finite number of past values of monthly precipitation. Forecasting accuracy is estimated through an expression of the Mean Square Error, which enables to derive confidence intervals of prediction. Validation of the derived expressions is carried out by comparing theoretical forecasts and observed SPI values. The methodologies have been applied to the series of SPI, based on monthly precipitation observed in Sicily over 40 rain gauges in the period 1921-2003. Results seem to confirm the reliability of the proposed methodologies, which therefore can find useful application within a drought monitoring system

Keywords: drought, SPI, stochastic techniques, transition probabilities, forecast

1. INTRODUCTION

Due to a slow evolution in time, drought is a phenomenon whose consequences take a significant amount of time with respect to its inception in order to be perceived by the socio-economic systems. Taking advantage of this feature, an effective mitigation of the most adverse drought impacts is possible, more than in the case of other extreme hydrological events such as floods, earthquakes, hurricanes, etc., provided a drought monitoring system, able to promptly warn of the onset of a

G. Rossi et al. (eds.), Methods and Tools for Drought Analysis and Management, 83–100.

drought and to follow its evolution in space and time, is in operation (Rossi, 2003). To this end, an accurate selection of indices for drought identification, providing a synthetic and objective description of drought conditions, represents a key point for the implementation of an efficient drought watch system.

Among the several proposed indices for drought monitoring, the Standardized Precipitation Index (SPI) has found widespread application (McKee et al., 1993; Heim, 2000; Wilhite et al., 2000; Rossi and Cancelliere, 2002). Guttman (1998) and Hayes et al. (1999) compared SPI with Palmer Drought Severity Index (PDSI) and concluded that the SPI has advantages of statistical consistency, and the ability to describe both short-term and long-term drought impacts through the different time scales of precipitation anomalies. Also, due to its intrinsic probabilistic nature, the SPI is the ideal candidate for carrying out drought risk analysis (Guttman, 1999). An evaluation of common indicators, according to six weighted evaluation criteria of performance (robustness, tractability, transparency, sophistication, extendability, and dimensionality), indicates strengths of the SPI and Deciles over the PDSI (Keyantash and Dracup, 2002).

Although most of the proposed indices have been developed with the intent to monitor current drought conditions, nevertheless some of them can be used to forecast the possible evolution of an ongoing drought, in order to adopt appropriate mitigation measures and drought policies for water resources management. Within this framework, Karl et al. (1986) assessed the amount of precipitation needed to restore normal conditions after a drought event, with reference to the Palmer Hydrologic Drought Index (PHDI). Cancelliere et al. (1996) proposed a procedure for short-middle term forecasting of the Palmer Index and tested its applicability to Mediterranean regions, by computing the probability that an ongoing drought will end in the following months. Other authors (Lohani et al., 1998) proposed a forecasting procedure of the Palmer index based on the first-order Markov chains, which enables one to forecast drought conditions for future months, based on the current drought class described by the PHDI values. Recently, Bordi et al. (2005) compared two stochastic techniques, namely an autoregressive model and a novel method called Gamma Highest Probability (GAHP), for forecasting SPI series at lag 1. The latter method forecasts precipitation of the next month as the mode of a Gamma distribution fitted to the observed precipitation series. They concluded that the GAHP performs better, especially in spring and summer months.

In the present chapter, a seasonal forecast of the SPI is addressed by means of stochastic techniques. In particular, transition probabilities from a drought class to another and from a current value of SPI to a drought class at different time horizons are analytically derived as a function of the statistical properties of the underlying monthly precipitation. The usefulness of such analytical derivation for estimating transition probabilities is evident in light of the fact that a Markov chain approach appears not adequate to model transitions of SPI values (Cancelliere et al., 2007). Moreover the analytical approach enables one to overcome the difficulties related to a frequency approach, whose reliability may be hindered by the generally limited sample size of the available precipitation series.

Also a model to evaluate SPI forecast on the basis of past values of precipitation has been developed. More specifically, analytical expressions of short-middle term forecasts of the SPI are derived as the expectation of future SPI values conditioned on a finite number of past observations. The accuracy of the model is evaluated in terms of the Mean Square Error (MSE) of prediction (Brockwell and Davis, 1996), which allows confidence intervals for forecasted values to be computed. Forecasting future values in terms of conditional expectation ensures that the corresponding forecasts will have minimum MSE. Validation of the model is carried out based on the historical series observed at 40 precipitation stations in Sicily (Italy), making use of a moving window scheme for parameters estimation.

2. THE STANDARDIZED PRECIPITATION INDEX

The SPI is able to take into account the different time scales at which the drought phenomenon occurs and, because of its standardization, is particularly suited to compare drought conditions among different time periods and regions with different climatic conditions (Bonaccorso et al., 2003).

The index is based on an equi-probability transformation of aggregated monthly precipitation into a standard normal variable. In practice, computation of the index requires fitting a probability distribution to aggregated monthly precipitation series (e.g. $k = 3$, 6, 12, 24 months, etc.), computing the non-exceedence probability related to such aggregated values and defining the corresponding standard normal quantile as the SPI. McKee et al. (1993) assumed an aggregated precipitation gamma distributed and used a maximum likelihood method to estimate the parameters of the distribution.

Although McKee et al. (1993) originally proposed a classification restricted only to drought periods, it has become customary to use the index to classify wet periods as well. Table 1 reports the climatic classification according to the SPI, provided by the National Drought Mitigation Center (NDMC, http://drought.unl. edu). Also, the probabilities ΔP, that the index lies within each class are listed. Since our present work focuses on forecasting drought conditions, the near normal and wet classes have been grouped into one class termed “Non-drought”.

Table 1. Wet and drought period classification according to the SPI index

Index value	Class	Probability	ΔP	
$SPI \geq 2.00$	Extremely wet	0.977-1.000	0.023	Non Drought
$1.50 \leq SPI < 2.00$	Very wet	0.933-0.977	0.044	Non Drought
$1.00 \leq SPI < 1.50$	Moderately wet	0.841-0.933	0.092	Non Drought
$-1.00 \leq SPI < 1.00$	Near normal	0.159-0.841	0.682	Non Drought
$-1.50 \leq SPI < -1.00$	Moderate drought	0.067-0.159	0.092	
$-2.00 \leq SPI < -1.50$	Severe drought	0.023-0.067	0.044	
$SPI < -2.00$	Extreme drought	0.000-0.023	0.023	

3. ANALYTICAL DERIVATION OF TRANSITION PROBABILITIES OF DROUGHT CLASSES

3.1 Transition Probabilities From A Drought Class to Another

Let $Z_{\nu,\tau}^{(k)}$ indicate the SPI value at year ν and month $\tau = 1, 2, \ldots 12$, for an aggregation time scale k of monthly precipitation. Also, let's indicate by C_i the generic drought class, for instance C_1 = Extreme, C_2 = Severe, C_3 = Moderate, C_4 = Non-drought. The probability that the SPI value after M months lies within a class C_j given that the SPI value at the current month lies within a class C_i, can be expressed as (Mood et al., 1974):

$$P\left[Z_{\nu,\tau+M}^{(k)} \in C_j \left| Z_{\nu,\tau}^{(k)} \in C_i \right.\right] = \frac{\int\int_{C_i,C_j} f_{Z_{\nu,\tau}^{(k)},Z_{\nu,\tau+M}^{(k)}}(t,s)\cdot dt \cdot ds}{\int_{C_i} f_{Z_{\nu,\tau}^{(k)}}(t)\cdot dt} \tag{1}$$

where $\mathrm{f}_{Z_{\nu,\tau}^{(k)},Z_{\nu,\tau+M}^{(k)}}(\cdot)$ is the joint density function of $Z_{\nu,\tau}^{(k)}$ and $Z_{\nu,\tau+M}^{(k)}$, $\mathrm{f}_{Z_{\nu,\tau}^{(k)}}(\cdot)$ is the marginal density function of $Z_{\nu,\tau}^{(k)}$, t and s are integration dummy variables, and the integrals are extended to the range of each drought class.

Since, by definition, SPI is marginally distributed as a standard normal variable, it is fair to assume the joint density function in eq. (1) to be bivariate normal, namely:

$$f_{Z_{\nu,\tau}^{(k)},Z_{\nu,\tau+M}^{(k)}}(t,s) = \frac{1}{2\pi\,|\Sigma|}\cdot\exp\left(-\frac{1}{2}\mathbf{X}^{\mathrm{T}}\Sigma^{-1}\mathbf{X}\right) \tag{2}$$

where $\mathrm{X} = [t, s]^{\mathrm{T}}$, and Σ represents the variance-covariance matrix:

$$\Sigma = \begin{bmatrix} 1 & \mathrm{cov}\left[Z_{\nu,\tau}^{(k)}, Z_{\nu,\tau+M}^{(k)}\right] \\ \mathrm{cov}\left[Z_{\nu,\tau}^{(k)}, Z_{\nu,\tau+M}^{(k)}\right] & 1 \end{bmatrix} \tag{3}$$

Thus, the computation of transition probabilities in eq. (1) requires the determination of the autocovariance at lag M of $Z_{\nu,\tau+M}^{(k)}$ namely $\mathrm{cov}\left[Z_{\nu,\tau}^{(k)}, Z_{\nu,\tau+M}^{(k)}\right]$. Such autocovariance can be efficiently estimated from an available sample of SPI series. Alternatively, under some hypothesis, it is possible to derive analytical expression of the autocovariance SPI, as a function of the statistics of the underlying precipitation. In general terms, such derivation is not straightforward, because of the equi-probability transformation underlying the SPI computation. However, under the hypothesis of monthly precipitation aggregated at time scale k normally distributed, the corresponding value of SPI can be computed through a simple standardization procedure:

$$Z_{\nu,\tau}^{(k)} = \frac{Y_{\nu,\tau}^{(k)} - \mu_\tau^{(k)}}{\sigma_\tau^{(k)}} \tag{4}$$

with $Y_{\nu,\tau}^{(k)} = \sum_{i=0}^{k-1} X_{\nu,\tau-i}$ aggregated precipitation at k months.

By assuming precipitation at month τ with mean μ_τ, the mean of the corresponding aggregated precipitation $Y_{\nu,\tau}^{(k)}$ will be:

$$\mu_\tau^{(k)} = \sum_{i=0}^{k-1} \mu_{\tau-i} \tag{5a}$$

Also, if σ_τ^2 is the variance of precipitation at month τ, under the hypothesis of precipitation values uncorrelated in time, the standard deviation of the corresponding aggregated precipitation $Y_{\nu,\tau}^{(k)}$ will be:

$$\sigma_\tau^{(k)} = \sqrt{\sum_{i=0}^{k-1} \sigma_{\tau-i}^2(x)} \tag{5b}$$

Substituting, eq. (4) becomes:

$$Z_{\nu,\tau}^{(k)} = \frac{\sum_{i=0}^{k-1} X_{\nu,\tau-i} - \sum_{i=0}^{k-1} \mu_{\tau-i}}{\sqrt{\sum_{i=0}^{k-1} \sigma_{\tau-i}^2}} \tag{6}$$

Therefore, the autocovariance can be expressed as:

$$\begin{aligned} cov\left[Z_{\nu,\tau+M}^{(k)}, Z_{\nu,\tau}^{(k)}\right] &= \frac{1}{\sqrt{\sum_{i=0}^{k-1} \sigma_{\tau+M-i}^2 \sum_{j=0}^{k-1} \sigma_{\tau-j}^2}} \cdot \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} cov\left[X_{\nu,\tau+M-j}, X_{\nu,\tau-i}\right] = \\ &= \frac{1}{\sqrt{\sum_{i=0}^{k-1} \sigma_{\tau+M-i}^2 \sum_{j=0}^{k-1} \sigma_{\tau-j}^2}} \cdot \sum_{i=0}^{k-M-1} \sigma_{\tau-i}^2 \end{aligned} \tag{7}$$

By substituting eq. (7) in the variance-covariance matrix Σ, it follows:

$$\Sigma = \begin{bmatrix} 1 & \frac{\sum_{i=0}^{k-M-1} \sigma_{\tau-i}^2}{\sqrt{\sum_{i=0}^{k-1} \sigma_{\tau+M-i}^2 \sum_{j=0}^{k-1} \sigma_{\tau-j}^2}} \\ \frac{\sum_{i=0}^{k-M-1} \sigma_{\tau-i}^2}{\sqrt{\sum_{i=0}^{k-1} \sigma_{\tau+M-i}^2 \sum_{j=o}^{k-1} \sigma_{\tau-j}^2}} & 1 \end{bmatrix} \tag{8}$$

Finally, by combining eq. (8) with eqs. (1) and (2), it is possible to express the SPI transition probabilities, in terms of the variances of monthly precipitation. Although the hypothesis of normality for aggregated monthly precipitation may appear restrictive, it is worth observing that it can be justified, especially for higher values of the aggregation time scale k, as a consequence of the central limit theorem.

3.2 Transition Probabilities from a Spi Value to a Drought Class

Eq. (1) enables to derive transition probabilities from one drought class to another in the future. One may expect however, that such probability should somewhat be affected by the value taken by SPI, and not just by the class it is laying in. For instance, transition toward Non drought class (SPI $\geq$ –1) from a moderate class (–1.5 $<$ SPI $<$ –1) should be more or less probable depending whether the observed SPI is closer to the class upper bound or to the class lower bound. Thus, it is of interest to derive an analytical expression in order to calculate the transition probabilities from a single value of SPI to a drought class in the future. The latter approach enables, with respect to the previous one, to take into account the influence of the specific past values on the forecasts.

Let's consider the random variable defined as:

$$Z = Z_{\nu,\tau+M} \left| Z_{\nu,\tau} \right.$$

According to the definition of SPI, the variables $Z_{\nu,\tau}$ e $Z_{\nu,\tau+M}$ are distributed as a normal distribution with zero mean and variance one; therefore, according to a well known statistical property of multivariate normal distributions, it is fair to assume Z as a conditional bivariate normal distribution, with mean and variance (Mood et al., 1974):

$$\mu_Z = \mu_{Z_{\nu,\tau+M}} + \left(\rho \cdot \sigma_{Z_{\nu,\tau+M}} \middle/ \sigma_{Z_{\nu,\tau}}\right) \left(z_{\nu,\tau} - \mu_{Z_{\nu,\tau}}\right) = \rho \cdot z_{\nu,\tau} \tag{9}$$

$$\sigma_Z^2 = \sigma_{Z_{\nu,\tau+M}}^2 \left(1 - \rho^2\right) = 1 - \rho^2 \tag{10}$$

where ρ can be calculated either directly by the observed series (correlation coefficient between SPI series at the present month τ and at month τ+M) or exploiting the previous equation (7) found for the autocovariance.

It follows that the transition probability to pass from a value of SPI Z_0 at the month τ, to a class of drought C_0, M months ahead, will be given by:

$$P[Z_{\nu,\tau+M} \in C_0 \left| Z_{\nu,\tau} = z_{\nu,\tau} \right.] = \int_{C_{0i}}^{C_{0s}} \frac{1}{\sqrt{2\pi}\sigma_Z} \cdot e^{-\frac{1}{2}\left(\frac{x-\rho z_{\nu,\tau}}{1-\rho^2}\right)^2} dx \tag{11}$$

where C_{oi} e C_{os} are respectively the lower and the upper bounds of the particular drought class C_0 considered.

Despite the apparent complexity of the proposed analytical approach for estimation of transition probabilities, it should be pointed out that the methodology yields results that can be considered in general more reliable than those obtained by alternative approaches, such as frequency analysis of observed transitions in an historical sample, or application of a Markov chain scheme (Cancelliere et al., 2007).

4. SPI FORECASTING

From a stochastic point of view, the problem of forecasting future values of a random variable is equivalent to the determination of the probability density function of future values conditioned by past observations. Once the conditional distribution is known, the forecast is usually defined as the expected value or a quantile of such distribution, and confidence intervals of the forecast values can be computed.

Let's consider a sequence of random variables Y_1, Y_2, …, Y_t. The interest here lies in determining a function f(Y_1, Y_2,…,Y_t) that forecasts a future value Y_{t+M} with minimum error. The latter is usually expressed as the Mean Square Error (MSE) of prediction, defined as (Brockwell and Davis, 1996):

$$\mathrm{MSE} = \mathrm{E}[(Y_{t+M} - \mathrm{f}(Y_1, Y_2, \ldots, Y_t))^2] \tag{12}$$

It can be shown, that the function f(·) that minimizes the MSE is the expected value of Y_{t+M} conditioned on $Y_1, Y_2, \cdots, Y_t$, i.e.:

$$\mathrm{f}(Y_1, Y_2, \ldots, Y_t) = \mathrm{E}[Y_{t+M}|Y_1, Y_2, \ldots, Y_t] \tag{13}$$

The above property enables to derive the "best" forecast (in MSE sense), provided the conditional expectation can be computed. Also, it may be worthwhile to note that if Y_{t+M} is independent of $Y_1, Y_2, \cdots, Y_t$, the best predictor of Y_{t+M} is its expected value, and furthermore, the MSE of prediction is just the variance of Y_{t+M}.

Besides MSE, a practical way of quantifying the accuracy of the forecast is by estimating the confidence interval of prediction, i.e. an interval that contains the future observed value with a fixed probability 1–α (e.g. 95 %). Obviously, the wider the interval, the less is the accuracy of the forecast and vice-versa. Confidence intervals of prediction for SPI can be estimated by capitalizing on the intrinsic normality of the index and by observing that, since the predictor is unbiased, its variance coincides with the MSE. Thus, the upper and lower confidence limits $Y_{1,2}$ of fixed probability 1– α can be computed as:

$$Y_{1,2} = \tilde{Y} \pm \sqrt{MSE} \cdot u_{1-\alpha/2} \tag{14}$$

where, for the sake of brevity, $\tilde{Y}$ represents the generic forecast and u(·) is the quantile of a standard normal (random) variable of probability 1–α/2.

Let's define $Z^{(k)}_{\nu,\tau}$ as the SPI value at year ν and month $\tau = 1, 2, \ldots, 12$, for an aggregation time scale k of monthly precipitation. The interest here lies in determining the conditional distribution (and its expectation) of a future value M months ahead $Z_1 = Z^{(k)}_{\nu,\tau+M}$, conditioned on the vector of θ past observations $\mathbf{Z}_2 = \left[\mathbf{Z}^{(k)}_{\nu,\tau}, Z^{(k)}_{\nu,\tau-1}, \ldots Z^{(k)}_{\nu,\tau-\theta+1}\right]^T$, i.e. the distribution of the following conditioned variable:

$$Z_{1|2} = Z^{(k)}_{\nu,\tau+M} \left| Z^{(k)}_{\nu,\tau}, Z^{(k)}_{\nu,\tau-1}, \ldots Z^{(k)}_{\nu,\tau-\theta+1} \right.$$

The conditional density function of $Z_{1|2}$ is defined as:

$$f_{Z_{1|2}}\left(z_{1|2}\right)=\frac{f_Z(z)}{f_{Z_2}(z_2)} \tag{15}$$

where $f_Z(\mathbf{z})$ is the multivariate pdf of the random vector $\mathbf{Z}=\left[\mathbf{Z}^{(k)}_{\nu,\tau+M}, Z^{(k)}_{\nu,\tau}, Z^{(k)}_{\nu,\tau-1}, \ldots Z^{(k)}_{\nu,\tau-\theta+1}\right]^T$ and $f_{Z2}(\mathbf{z}_2)$ is the joint density function of $\mathbf{Z}_2$.

As already mentioned, since, by definition, the SPI is marginally distributed as a standard normal variable, it is fair to assume that the vector $\mathbf{Z}$ is multivariate normal with zero mean and variance-covariance matrix Σ. The latter can be partitioned as follows:

$$\Sigma=\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix} \tag{16}$$

where:

$$\Sigma_{11}=Var\left[Z_1\right]=1 \tag{17}$$

$$\Sigma_{12}=\Sigma_{21}^{\mathrm{T}}=Cov\left[\mathbf{Z_1},\mathbf{Z_2}\right] \tag{18}$$

According to the property shown before, the conditional density function of $Z_{1|2}$ is normal itself. Besides, reminding that the SPI has zero mean, it can be shown that the expected value of $Z_{1|2}$ is (Kotz et al., 2000):

$$\mathrm{E}\left[Z_{1|2}\right]=\Sigma_{12}\cdot\Sigma_{22}^{-1}\cdot\mathbf{z} \tag{19}$$

where $\mathbf{z}=\left[z^{(k)}_{\nu,\tau}, z^{(k)}_{\nu,\tau-1}, \ldots z^{(k)}_{\nu,\tau-\theta+1}\right]^{\mathrm{T}}$ is the vector of past observations. Equation (19) yields the best forecast (in MSE sense) of a future value $\tilde{Z}^{(k)}_{\nu,\tau+M}$ given θ past observations $\mathbf{z}=\left[z^{(k)}_{\nu,\tau}, z^{(k)}_{\nu,\tau-1}, \ldots z^{(k)}_{\nu,\tau-\theta+1}\right]^T$. Furthermore, it is easy to see that the above forecast is unbiased and therefore the Mean Square Error of forecast coincides with the variance of $Z_{1|2}$, which is given by (Kotz et al. 2000):

$$MSE=Var\left[Z_{1|2}\right]=1-\Sigma_{12}\cdot\Sigma_{22}^{-1}\cdot\Sigma_{21} \tag{20}$$

Obviously, if $Z^{(k)}_{\nu,\tau+M}$ is uncorrelated with $Z^{(k)}_{\nu,\tau}, Z^{(k)}_{\nu,\tau-1}, \ldots, Z^{(k)}_{\nu,\tau-\theta}$, the best forecast of $Z^{(k)}_{\nu,\tau+M}$ is its expected value, and furthermore, the MSE of prediction is just the variance of $Z^{(k)}_{\nu,\tau+M}$.

If $\theta=1$, i.e. the forecast is based only on the present value, eq. (19) simplifies as:

$$\tilde{Z}^{(k)}_{\nu,\tau+M}=\mathrm{E}\left[Z^{(k)}_{\nu,\tau+M}|Z^{(k)}_{\nu,\tau}=z^{(k)}_{\nu,\tau}\right]=\rho\cdot z^{(k)}_{\nu,\tau} \tag{21}$$

while the Mean Square Error becomes:

$$MSE = \mathrm{Var}\left[Z^{(k)}_{\nu,\tau+M}\left|Z^{(k)}_{\nu,\tau} = z^{(k)}_{\nu,\tau}\right.\right] = 1 - \rho^2 \tag{22}$$

where ρ is the correlation coefficient between $Z^{(k)}_{\nu,\tau}$ and $Z^{(k)}_{\nu,\tau+M}$.

5. APPLICATIONS TO PRECIPITATION SERIES OBSERVED IN SICILY

The proposed methodologies have been applied to monthly precipitation observed from 1921 until 2003, for 40 precipitation stations in Sicily. The selected stations are included in the drought monitoring bulletin published on the website of the Sicilian Regional Hydrographic Office (Rossi and Cancelliere, 2002, http://www.uirsicilia.it).

5.1 Evaluation of Transition Probabilities

By fixing several combinations of forecasting time horizon M (months) and aggregation time scales k, for each station the transition probabilities from a class of SPI at month τ to another one at month $\tau + M$ have been computed by eq. (1), for every month. In order to compute the double integral in eq. (1), which expresses the normal joint cdf (cumulative distribution function), the algorithm MULNOR has been adopted (Schervish, 1984).

The mean values of transition probabilities corresponding to the 40 stations, for $M = 6$ and $k = 24$, are presented in Figure 1, as a function of the current month τ. In particular, transition probabilities related to different combinations of initial and final drought conditions (extreme Ex, severe Se, moderate Mo, non-drought N) have been considered. In order to show also the variability of transition probability among different stations, in the same plots, the limits related to ± 1 standard deviation are indicated by dashed lines. It can be observed that transition probabilities vary from one month to another, and for some transitions, also from one station to another (as indicated by the width of the limits). In particular, the mean probability value, indicated by the continuous line, of remaining in the extreme class (Ex/Ex) ranges from 60% (for February-March) to less than 25% (for August-September). The mean probability value of remaining in the non-drought class (N/N) presents a limited variability across the months, since it ranges from 95% (for February-March) to about 90% (for August-September). The transition probabilities from one class to another, especially in the lower triangular part of the matrix constituted by the graphics, are very low, at least for the considered time horizon, namely $M = 6$ months. In general, it can be concluded that starting from a wet month there is an higher probability to remain in the same drought class M months ahead (plots along the diagonal), than when starting from dry months, and conversely a lower probability to return to normal conditions.

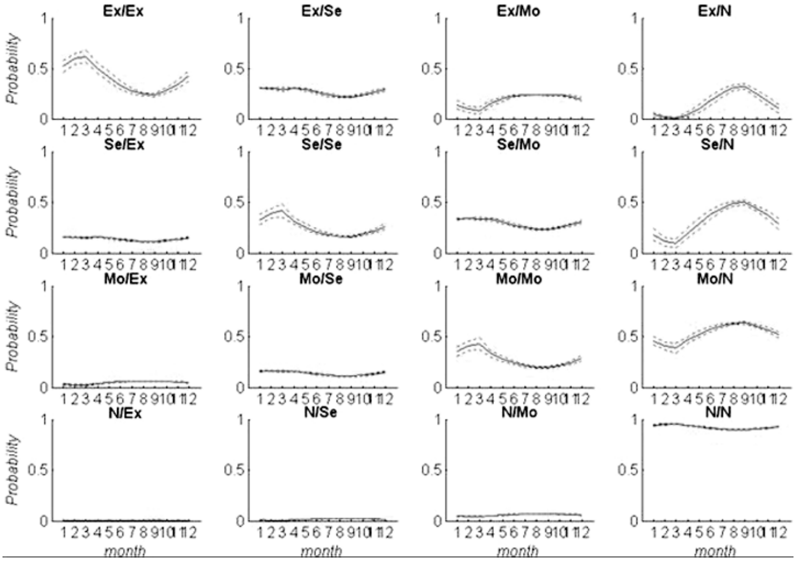

Figure 1. Mean of transition probabilities (continuous line) for $M = 6$ and $k = 24$ computed on 40 stations, and limits corresponding to ± 1 standard deviation (dashed line)

Such different behavior can be justified by considering that starting from a wet month and considering a 6 months time horizon, the occurrence probability of precipitation events, able to restore normal conditions, is very low. On the contrary, starting from a dry month, there is a high chance to observe values of precipitation such to modify drought conditions during the next 6 months.

In order to analyze the effects of the aggregation time scale k and of the forecasting time horizon M, transition probabilities computed for the 40 stations have been averaged and represented on a 3D-plot for a specific starting month and class, as a function of the final class and of the M values. In Figure 2 the case corresponding to extreme drought as starting class, August as starting month and

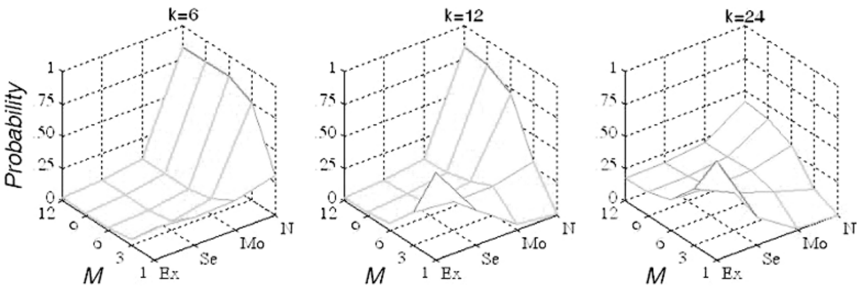

Figure 2. Mean transition probabilities (computed on the 40 stations) as a function of forecasting time horizon M (starting class: extreme drought (Ex), starting month: August)

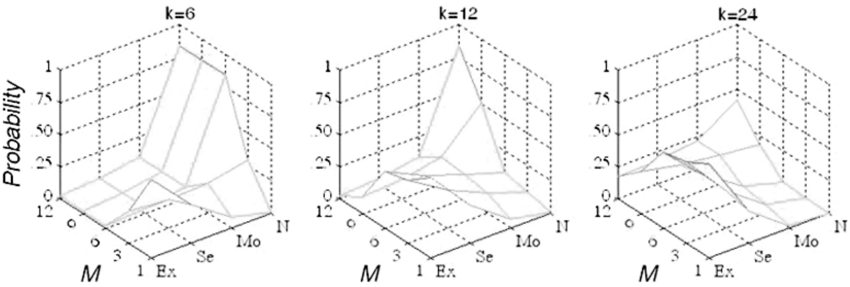

Figure 3. Mean transition probabilities (computed on the 40 stations) as a function of forecasting time horizon M (starting class: extreme drought (Ex), starting month: February)

$k = 6$, 12 and 24 are illustrated. In Figure 3 similar 3D-plots for February as starting month are shown.

It can be seen that probabilities to remain in the same class generally decrease as the forecasting time horizon M increases, while transition probabilities to non-drought condition show an opposite behaviour. Further, as the aggregation time scale k increases, the probabilities of remaining in the same class increase, whereas transition probabilities to return to non-drought condition generally decrease. Finally, by comparing transition probabilities related to August to those corresponding to February, it can be inferred that it is generally easier to recover from drought conditions starting from August than starting from February.

When the interest turns to the computation of transition probabilities from one value of current SPI to a drought class in the future, equation (11) is applied. In particular, by fixing the starting month τ, the aggregation time scale k (3, 6, 9, 12 and 24 months), the time horizon M (1, 3, 6, 9, 12 and 24 months) and the drought class C_0 in the future, it is possible to derive the corresponding transition probabilities for different current SPI values z_τ.

Figure 4 and Figure 5 illustrate the mean transition probabilities to different drought classes (Ex, Se, Mo, N) as a function of current SPI value, computed based on the precipitation series of the considered 40 stations. For the sake of clarity, only the portion of the plots corresponding to $-3 \leq Z_\tau \leq 0$ have been plotted. Besides, each application is performed with reference to two starting months, namely: February and August.

More specifically, in Figure 4 transition probabilities parametrized with respect to different time horizons M (1, 3, 6 months) are reported for an aggregation time scale k=9 months. Figure 5 shows the transition probabilities corresponding to the aggregation time scale k (9, 12, 24 months), for a fixed time horizon M=6 months.

From both figures it is possible to infer similar results to those previously obtained for the transition probability from one class to another. In fact, the probability to remain in the same class decreases as the forecasting time horizon M increases, while transition probabilities to non-drought condition show an opposite behaviour. For instance, from the first plots at the left hand side in Figure 4, representing the transitions to a drought class Extreme, it can be clearly observed that for a current extremely

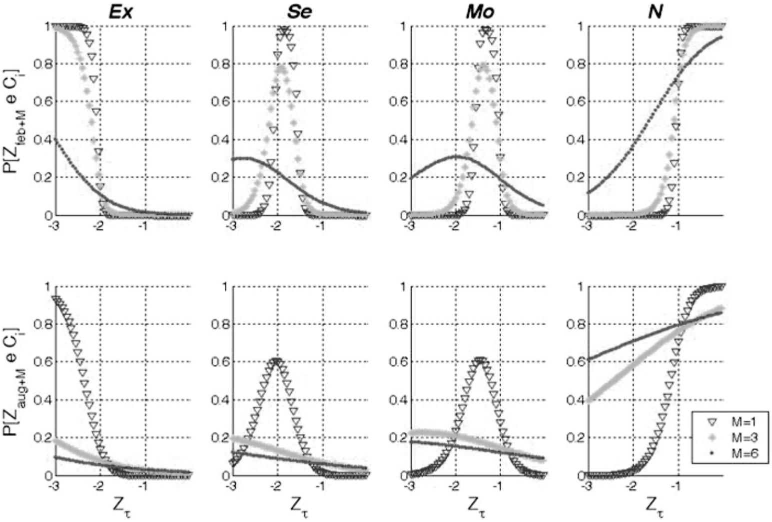

Figure 4. Mean transition probabilities (computed on the 40 stations) from the starting SPI Z_τ to a drought class C_i 1 month ahead (column), as a function of forecasting horizon time *M* (starting months: February, August – rows; aggregation scale *k*=9 months)

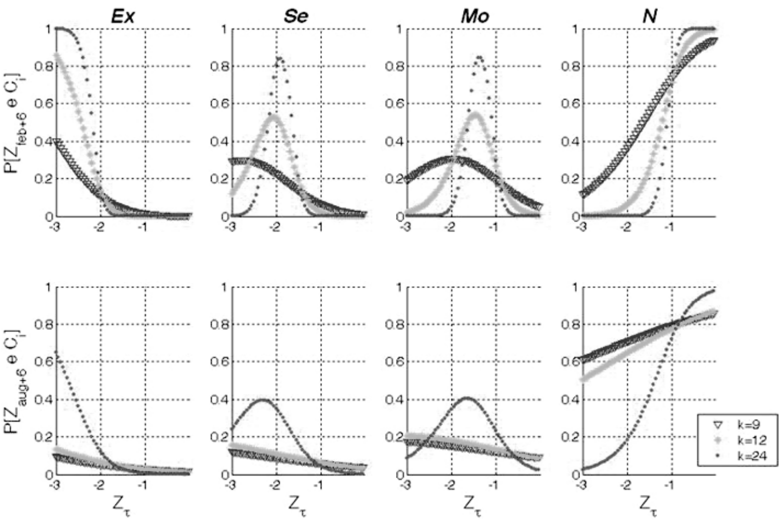

Figure 5. Mean transition probabilities (computed on the 40 stations) from the starting SPI Z_τ to a drought class C_i (column), as a function of aggregation scale *k* (starting months: February, August–rows)

dry condition, namely SPI< -2, the probabilities decrease as M increases, while for all the other hydrometeorological conditions, namely SPI> -2, the probabilities increase as M increases, even if in this case such values are very low (less than ~ 0.20).

On the contrary, from the last plots at the right hand side, representing the transitions to a Normal condition, it can observed that, as M increases, the probabilities increase for SPI< -1 and decrease for SPI> -1.

Figure 5 shows the influence of the aggregation time scale k on the transition probabilities for $M = 6$ month. In particular, as it was expected, it can be noticed a general increase of the probabilities to remain in the same class and a decrease of transition probabilities to different classes, as k increases.

From both figures, it is worth pointing out that transition probabilities are almost independent of M and k, for current values of SPI close to the lower limits of the future classes of transition.

Another important result consists in the high variability of the transition probabilities within the same starting class. In Figure 4, for example, the probability to pass to a Non drought condition 1 month ahead with aggregation scale 9 months starting from August (2nd row, 4th column, line marked by triangles) increases from the value 20% for $Z_\tau = -1.5$ to the value 70% for $Z_\tau = -1$, a considerable difference within the same starting Moderate class. Else by considering Figure 5, the probability to remain in Severe conditions 6 months ahead with aggregation scale 24 months starting from February (1st row, 2nd column, line marked by circles) decreases from the value 80% for $Z_\tau = -2$ to the value 20% for $Z_\tau = -1.5$.

In order to emphasize this feature, in Figure 6 transition probabilities for two different values of SPI inside the same drought class, namely $Z_\tau = -2.95$ and -2.05,

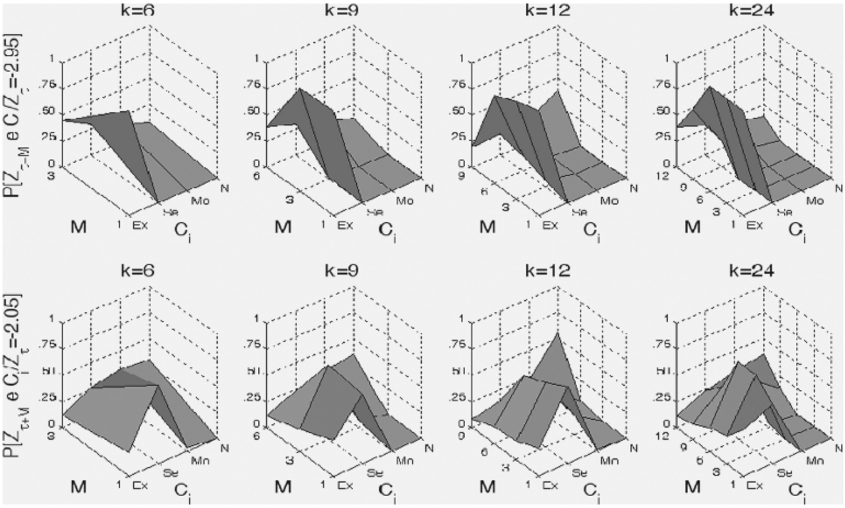

Figure 6. Mean transition probabilities (computed on the 40 stations) as a function of forecasting time horizon M and drought class C_i (starting SPI: -2.95 on the first row, -2.05 on the second row; starting month: February)

are shown. In particular, the three-dimensional graphics represent the variability of transition probabilities (z-axis) versus the time horizon M (x-axis) and the drought class C_i (y-axis) for different aggregation time scale k and by fixing February as starting month. For instance, for an aggregation time scale $k = 12$ months and a horizon time $M = 9$ months, it is possible to note how the probability to pass to non-drought condition increases from the value 35% for $Z_\tau = -2.95$ to the value 60% for $Z_\tau = -2.05$, as it was easily expected, since the second condition is less extreme than the other. Same features can be obtained for other time scales and time horizons.

5.2 SPI Forecast

The forecasting models proposed has been applied by considering several aggregation time scales, $k = 6$, 9, 12 and 24 months, as well as different forecasting time horizons $M = 3$, 6, 9 and 12 months.

First, theoretical MSE values (see eq. (20)) have been computed for all the 40 stations. In particular, for each available series, SPI values have been calculated and the variance- covariance matrix Σ (eq. (16)) has been estimated by means of its sample counterpart for fixed τ, k θ and M.

Figure 7 illustrates the boxplots of monthly MSE values for different initial month as a function of the number of past observations θ adopted to compute the

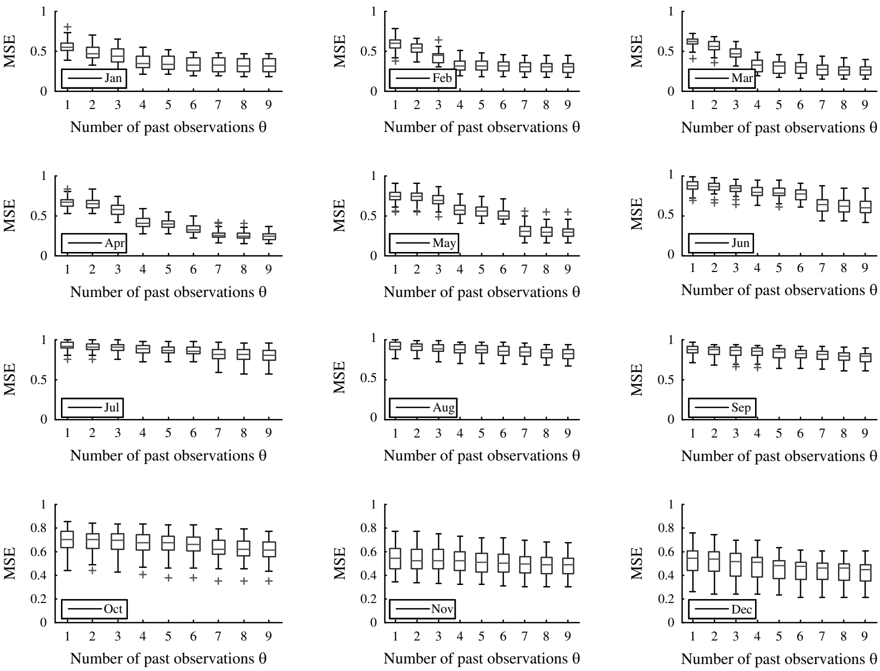

Figure 7. Boxplots of MSE of forecast as a function of the number of past observations θ for different initial month (by considering k=6 and M=3)

forecast. MSE have been computed considering an aggregation time scale $k = 6$ and a forecasting time horizon $M = 3$ months. The overall height of each boxplot indicates the variability of MSE among the different stations. As expected, the performance of the forecasting model generally improves (lower MSE's) as a larger number of past observations is adopted, although the decrease of MSE exhibits a different behaviour according to the considered initial month.

This reflects the intrinsic seasonality in the autocovariance function of SPI, as pointed out by Cancelliere et al., (2005). From the plots, it can be inferred that no significant improvement in the performance of the forecast are obtained for $\theta > 7$ in all cases, and therefore such a value seems to be preferable. Results related to different combinations of k and M (here not shown for brevity), lead to similar conclusions and therefore $\theta = 7$ has been adopted in what follows.

Then, the forecasting model has been validated by comparing observed and forecasted SPI computed by eq. (19) during the period 1921–2003. As an example, Figure 8 shows the comparison between observed and forecasted values SPI for one of the 40 stations, namely Caltanissetta, for different combinations of the time

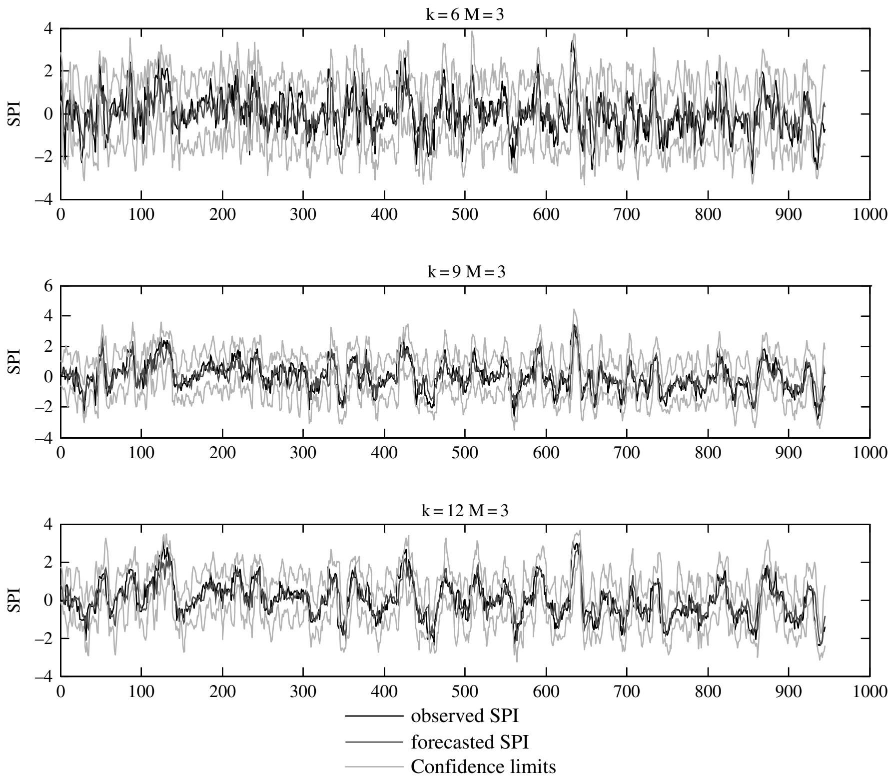

Figure 8. Comparison between observed and forecasted SPI for Caltanissetta station and related 95% confidence limits for different aggregation time scales k

scales k and $M = 3$ months. On the same plots, 95% confidence intervals, estimated by means of eq. (14) are also shown.

From the figure it can be inferred a fairly good agreement between observed and forecasted SPI values as is also evident from the fact that almost all of the observed values lie within the confidence limits. As expected, the accuracy of the forecast decreases as larger k are adopted with respect to the fixed M.

6. CONCLUSIONS

Drought monitoring and forecasting are essential tools for implementing appropriate mitigation measures in order to reduce negative drought impacts. Knowledge of transition probabilities from a drought class to another, for a given site or region, as well as the availability of forecasts of drought indices, and of the related confidence intervals, can help to improve the decision making process for drought mitigation, since appropriate measures can be selected based on the risk associated with the possible evolution of a current drought condition.

In the chapter, stochastic methodologies (i) to compute drought transition probabilities (from one class to another and from a current SPI value to a drought class), and (ii) to forecast future SPI values on the basis of past precipitation, are presented. Analytical expressions of the Mean Square Error of prediction are also presented, which enables one to derive confidence intervals for the forecasted values.

The effects of the aggregation time scale k and of the forecasting time horizon M on transition probabilities have been analyzed for 40 stations, by fixing the starting month and class, as a function of the final class and of the M values. Also, an analysis done by fixing the starting value of SPI, shows the importance of considering a current value of SPI rather than a class in the calculation of transition probabilities. In fact the transition probabilities are strongly affected by the value taken by the SPI within the same starting class. Thus taking into account the starting value rather than the starting class leads to a significant improvement in the identification of most probable drought classes in the future.

The proposed methods to compute transition probabilities is particularly valuable from a practical standpoint, in light of the difficulties of applying a frequency approach due to the limited number of transitions generally observed even on relatively long SPI records. In fact, the number of observed transitions is generally not sufficient to compute reliable frequency estimates. Furthermore, the lack of observed transitions in some cases would lead to the misleading conclusion that such transitions have zero occurrence probability, which is obviously not correct. Application of the analytical approach on the other hand, allows one to always estimate transition probabilities, even from relatively short records, since the whole available precipitation series, and not just the few observed transitions, are utilized. Regarding the applicability of Markov chain hypothesis to model transitions of SPI values from one drought class to another, Cancelliere et al. (2007) have also demonstrated that such an assumption may not be valid in general.

Validation of the forecasting model has been carried out by comparing SPI values computed on precipitation observed in the same 40 stations in Sicily and the corresponding forecasts. The results show a fairly good agreement between observations and forecasts, as it has also been confirmed by the values of some performance indices, which indicates the suitability of the model for short- medium term forecast of drought conditions.

Ongoing research is being carried out to improve the forecasting capabilities of the model, by taking into account exogenous covariates such as large scale climatic indices.

ACKNOWLEDGEMENTS

The present research has been developed with the partial financial support of the EU Projects "Mediterranean drought preparedness and mitigation planning" (MEDROPLAN, EuropeAid MEDA, contract no. ME8/AIDCO/2001/0515/59770-P027), of the Italian project MIUR PRIN 2005 “Previsione e mitigazione della siccità” and of the research contract INTERREG III B MEDOCC project SEDEMED II.

REFERENCES

Bonaccorso, B., Bordi, I., Cancelliere, A., Rossi, G. and Sutera, A. (2003) Spatial variability of drought: an analysis of SPI in Sicily. *Water Resource Management* 17, pp. 273–296.

Bordi, I., Fraedrich, K., Petitta, M. and Sutera, A. (2005) Methods for predicting drought occurrences. *Proc. of the 6th International Conference of the European Water Resources Association*, Menton, France, 7–10 September 2005.

Brockwell, P.J. and Davis, R.A. (1996) *Introduction to Time Series and Forecasting.* Springer-Verlag ed., New York, pp. 1–422.

Cancelliere, A., Di Mauro, G., Bonaccorso, B. and Rossi, G. (2007). Drought forecasting using the Standardized Precipitation Index. *Water Resource Management* 21(5), pp. 801–819.

Cancelliere, A., Di Mauro, G., Bonaccorso, B. and Rossi, G. (2005) Stochastic forecasting of Standardized Precipitation Index. *Proc. of XXXI IAHR Congress “Water Engineering for the future: Choice and Challenges”*, Seoul, Korea, 11–16 September 2005, pp. 3252–3260.

Cancelliere, A., Rossi, G. and Ancarani, A. (1996) Use of Palmer Index as drought indicator in Mediterranean regions. *Proc. IAHR Congress “From Flood to Drought”*, Sun City, South Africa, August 5–7, 1996, pp. S4.12. 1–25.

Guttman, N.B. (1998) Comparing the Palmer Drought Severity Index and the Standardized Precipitation Index. *Journal of the American Water Resources Association* 34(1), pp. 113–121.

Guttman, N.B. (1999) Accepting the Standardized Precipitation Index: a calculating algorithm. *Journal of the American Water Resources Association* 35(2), pp. 311-323.

Hayes, M.J., Svoboda, M., Wilhite, D.A. and Vanyarkho, O. (1999) Monitoring the 1996 drought using the SPI. *Bulletin of the American Meteorological Society* 80, pp. 429–438.

Heim, R.R. (2000) Drought indices: a review. In: D.A. Wilhite (Ed.), *Drought: A Global Assessment*, Routledge, London and New York.

Karl, T., Quinlan, F. and Ezell, D.S. (1986) Drought termination and amelioration: its climatological probability. *Journal of Climate and Applied Meteorology* 26, pp. 1198–1209.

Keyantash, J. and Dracup, J.A. (2002) The quantification of drought: an evaluation of drought indices. *Bulletin of the American Meteorological Society* 83(8), pp. 1167–1180.

Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000) *Continuous multivariate distributions*, Volume 1. John Wiley and Sons, Canada.

Lohani, V.K., Loganathan, G.V. and Mostaghimi, S. (1998) Long-term analysis and short-term forecasting of dry spells by Palmer Drought Severity Index. *Nord. Hydrol.* 29(1), pp. 21–40.

McKee, T.B., Doesken, N. J. and Kleist, J. (1993) The relationship of drought frequency and duration to time scales. *Proc. 8th Conference on Applied Climatology,* Anaheim, California, pp. 179–184.

Mood, A. M., Graybill, F. A. and Boes (1974) *Introduction to the Theory of Statistics.* McGraw-Hill, New York.

Rossi, G., Cancelliere, A. (2002) Early warning of drought: development of a drought bullettin for Sicily. *Proc. 2nd International Conference "New trends in Water and Environmental Engineering for Safety and Life: Eco-compatible Solutions for Aquatic Environments"*, Capri, Italy, June 24–28, 2002, pp. 1–12.

Rossi, G. (2003) Requisites for a drought watch system. In: G. Rossi et al. (Eds.), *Tools for Drought Mitigation in Mediterranean Regions*, Kluwer Academic Publishing, Dordrecht, pp. 147–157.

Schervish, M.J. (1984) Algorithm AS 195. Multivariate Normal Probabilities with Error Bound. *Applied Statistics* 33(1), pp. 81–94.

Wilhite, D.A., Hayes, M. J. and Svoboda, M.D. (2000) Drought monitoring and assessment: status and trends in the United States. In: J.V.Vogt, F. Somma (Eds.), *Drought and Drought Mitigation in Europe*, Kluwer Academic Publishers, Dordrecht, Boston, London, pp. 149–160.