Comparison between Parametric and Nonparametric Approaches for the Calculation of Two Drought Indices: SPI and SSI

Tatiana Soľáková[1]; Carlo De Michele[2]; and Renata Vezzoli, Ph.D.[3]

Introduction

Drought is generally viewed as a sustained and regionally extensive occurrence of below-average natural water availability, either in the form of precipitation, river runoff, or groundwater (Beran and Rodier 1985). Droughts can be classified as (1) hydrological droughts if these involve periods of below-normal flow, and depleted reservoir storage; (2) meteorological droughts if these involve periods of below-normal precipitation; (3) agricultural droughts if the soil moisture is not sufficient to support crop growth; and (4) socioeconomical droughts when the low water supply affects society's productive and consumptive activities (Dracup et al. 1980; Beran and Rodier 1985; Wilhite and Glantz 1987).

Drought is ranked first among other natural hazards (Mishra and Singh 2010). Drought differs from other natural hazards in several specifications (Wilhite 2000). First, the lack of a precise definition of drought is the main obstacle to its investigation. Second, identifying the onset and the end of a drought is difficult. Third, the spatial extension of a drought is difficult to determine; the impacts of droughts are nonstructural, and can interest large geographical areas more than other natural hazards. Fourth, human activities can directly trigger a drought, unlike other natural hazards, with exacerbating factors such as overfarming, excessive irrigation, deforestation, and overexploiting available water (Wilhite 2000).

In order to monitor the dynamics of droughts, several indices have been developed, accommodating the different typologies of droughts. The main indices include the Palmer drought severity index (Palmer 1965; Alley 1984), deciles (Gibbs and Maher 1967), the crop moisture index (Palmer 1968), the standardized precipitation index (McKee et al. 1993), the reconnaissance drought index (Tsakiris 2004), the Palmer hydrological drought index (Palmer 1965), the surface water supply index (Shafer and Dezman 1982; Tsakiris et al. 2007), and the standardized runoff index (Shukla and Wood 2008). Comprehensive lists of drought indices are reported in the following review papers: Heim (2002), Mishra and Singh (2010), and Dai (2011).

The standardized precipitation index (SPI) is the most commonly used drought index. It has been introduced by McKee et al. (1993) for the quantification of meteorological droughts. Starting from time series of monthly precipitation, time series averaged at a moving window of 3, 6, 12, and 24 months are derived. Then the gamma distribution is commonly used to accommodate the statistical variability of each time series. The application of the inverse of standard normal distribution to the gamma distribution gives a standard normal random variable, which represents the SPI (McKee et al. 1993). According to the time scale, the index is indicated as SPI1 for a time scale of 1 month, SPI3 for 3 months, and so on.

Guttman (1999) has applied six different probability distributions to long-term datasets and examined the impact of using the different distributions on the calculation of SPI. Haynes (2000) has investigated the effect of the length of precipitation record on the determination of the probability distributions. Guttman (1999) stated that at least 40–60 years are necessary for the stability of the parameters estimation in the central part of the distributions, and that about 70–80 years of records are necessary for achieving the stability in the tails of the precipitation distributions.

[1]Ph.D. Student, Institute of Environmental Engineering, Civil Engineering Faculty, Technical Univ. of Košice, Vysokoškolská 4, Košice, Slovakia.

[2]Associate Professor, Dept. of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo Da Vinci 32, 20133 Milano, Italy (corresponding author). E-mail: carlo.demichele@polimi.it

[3]Impacts on Soil and Coasts Division, Centro Euro-Mediterraneo sui Cambiamenti Climatici (ISC-CMCC), Via Maiorise s.n.c., 81043 Capua (CE), Italy.

Wu et al. (2005) noted that if SPI is computed from different sample sizes, then inconsistent results could be obtained. This arises as a consequence of temporal changes in the parameters of probability distribution. Consequently, it is recommended to test different probability distributions when the length of data records is short (Mishra and Singh 2010).

Despite all the limitations, the SPI is widely used to es-tablish extreme dryness, or wetness, in precipitation time series (e.g., Vicente-Serrano 2006; Lopez-Moreno and Vicente-Serrano 2008; Mo and Schemm 2008; Bordi et al. 2009; Bothe et al. 2010; Santos et al. 2010; Zhu et al. 2011; Sienz et al. 2012). In anal-ogy to the SPI, recently two hydrological drought indices have been developed: the standardized streamflow index (SSI) (Nalbantis 2008), and the standardized runoff index (SRI). The first index considers as a variable of interest the monthly streamflow, while the second one considers the monthly runoff (Shukla and Wood 2008). Also, SSI and SRI can be calculated at different time scales (1, 3, 6, 12, and 24 months) according to the needs of the water re-source management. Drought indices are fundamental tools in assessing the effects of a drought and defining drought parameters including duration, severity, and interarrival time between two successive droughts. Understanding of the historical behavior of droughts is important in planning management of water resources and predicting and avoiding the potential damage of future droughts (Mishra and Singh 2010). Several authors (Kroll and Vogel 2002; Yue and Wang 2004; McMahon et al. 2007; Nalbantis 2008; Shukla and Wood 2008; Madadgar and Moradkhani 2011; Vicente-Serrano et al. 2012) have stated that the log-normal distribution is the better choice to represent hydrological data.

The problem of selecting the appropriate probability distribution to fit monthly streamflow data has not been investigated in depth. Vicente-Serrano et al. (2012) have considered two different parametric methods to calculate the distribution of monthly streamflow, and evaluate the SSI, relative to the Ebro River in Spain. In the first method, the same probability distribution has been used for each month of the year, but with different parameters month by month. In the second method, the probability distribution for each month has been selected among a pool of six three-parameter distributions (namely, log-normal, Pearson type III, log-logistic, general extreme value, generalized Pareto, and Weibull) using the nonparametric Kolmogorov-Smirnov test, and the minimal orthogonal distance criteria.

Motivation

In literature, parametric approaches have been preferred to nonparametric methods for the calculation of drought indices. Although nonparametric methods are more flexible in application, because they do not require the specification of the probability distribution, problems of extrapolation have favored the application of parametric methods. Differences between parametric and nonparametric calculations of drought indices are due to the fact that in the former method a probability distribution for the variable of interest must be assumed. In literature no studies quantify the differences between nonparametric and parametric methods on the evaluation of droughts.

Here SPI and SSI indices are evaluated at different time scales (1, 3, and 12 months) using both nonparametric and parametric approaches, and the mean percentage difference between the two approaches is calculated. The calculation is made considering the monthly time series at Roma Collegio Romano for precipitation, and at the Ripetta gauging station on the Tevere River for streamflow, both in the period 1921–2000. The parametric approach has been implemented considering two possibilities: (1) a unique probability distribution for the variable of interest (gamma for precipitation, and log-normal for streamflow) with parameters depending on the month of the year, and (2) the probability distribution, depending on the month of the year, selected from a pool of continuous distributions (Burr, Erlang, gamma, generalized gamma, inverse Gaussian, log-normal, Pearson, Weibull, and generalized extreme value) using the Kolmogorov-Smirnov goodness-of-fit test. In both cases, the parameters of probability distributions are calculated using the maximum likelihood method. The nonparametric approach has been implemented using the Weibull plotting position to calculate the cumulative frequency.

Case Study

The case study considers two monthly time series in central Italy: Roma Collegio Romano (41°53′54″ N, 12°28′46″ E) for precipitation, and Tevere River Basin at Ripetta (41°54′20″ N, 12°28′32″ E) for streamflow, relative to the period 1921–2000. Data are provided by the Regional Agency of Environmental Protection of the Lazio Region. The precipitation dataset presents some years missing: 1928, 1934, 1937, 1942, and 1946. The streamflow dataset has only two months missing: August and September 1992. Note that authors have limited the period of observation to the 80 years 1921–2000, even if the time series are longer. This is to avoid possible problems of nonhomogeneity in the two time series (Wijngaard et al. 2003). Pettit's test (Pettit 1979) for homogeneity has been applied to data and both series can be considered homogeneous over the period 1921–2000, with a p-value equal to 0.959 for precipitation, and equal to 0.972 for streamflow, always greater than the canonical significance levels (0.05, 0.1).

Materials and Methods

The variable $X_d = \bar{P}_d, \bar{Q}_d$ indicates, respectively, the mean precipitation, or the average streamflow, in the time window d. The series of mean values, over the time window d, are calculated as $X_d(t) = (1/d) \sum_{\tau=t-d+1}^{t} X(\tau)$.

Here three time scales are considered: $d = 1$, 3, and 12 months. Let $F(x)$ be the cumulative distribution of X_d. The standardized drought index Z_d (here SPI or SSI according to X_d) is defined as the inverse of cumulative standard normal distribution (Φ^{-1}) of $F(x)$

$$Z_d = \Phi^{-1}[F(X_d)] \quad (1)$$

According to Madadgar and Moradkhani (2011), as the threshold level of drought $Z_d = -1$ is used. Thus, if $Z_d \leq -1$, then drought conditions prevail, while if $Z_d > -1$, normal, or wet, conditions prevail. Consequently, it is possible to define the duration of a drought, D, as the period during which the drought index is continuously below the threshold level. The interarrival time, T, is defined as the continuous period the index persists above the threshold. The severity of a drought is defined as $S = -\sum_{i=1}^{D} Z_d(i)$, the cumulative value of the index in the drought period.

The central problem in the calculation of Z_d is the determination of its distribution function, $F(x)$. It can be calculated using both parametric and nonparametric approaches. In the parametric approaches, it is necessary to select a probability distribution for the variable X_d, while in the nonparametric approach, the empirical cumulative frequency gives an estimation of $F(x)$. Here the parametric approach has been implemented considering two possibilities for the

choice of the probability distribution. The first parametric approach, denominated as P1, considers a unique probability distribution for the variable of interest [gamma for $\bar{P}_d(t)$, and log-normal for $\bar{Q}_d(t)$] with parameters depending on the month (t) of the year. The second parametric approach, denominated as P2, considers the probability distribution of the variable of interest depending on the month of the year, and is selected from a pool of continuous distributions using the Kolmogorov-Smirnov goodness-of-fit test. The probability distributions considered are nine: Burr, Erlang, gamma, generalized gamma, inverse Gaussian, log-normal, Pearson, Weibull, and generalized extreme value. The parameters of the probability distributions are determined using the maximum likeli-hood method (Kottegoda and Rosso 1997). Note that the precipitation is characterized by a probability mass in zero, P_0; thus, its cumulative distribution function is a mixed distribution:

$F(x) = P_0 + (1 - P_0)F(x|x > 0)$, where $F(x|x > 0)$ is the distribution of for strictly positive values of the variable (Blain 2012).

An estimate of P_0 can be obtained as $P_0 = N_0/(N + 1)$, where N_0 is the number of zeros in a dataset composed by N observations.

The nonparametric approach, denominated as NP, has been implemented calculating the cumulative frequency through the Weibull plotting position formula (Kottegoda and Rosso 1997). Let $x_1, x_2, \ldots, x_N$ be the sequence of N observations of the variable X_d, with the relative order statistics $x_{(1)} \le x_{(2)} \le \cdots \le x_{(N)}$. For the ith order statistics $x_{(i)}$ the cumulative frequency is $F_N(x_{(i)}) = i/(N + 1)$.

The Kolmogorov-Smirnov (K-S) test has been used here to measure the goodness of fit between a sample and a probability distribution (Kottegoda and Rosso 1997). The Kolmogorov-Smirnov statistic, denominated with D_N, is the maximum distance, in absolute value, between the empirical distribution function of the sample and the cumulative distribution function. Mathematically speaking it is calculated as $D_N = \sup_x|F_N(x) - F(x)|$. By the Glivenko-Cantelli theorem, if the sample comes from distribution $F(x)$, then D_N converges to 0 almost surely. The standard significance level used in this paper is $\alpha = 5\%$. The critical value of D_N is reported in standard probability books (e.g., Kottegoda and Rosso 1997) under the hypothesis that the parameters of the probability distribution are known. When the parameters are estimated from data, as in this case, then the critical value of D_N must be calculated via Monte Carlo simulation over a pool of 10,000 simulations as indicated by Keutelian (1991).

Following Vicente-Serrano et al. (2012), among a pool of probability distributions, the one characterized by the smallest value of the Kolmogorov-Smirnov statistic has been selected.

In order to show the differences between the nonparametric method (NP) and the parametric ones (P1 and P2), as well as between P1 and P2, in the calculation of the two drought indices, the authors evaluate for each of the three time scales, for each couple of methods and for each variable, the percentage difference in absolute value, Δ, in terms of duration D, severity S, and interarrival time T. For example, $\Delta_D^{NP-P1}(j) = 100 \cdot [|D_{NP}(j) - D_{P1}(j)|]/D_{NP}(j)$ represents the percentage difference, in absolute value, between NP and P1 in terms of drought duration D, relatively to the event j. Note that Δ takes into account the differences in terms of duration, but also in terms of temporal allocation in the year. Then, to summarize the differences, the mean value, $\bar{\Delta}$, is evaluated.

Results

First the monthly time series of precipitation at Roma Collegio Romano were considered, and the average precipitation at the 3-month

and 12-month time scale was calculated. Then for each month (t) the authors (1) evaluated the probability mass in zero, P_0; (2) estimated the parameters of the cumulative distribution functions considered, both for P1 and P2 approaches, using the maximum likelihood method; (3) checked that the domain of the variable includes only nonnegative values, making, in contrary cases, a lower truncation of the distribution to avoid negative values; and (4) checked the agreement between each distribution and data using the Kolmogorov-Smirnov goodness-of-fit test as indicated by Keutelian (1991). In particular for the P2 approach, the probability distribution with the smallest value of the K-S statistic was selected. As an example, in Table 1 only the parameters' values of the selected distributions, both for P1 and P2, were reported, relative to the time series of monthly precipitation. In particular, the second column reports the estimates of the parameters of gamma 3P, considered in the P1 approach. Columns 3–5 of Table 1 give, for the P2 approach, respectively, the distribution for each month, se-lected with the K-S test, its cumulative distribution function, and the parameters' estimates. The last column of Table 1 reports the monthly estimate of P_0. For two months (July and August), P_0 is significantly different from zero, for six months (January, February, April, May, October, and November) P_0 is zero, and for the remaining months (March, June, September, and December), P_0 is slightly greater than zero. Note also that the value of the shift parameter, γ, of gamma 3P, is equal to zero for values of P_0 different from zero, while it is different from zero for the months where $P_0 = 0$.

Note that for January, and the P2 approach, the selected distribution is generalized extreme value, with parameters $\mu = 49.53$ mm, $\sigma = 35.16$ mm, and $k = -0.11$. In this case, the domain of the variable X is upper bounded: $x < \mu - \sigma/k = 369$ mm. To consider only nonnegative values, the authors have operated a truncation of the generalized extreme value on the left side, to restrict the variability of X to the interval $0 < x < \mu - \sigma/k$. The conditional distribution $F(x|x > 0)$ is $F(x|0 < x < \mu - \sigma/k) = [F(x) - F(0)]/[1 - F(0)]$. The truncation of the distribution is performed in the following months for the P1 approach: January, February, May, and October. Similarly, for the P2 approach, it is performed in the following months: January, May, and October.

Successively, for each series of average precipitation, the cumulative frequency is calculated month by month using the Weibull plotting position.

Then SPI1, SPI3, and SPI12 are calculated using both parametric and nonparametric approaches. For each method it is verified that the distribution of SPI1 (SPI3 and SPI12) is normal using the K-S test, as suggested by Blain (2012). It is found that the test is satisfied with a p-value greater than 0.1. Figs. 1–3 report, respectively, SPI1, SPI3, and SPI12. Each figure is divided in three panels. The top panel shows the calculation of SPI using the NP approach reporting the drought months in black, and the no-drought months in gray. The intermediate panel gives the differences between P1 and NP, i.e., SPI(P1)–SPI(NP), while the bottom panel gives the differences between P2 and NP, i.e., SPI(P2)–SPI(NP). In the intermediate and bottom panels, the differences are reported with the following legend of colors: light gray if $\{SPI(NP) > -1 \cap SPI(P) > -1\}$ or if $\{SPI(NP) \le -1 \cap SPI(P) \le -1\}$, and black if $\{SPI(NP) > -1 \cap SPI(P) \le -1\}$ or if $\{SPI(NP) \le -1 \cap SPI(P) > -1\}$, where P indicates P1 or P2.

From Figs. 1–3, it is possible to see that the SPI differences between the parametric and nonparametric approach are in the range $(-2, +1)$. The parametric approaches show more extreme values of the SPI (i.e., < -2 and >2) than the nonparametric one. Fig. 4 reports the couples (T, S) in the left panels, and the couples (D, S) in the right panels, identified using the nonparametric approach (open circle), P1 (cross), and P2 (solid circle), for SPI1 in the top panels,

Table 1. Parameter Estimates of Selected Cumulative Distribution of Monthly Precipitation at Roma Collegio Romano for P1 and P2 Approaches

Month	P1 Parameters estimates of gamma 3P $F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$	P2 Distribution	Cumulative distribution function	Parameters estimates	P_0
1	$\alpha = 6.33$ $\beta = 16.19$ $\gamma = -36.13$	Generalized extreme value	$F(x) = \exp[-(1 + kz)^{-1/k}]$ $z = \frac{x-\mu}{\sigma}$	$k = -0.14$ $\sigma = 36.42$ $\mu = 49.77$	0.000
2	$\alpha = 1.71$ $\beta = 37.36$ $\gamma = -0.89$	Weibull 3P	$F(x) = 1 - \exp[-(\frac{x-\gamma}{\beta})^\alpha]$	$\alpha = 1.37$ $\beta = 68.58$ $\gamma = 0.16$	0.000
3	$\alpha = 2.65$ $\beta = 20.78$ $\gamma = 0$	Generalized extreme value	$F(x) = \exp[-(1 + kz)^{-1/k}]$ $z = \frac{x-\mu}{\sigma}$	$k = -0.04$ $\sigma = 28.11$ $\mu = 40.17$	0.013
4	$\alpha = 2.33$ $\beta = 22.82$ $\gamma = 0.44$	Weibull 3P	$F(x) = 1 - \exp[-(\frac{x-\gamma}{\beta})^\alpha]$	$\alpha = 1.53$ $\beta = 56.18$ $\gamma = 3.05$	0.000
5	$\alpha = 1.64$ $\beta = 28.33$ $\gamma = -1.36$	Weibull 3P	$F(x) = 1 - \exp[-(\frac{x-\gamma}{\beta})^\alpha]$	$\alpha = 1.33$ $\beta = 49.64$ $\gamma = -0.53$	0.000
6	$\alpha = 0.99$ $\beta = 32.74$ $\gamma = 0$	Gamma 3P	$F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$	$\alpha = 1.06$ $\beta = 30.31$ $\gamma = 0$	0.013
7	$\alpha = 0.44$ $\beta = 31.85$ $\gamma = 0$	Generalized extreme value	$F(x) = \exp[-(1 + kz)^{-1/k}]$ $z = \frac{x-\mu}{\sigma}$	$k = 0.47$ $\sigma = 7.02$ $\mu = 3.89$	0.158
8	$\alpha = 0.66$ $\beta = 38.83$ $\gamma = 0$	Generalized extreme value	$F(x) = \exp[-(1 + kz)^{-1/k}]$ $z = \frac{x-\mu}{\sigma}$	$k = 0.34$ $\sigma = 13.76$ $\mu = 10.73$	0.105
9	$\alpha = 1.66$ $\beta = 41.06$ $\gamma = 0$	Generalized gamma	$F(x) = \frac{\Gamma_{(x/\beta)^k}(\alpha)}{\Gamma(\alpha)}$	$k = 1.59$ $\alpha = 0.81$ $\beta = 88.84$	0.039
10	$\alpha = 1.86$ $\beta = 56.43$ $\gamma = 0$	Generalized extreme value	$F(x) = \exp[-(1 + kz)^{-1/k}]$ $z = \frac{x-\mu}{\sigma}$	$k = 0.01$ $\boldsymbol{\sigma} = 59.42$ $\mu = 69.87$	0.000
11	$\alpha = 1.47$ $\beta = 64.27$ $\gamma = 11.76$	Burr	$F(x) = 1 - [1 + (\frac{x}{\beta})^\alpha]^{-k}$	$k = 253.09$ $\alpha = 1.61$ $\beta = 3728.5$	0.000
12	$\alpha = 2.44$ $\beta = 34.02$ $\gamma = 0$	Generalized extreme value	$F(x) = \exp[-(1 + kz)^{-1/k}]$ $z = \frac{x-\mu}{\sigma}$	$k = 0.05$ $\boldsymbol{\sigma} = 39.72$ $\mu = 58.01$	0.013

Note: The last column gives, for each month, the estimated probability of zero precipitation; the symbol Γ indicates the gamma function, $\Gamma_{\bullet}$ the incomplete gamma function; Φ the cumulative distribution function of the standard normal distribution.

SPI3 in the intermediate panels, and SPI12 in the bottom panels. Fig. 4 gives an idea, for each meteorological drought, if there is agreement or not among the different approaches. For example, in the top right panel, the couple ($D = 5$, $S = 7.5$) is identified by the nonparametric method but not by the two parametric approaches. Conversely, the couple ($D = 3$, $S = 6$) is considered by the NP and P2 methods but not by P1. To quantify the differences between nonparametric and parametric approaches, as well as between the two parametric approaches, in terms of duration, and interarrival time between two successive droughts, and severity, Table 2 reports the mean values of percentage differences in absolute value.

From Table 2, it is possible to see how the greatest differences are in terms of severity. In particular the mean value of percentage differences in absolute value, between nonparametric and parametric approaches, is quite small in terms of T (~1–2%) and D (~6–12%), while in terms of S, it is ~30–175%. The differences evaluated choosing P1 rather than P2 are quite similar in D and T, while in terms of S, they are smaller using P2, with 3-month and 12-month time scales. The fact that the differences between parametric and nonparametric approaches are smaller in

correspondence of the P2 approach is not to be considered a rule, because the selection of the probability distribution in P2 has been done minimizing the statistic D_N, i.e., the maximum difference in absolute value between the theoretical and the empirical distribution function, not considering the other differences or the mean value of the differences. Thus, it is possible that in some cases the differences between parametric and nonparametric approaches are smaller in correspondence of the P1 approach.

The mean value of the percentage differences in absolute value between the two parametric approaches is quite small in terms of T (~1–1.5%) and D (~6–10%), while in terms of S, it is ~30–40%. Note that the differences tend to reduce as the time scale increases from 1 to 12 months, as expected.

Here the monthly time series of streamflow at the gauging station of Ripetta the Tevere River Basin is analyzed, and the average precipitation at Roma Collegio Romano at the 1-month, 3-month, and 12-month time scale is calculated. For each month the authors (1) estimated the parameters of the cumulative distribution functions considered, for P1 and P2 approaches, using the maximum likelihood method; (2) checked that the domain of the variable includes only nonnegative values, making, in contrary cases, a lower
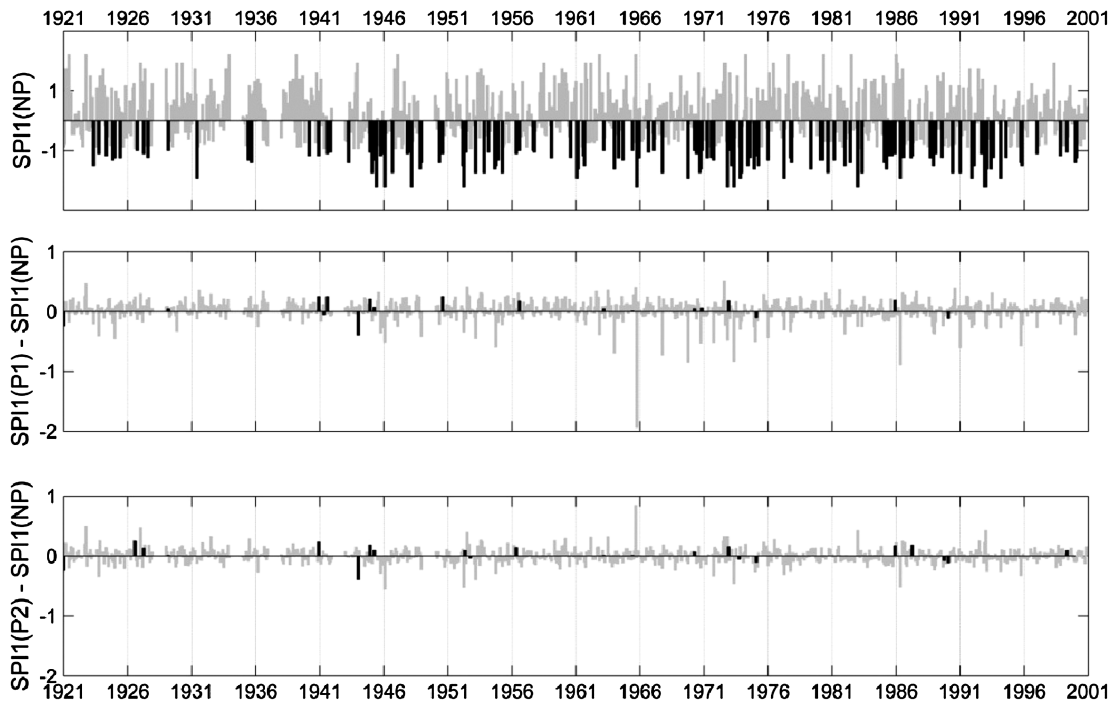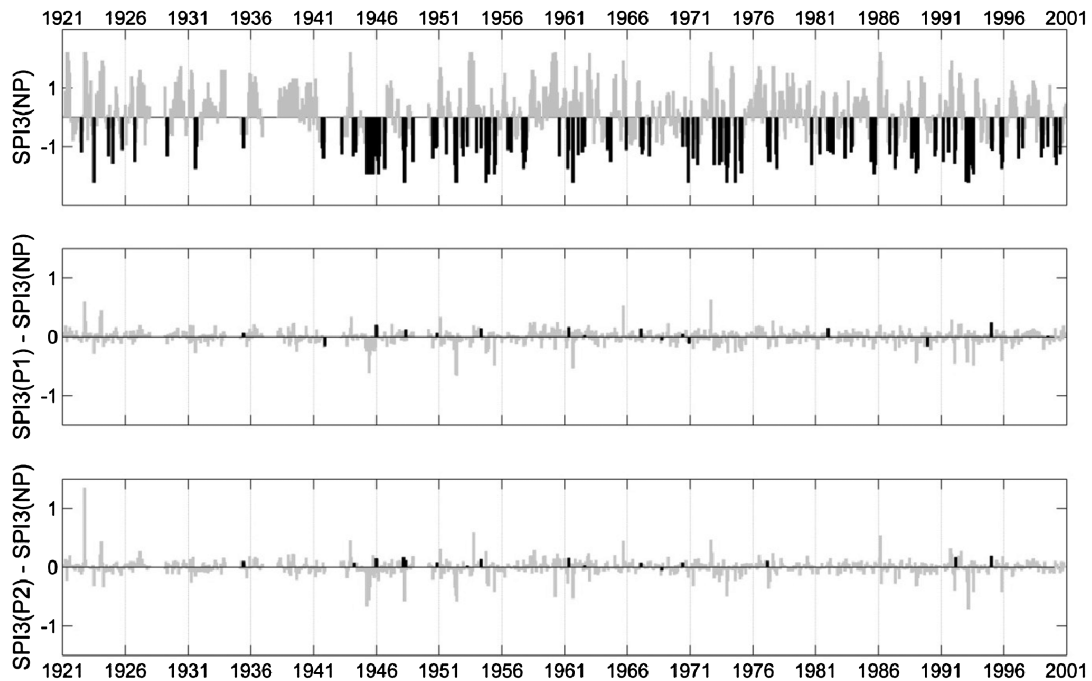
Fig. 1. SPI1 for Roma at Collegio Romano

Fig. 2. SPI3 for Roma at Collegio Romano

truncation of the distribution to avoid negative values; and (3) checked the agreement between each distribution and data using the Kolmogorov-Smirnov goodness-of-fit test as indicated by Keutelian (1991). In particular for P2 approach, the probability distribution with the smallest value of the K-S statistic was selected. Table 3 gives the parameters' values of the selected distributions, both for P1 and P2, relative to the time series of monthly streamflow.

In particular, the second column reports the estimates of the parameters of log-normal 3P, used in the P1 approach. Note that the value of the shift parameter, γ, of log-normal 3P, is always positive, except for August. Columns 3–5 of Table 3 give, for the P2 approach, respectively, the distribution for each month, its cumulative distribution function, and the parameters' estimates. The truncation of the distribution was performed only in August for the P1 approach.
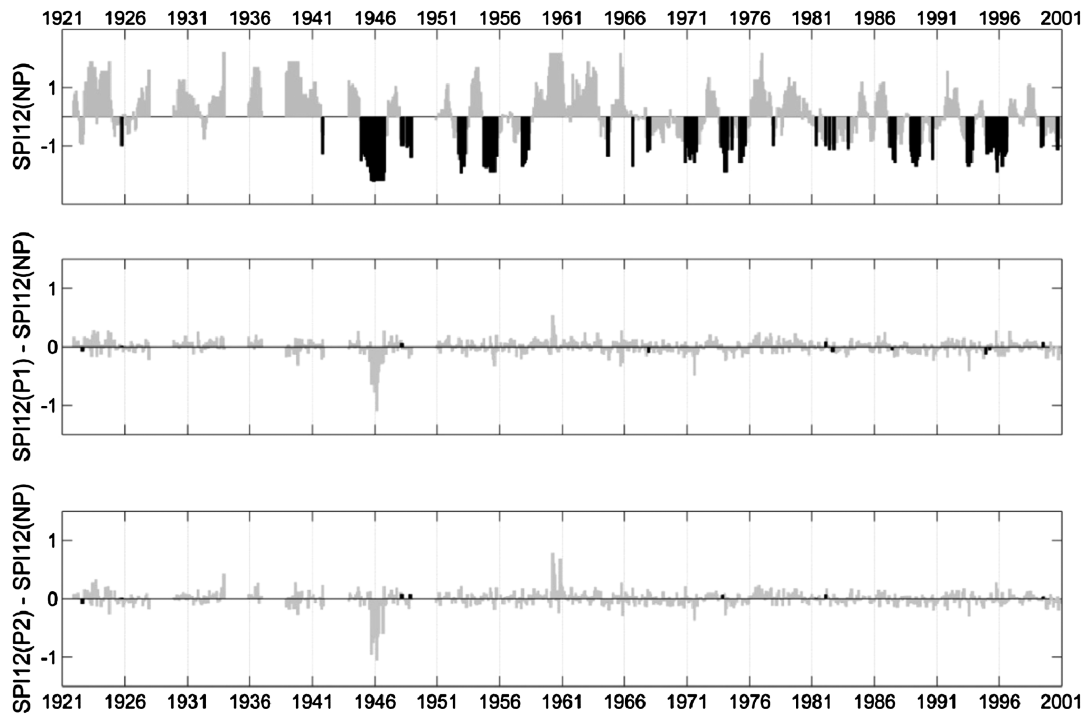
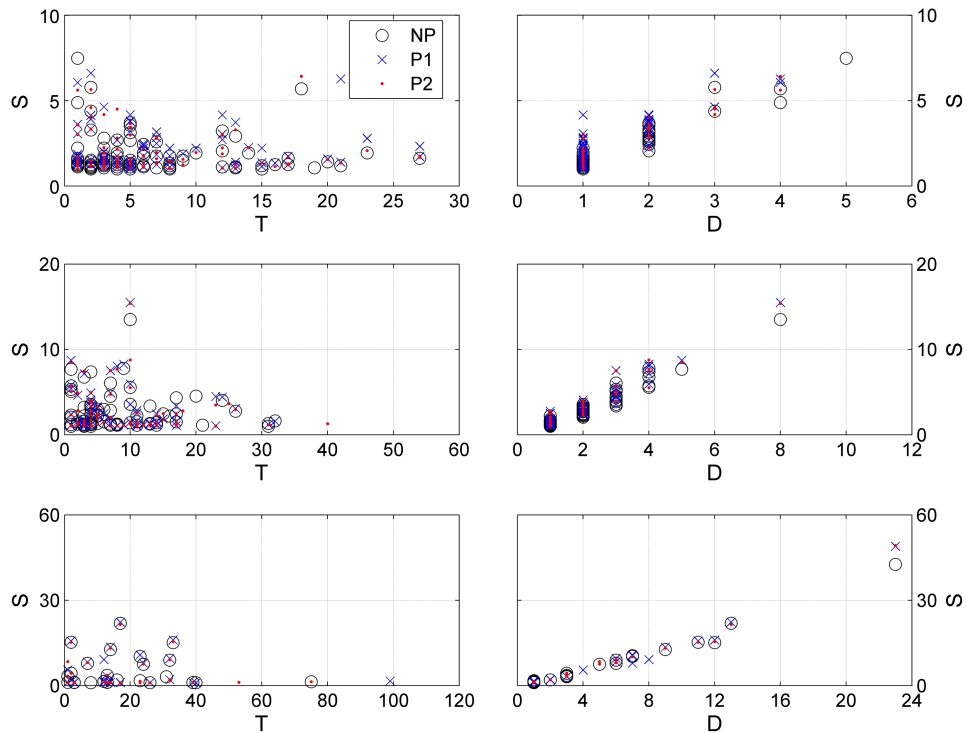**Fig. 3.** SPI12 for Roma at Collegio Romano

Fig. 4. Couples (T, S) and (D, S), for SPI1 in the top panel, SPI3 in the intermediate panel, and SPI12 in the bottom panel, identified using both nonparametric and parametric approaches

Successively, for each of the three series of average streamflow (at time scales of 1, 3, and 12 months), the cumulative frequency was calculated month by month using the Weibull plotting position. Then SSI1, SSI3, and SSI12 were calculated using both parametric and nonparametric approaches, and verified that SPI1 (SPI3 and SPI12) is normally distributed using the K-S test. Also in this case the test is satisfied with a p-value always greater than 0.1.

Table 2. Mean Value of Percentage SPI Differences in Absolute Value between Nonparametric and Parametric Approaches

	$\bar{\Delta}^{\text{NP}-\text{P1}}$			$\bar{\Delta}^{\text{NP}-\text{P2}}$			$\bar{\Delta}^{\text{P2}-\text{P1}}$		
d	D (%)	T (%)	S (%)	D (%)	T (%)	S (%)	D (%)	T (%)	S (%)
1	11.8	1.9	141.9	12.1	2.0	174.6	9.6	1.6	30.4
3	11.0	1.9	42.3	11.2	1.9	34.7	7.4	1.2	33.0
12	7.2	1.2	54.6	6.0	1.0	33.6	5.9	1.0	41.3

Note: D = drought duration; d = time scale in months; NP = nonparametric; P1 and P2 = parametric; S = severity; T = interarrival time between two successive droughts; $\bar{\Delta}$ = absolute value.

Figs. 4–7 report, respectively, SSI1, SSI3, and SSI12. Each figure is divided in three panels: the top panel shows the calculation using the NP approach, reporting the drought months in black and the no-drought months in gray. The intermediate panel gives the SSI differences between P1 and NP, i.e., SSI(P1)–SSI(NP), while the bottom panel gives the SSI differences between P2 and NP,

i.e., SSI(P2)–SSI(NP). In the intermediate and bottom panels, the differences are reported with the following legend of colors: light gray if {SSI(NP) > −1 ∩ SSI(P) > −1}, or if {SSI(NP) ≤ −1 ∩ SSI(P) ≤ −1}, and black if {SSI(NP) > −1 ∩ SSI(P) ≤ −1} or if {SSI(NP) ≤ −1 ∩ SSI(P) > −1}, where P indicates P1 or P2.

The visual comparison shows how the SSI differences between parametric and nonparametric approaches are in the range (−1, +2), with more extreme values of the index (i.e., < − 2 and >2) associated with parametric approaches. Fig. 8 reports the couples (T, S) in the left panels, and the couples (D, S) in the right panels, identified using the nonparametric approach (open circle), P1 (cross), and P2 (solid circle), for SSI1 in the top panels, SSI3 in the intermediate panels, and SSI12 in the bottom panels. Fig. 8 gives an idea, for each hydrologic drought, if there is agreement or not among the different approaches.

Table 4 reports the mean values of SSI differences in absolute value between nonparametric and parametric approaches, as well between the two parametric approaches, in terms of

Table 3. Parameter Estimates of Selected Cumulative Distribution of Monthly Streamflow for Tevere River Basin at Ripetta Using P1 and P2 Approaches

Month	P1 — Parameters estimates of log-normal 3P $F(x) = \Phi\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)$	P2 — Distribution	P2 — Cumulative distribution function	P2 — Parameters estimates
1	$\sigma = 0.53$ $\mu = 5.79$ $\gamma = 1,197.80$	Weibull 3P	$F(x) = 1 - \exp[-(\frac{x-\gamma}{\beta})^{\alpha}]$	$\alpha = 1.38$ $\beta = 5,955.80$ $\gamma = 3,265.10$
2	$\sigma = 0.47$ $\mu = 8.99$ $\gamma = 0.00$	Erlang 3P	$F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(m)}{\Gamma(m)}$	$m = 2.00$ $\beta = 3,254.00$ $\gamma = 98.83$
3	$\sigma = 0.49$ $\mu = 8.84$ $\gamma = 1,104.10$	Inverse Gaussian 3P	$F(x) = \Phi[\sqrt{\lambda/(x-\gamma)}(\frac{x-\gamma}{\mu}+1)] + \Phi[-\sqrt{\lambda/(x-\gamma)}(\frac{x-\gamma}{\mu}+1)]\exp(2\lambda/\mu)$	$\lambda = 31,688.00$ $\mu = 8,068.80$ $\gamma = 857.03$
4	$\sigma = 0.39$ $\mu = 8.72$ $\gamma = 836.64$	Inverse Gaussian 3P	$F(x) = \Phi[\sqrt{\lambda/(x-\gamma)}(\frac{x-\gamma}{\mu}+1)] + \Phi[-\sqrt{\lambda/(x-\gamma)}(\frac{x-\gamma}{\mu}+1)]\exp(2\lambda/\mu)$	$\lambda = 44,515.00$ $\mu = 6,769.70$ $\gamma = 707.93$
5	$\sigma = 0.64$ $\mu = 8.11$ $\gamma = 2,764.10$	Weibull 3P	$F(x) = 1 - \exp[-(\frac{x-\gamma}{\beta})^{\alpha}]$	$\alpha = 1.29$ $\beta = 3,562.40$ $\gamma = 3,500.00$
6	$\sigma = 0.45$ $\mu = 8.04$ $\gamma = 1,648.70$	Pearson 5	$F(x) = 1 - \frac{\Gamma_{\beta/x}(\alpha)}{\Gamma(\alpha)}$	$\alpha = 12.31$ $\beta = 57,568.00$
7	$\sigma = 0.21$ $\mu = 8.33$ $\gamma = 0.00$	Gamma 3P	$F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$	$\alpha = 21.24$ $\beta = 200.56$ $\gamma = 0.00$
8	$\sigma = 0.09$ $\mu = 9.09$ $\gamma = -5,122.30$	Gamma 3P	$F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$	$\alpha = 21.75$ $\beta = 177.60$ $\gamma = 0.00$
9	$\sigma = 0.49$ $\mu = 7.59$ $\gamma = 2,045.70$	Burr	$F(x) = 1 - [1 + (\frac{x}{\beta})^{\alpha}]^{-k}$	$k = 0.43$ $\alpha = 11.59$ $\beta = 3,556.10$
10	$\sigma = 0.73$ $\mu = 7.61$ $\gamma = 2,773.05$	Generalized gamma	$F(x) = \frac{\Gamma_{[(x-\gamma)/\beta]^k}(\alpha)}{\Gamma(\alpha)}$	$k = 0.49$ $\alpha = 6.54$ $\beta = 46.06$ $\gamma = 2,931.80$
11	$\sigma = 0.69$ $\mu = 8.29$ $\gamma = 2,325.10$	Generalized Gamma	$F(x) = \frac{\Gamma_{[(x-\gamma)/\beta]^k}(\alpha)}{\Gamma(\alpha)}$	$k = 0.49$ $\alpha = 7.24$ $\beta = 77.39$ $\gamma = 2,639.20$
12	$\sigma = 0.47$ $\mu = 8.99$ $\gamma = 0.00$	Weibull 3P	$F(x) = 1 - \exp[-(\frac{x-\gamma}{\beta})^{\alpha}]$	$\alpha = 1.34$ $\beta = 6,603.30$ $\gamma = 2,957.50$

Note: The symbol $\Gamma_{\bullet}$ is the incomplete gamma function; Φ is the cumulative distribution function of the standard normal distribution.
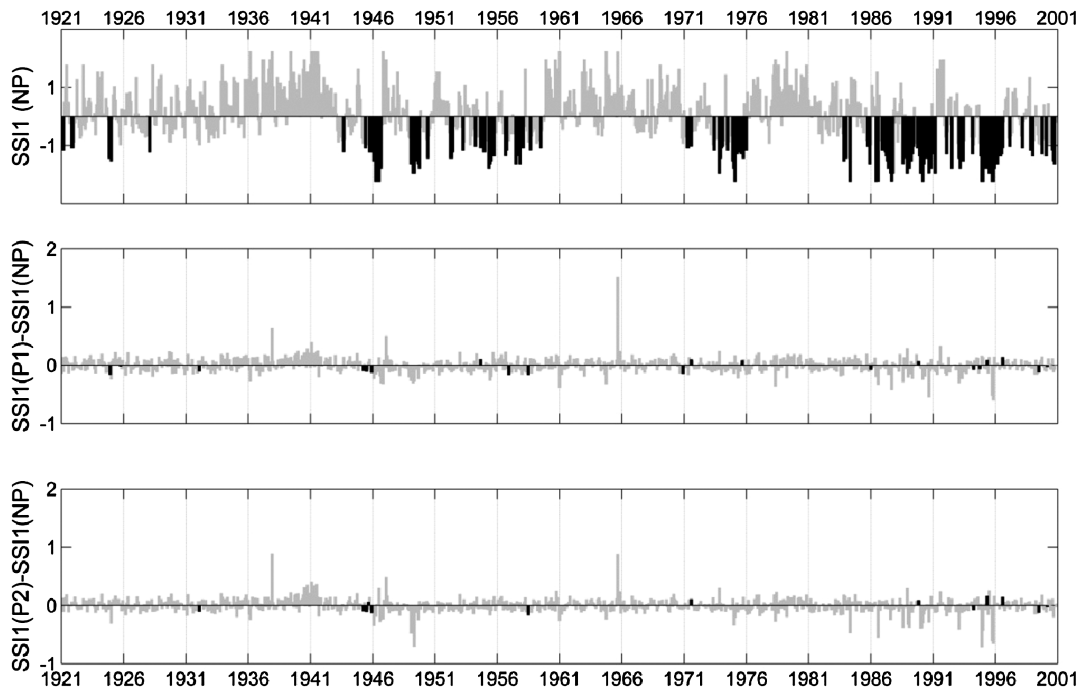
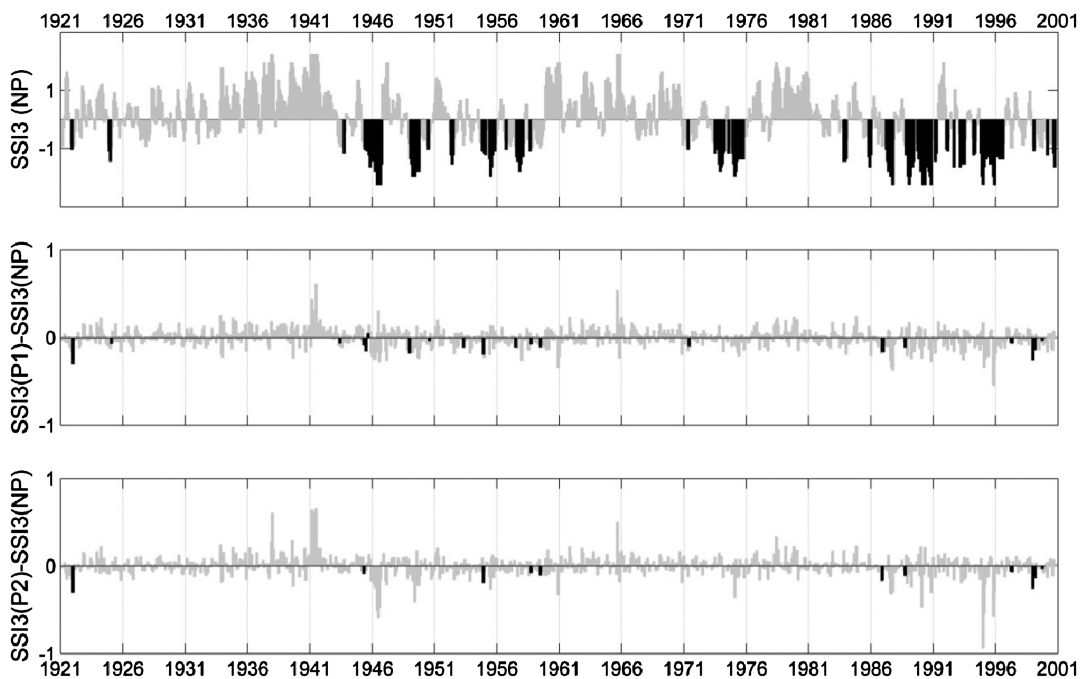**Fig. 5.** SSI1 for Tevere River Basin at Ripetta

Fig. 6. SSI3 for Tevere River Basin at Ripetta

duration, interarrival time between two successive droughts, and severity.

From Table 4, it is possible to see that for SSI, similarly to the SPI analysis, the greatest differences are in terms of severity. The mean value of percentage differences in absolute value between nonparametric and parametric approaches is quite small in terms of T (~1–3%) and D (~7–13%), while in terms of S, it is ~35–80%. The differences calculated using P1 rather than P2 are quite similar, and are smaller using P2, in correspondence of 1-month and 3-month time scales. The mean value of the percentage differences in absolute value between the two parametric approaches is quite small in terms of T (~1%), and D (~6–8%), while in terms of S, it is ~20–100%.

Last, Tables 2 and 4 are compared in order to identify differences passing from SPI to SSI, i.e., from a meteorological to a hydrological drought index. From Tables 2 and 4, it is possible to see how
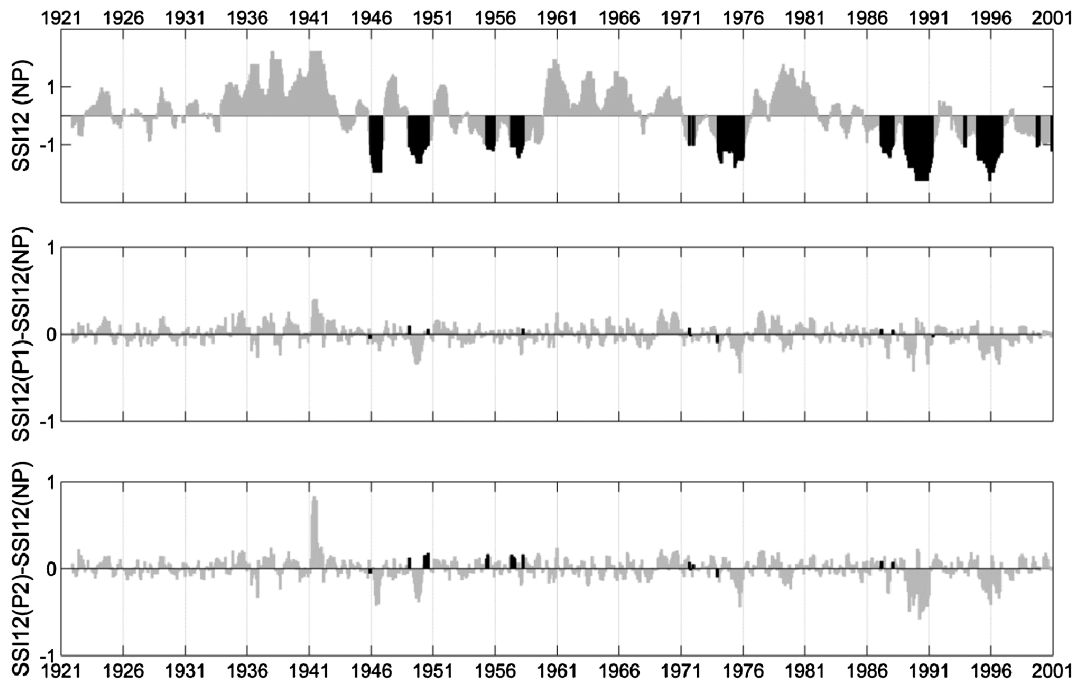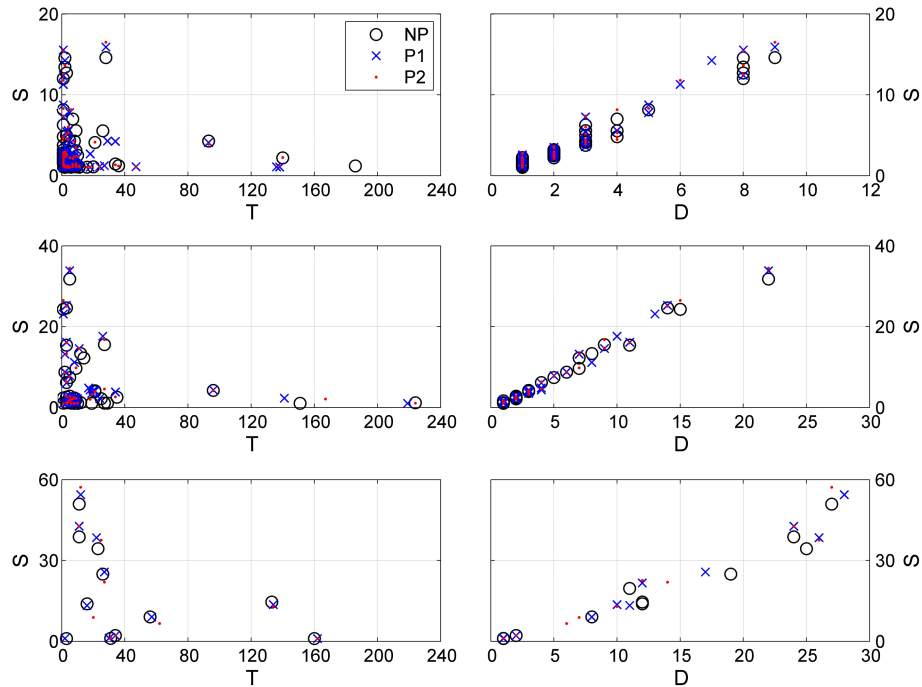
Fig. 7. SSI12 for Tevere River Basin at Ripetta

Fig. 8. Couples (*T*, *S*) and (*D*, *S*), for SSI1 in the top panel, SSI3 in the intermediate panel, and SSI12 in the bottom panel, identified using both nonparametric and parametric approaches

the mean values of the percentage differences in absolute value between nonparametric and parametric approaches, in terms of *T* and *D*, are very similar and in the range 8–10%. In terms of *S*, the mean values for SPI are higher in respect to the ones found for SSI. This difference can be explained by the different type of variable considered in SPI and SSI: the monthly precipitation is a variable with a mixed distribution, while the monthly streamflow

for the case study (Tevere at Ripetta) does not present masses in zero (P_0).

Conclusions

This paper presents a first comparison between parametric and nonparametric approaches for the calculation of two drought indices,

Table 4. Mean Value of Percentage SSI Differences in Absolute Value between Nonparametric and Parametric Approaches

d	$\bar{\Delta}^{NP-P1}$			$\bar{\Delta}^{NP-P2}$			$\bar{\Delta}^{P2-P1}$		
	D (%)	T (%)	S (%)	D (%)	T (%)	S (%)	D (%)	T (%)	S (%)
1	12.7	2.4	45.9	8.5	1.6	34.9	5.8	1.1	29.9
3	13.3	2.7	78.4	7.7	1.5	54.3	6.1	1.2	96.3
12	6.8	1.2	34.8	12.9	2.3	40.0	7.7	1.3	20.0

Note: D = drought duration; d = time scale in months; NP = nonparametric; P1 and P2 = parametric; S = severity; T = interarrival time between two successive droughts; $\bar{\Delta}$ = absolute value.

the standardized precipitation index, and the standardized streamflow index, at three time scales: 1, 3, and 12 months. In literature, the parametric approaches are generally preferred to nonparametric methods, for the extrapolation problems presented by the latter ones, even if the former ones can suffer problems of probability distribution misspecification.

Here two commonly used parametric approaches are considered: P1, with a unique probability distribution for the variable of interest with parameters depending on the month of the year, and P2, with the probability distribution depending on the month of the year and selected among a pool of continuous probability distributions using the Kolmogorov-Smirnov test. In addition, a nonparametric approach NP is implemented, in which the cumulative distribution function is approximated by the Weibull plotting position.

The present study is based on two monthly time series, 80 years long: Roma Collegio Romano for precipitation, and Tevere River Basin at Ripetta for streamflow, both of them in the period 1921–2000. The differences between parametric and nonparametric approaches, as well as between the two parametric methods in the calculation of the drought indices, at the three time scales, are summarized in terms of the mean value of percentage differences in absolute value, relative to three drought characteristics: duration, severity, and interarrival time between two successive droughts.

From the analysis of the two time series, the authors find that the differences between parametric and nonparametric approaches are more evident in terms of severity, and less in terms of duration and interarrival time, i.e., more on drought entity, and less on drought identification.

In particular, the mean value of percentage differences in absolute value is, in terms of severity, ~30–175% for SPI, and ~30–80% for SSI, depending on the time scale considered. In terms of duration, it is ~6–13%, while in terms of interarrival time it is ~1–3%, both for SPI and SSI. The differences between the two parametric approaches, in SPI and SSI, are quite similar in terms of duration (~6–10%) and interarrival time (~1–2%), while in terms of severity they are ~30–40% for SPI, and ~20–100% for SSI.

References

Alley, W. M. (1984). "The palmer drought severity index. Limitations and assumptions." *J. Clim. Appl. Meteorol.*, 23(7), 1100–1109.

Beran, M., and Rodier, J. (1985). "Hydrological aspects of drought: A contribution to the international hydrological programme." *Studies and reports in hydrology*, Vol. 39, UNESCO-World Meteorological Organization, Paris.

Blain, G. C. (2012). "Revisiting the probabilistic definition of drought: Strengths, limitations and an agrometeorological adaptation." *Bragantia*, 71(1), 132–141.

Bordi, I., Fraederich, K., and Sutera, A. (2009). "Observed drought and wetness in Europe: An update." *HESS*, 13(8), 1519–1530.

Bothe, O., Fraedrich, K., and Zhu, X. (2010). "The large-scale circulations and summer drought and wetness on the Tibetan plateau." *Int. J. Climatol.*, 30(6), 844–855.

Dai, A. (2011). "Drought under global warming: A review." *Wiley Interdiscip. Rev. Clim. Change*, 2(1), 45–65.

Dracup, J. A., Lee, K. S., and Paulson, E. G. (1980). "On the definition of droughts." *Water Resour. Res.*, 16(2), 297–302.

Gibbs, W. J., and Maher, J. V. (1967). "Rainfall deciles as drought indicators." *Bureau of Meteorology Bulletin No. 48*, Commonwealth of Australia, Melbourne, Australia.

Guttman, G. B. (1999). "Accepting the standardized precipitation index: A calculation algorithm index." *J. Am. Water Resour.*, 35(2), 311–322.

Hayes, M. (2000). "Revisiting the SPI: Clarifying the process." *Drought Netwrok News (1994-2001)*, Paper 18.

Heim, R. (2002). "A review of twentieth century drought indices used in the United States." *Bull. Am. Meteorol. Soc.*, 83, 1149–1165. Keutelian,

H. (1991). "The Kolmogorov-Smirnov statistic when estimating parameters from data." *Collider detector facility note 1285*, Fermilab, Batavia, IL.

Kottegoda, N., and Rosso, R. (1997). *Statistics, probability and reliability for civil and environmental engineers*, McGraw-Hill, New York.

Kroll, C. N., and Vogel, R. M. (2002). "The probability distribution of low streamflow series in the United States." *J. Hydrol. Eng.*, 10.1061/(ASCE)1084-0699(2002)7:2(137), 137–146.

Lopez-Moreno, J. I., and Vicente-Serrano, S. M. (2008). "Positive and negative phases of the wintertime North-Atlantic oscillation and drought occurrences over Europe: A multitemporal-scale approach." *J. Clim.*, 21(6), 1220–1243.

Madadgar, S., and Moradkhani, H. (2013). "Drought analysis under climate change using copula." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000532, 746–759.

McKee, T. B. N., Doesken, J., and Kleist, J. (1993). "The relationship of drought frequency and duration to time scales." *Proc., 8th Conf. on Applied Climatology*, American Meteorological Society, Boston, 179–184.

McMahon, T. A., Pegram, G. G. S., Vogel, R. M., and Peel, M. C. (2007). "Revisiting reservoir storage-yield relationships using a global streamflow database." *Adv. Water Resour.*, 30(8), 1858–1872.

Mishra, K., and Singh, P. (2010). "A review of drought concepts." *J. Hydrol.*, 391(1–2), 202–216.

Mo, K. C., and Schemm, J. E. (2008). "Relationship between ENSO and drought over the southeastern United States." *Geophys. Res. Lett.*, 35(15), L15701.

Nalbantis, I. (2008). "Evaluation of a hydrological drought index." *Eur. Water*, 23–24, 67–77.

Palmer, W. C. (1965). "Meteorological drought." *Research Paper No. 45*, U.S. Dept. of Commerce Weather Bureau, Washington, DC.

Palmer, W. C. (1968). "Keeping track of crop moisture conditions, nationwide: The new crop moisture index." *Weatherwise*, 21(4), 156–161.

Pettit, A. N. (1979). "A non-parametric approach to the change-point detection." *Appl. Statist.*, 28(2), 126–135.

Santos, J. F., Pulido-Calvo, I., and Portela, M. M. (2010). "Spatial and temporal variability of droughts in Portugal." *Water Resour. Res.*, 46(3), W03503.

Shafer, B. A., and Dezman, L. E. (1982). "Development of surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas." *Proc., Western Snow Conf.*, Colorado State Univ., Reno, NV, 164–175.

Shukla, S., and Wood, A. W. (2008). "Use of standardised runoff index for characterizing hydrologic drought." *Geophys. Res. Lett.*, 35(2), L02405.

Sienz, F., Bothe, O., and Fraedrich, K. (2012). "Monitoring and quantifying future climate projections of dryness and wetness extremes: SPI bias." *Hydrol. Earth Syst. Sci.*, 16(7), 2143–2157.

Tsakiris, G. (2004). *Meteorological drought assessment*, Project MEDROPLAN (Mediterranean Drought Preparedness and Mitigation Planning), Zaragoza, Spain.

Tsakiris, G., et al. (2007). "Drought characterization [Part 1. Components of drought planning. 1.3. Methodological component]." *Drought management guidelines technical annex*, A. Iglesias, M. Moneo, and A. López-Francos, eds., CIHEAM/EC MEDA Water, Zaragoza, 85–102.

Vicente-Serrano, S. M. (2006). "Differences in spatial patterns of drought on different time scales: An analysis of the Iberian Peninsula." *Water Resour. Manage.*, 20(1), 37–60.

Vicente-Serrano, S. M., et al. (2012). "Accurate computation of a streamflow drought index." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584 .0000433, 229–358.

Wijngaard, J., Klein Tank, A. M. G., and Können, G. P. (2003). "Homogeneity of 20th century European daily temperature and precipitation series." *Int. J. Climatol.*, 23(6), 679–692.

Wilhite, D. A. (2000). "Drought as a natural hazard: Concepts and definitions." *Drought: A global assessment*, Vol. 1, D. Wilhite, ed., Routledge, London, New York, 3–18.

Wilhite, D. A., and Glantz, M. H. (1987). "Understanding the drought phenomenon: The role of definitions." *Planning for drought*, D. A. Wilhite and W. E. Easterling, eds., Westview Press, Boulder, CO, 11–27.

Wu, H., Hayes, M. J., Wilhite, D. A., and Svoboda, M. D. (2005). "The effect of the length of record on the standardized precipitation index calculation." *Int. J. Climatol.*, 25(4), 505–520.

Yue, S., and Wang, C. Y. (2004). "Possible regional probability distribution type of Canadian annual streamflow by L-moments." *Water Resour. Manage.*, 18(5), 425–438.

Zhu, X., Bothe, O., and Fraedrich, K. (2011). "Summer atmospheric bridging between Europe and east Asia: Influences on drought and wetness on the Tibetan Plateau." *Quat. Int.*, 236(1–2), 151–157.