Global Change: Facing Risks and Threats to Water Resources (Proc. of the Sixth World FRIEND Conference, Fez, Morocco, October 2010). IAHS Publ. 340, 2010.

Regional analysis of low flow in Tuscany (Italy)

GIUSEPPE ROSSI & ENRICA CAPORALI

Dipartimento di Ingegneria Civile e Ambientale, Università degli Studi di Firenze Via S. Marta 3, 50139 Firenze, Italy
giuseppe.rossi@dicea.unifi.it

Abstract In environment protection, estimates of low flows in rivers are needed for many purposes. Generally estimation is based on observed streamflow data. For sites where data are not available, alternative techniques are necessary to infer this information. A regional approach is often used for ungauged basins and is the one adopted for this study. The analysis is carried out on the discharge data of 65 consistent hydrometric stations located in the Tuscany region, central Italy, with recorded data from 1949 to 2008. The area is subdivided into different regions using the L-moments method applied to the 7-day annual minimum flows and to the Q_{70} annual series. The division into sub-regions was tested using discordancy and heterogeneity statistics. A unique region and a subdivision into three different sub-regions, following previous studies on rainfall extremes were considered. The subdivision into five homogeneous sub-regions was undertaken by accounting for hydrological features.

Key words drought; low-flow; ungauged rivers; regional analysis; L-moments; homogeneity measure

INTRODUCTION

Streamflow data analysis, characterization of low-flow behaviour and hydrological drought indices play an important role in engineering practice for water resources design and management and their definition is necessary for several purposes, including water supply planning, river basin management, hydropower development, and environmental flow characterization (Smakhtin, 2001). Another important field and difficult task in which they are utilized is in identifying the occurrence, the extent, and the magnitude of a drought. In the framework of risk monitoring and management, low flow data, such as drought regime magnitude, frequency, spatial extent and seasonality, can effectively support the activity of decision-makers and politicians (Garrote *et al.*, 2006).

Drought indicators are defined as a single observation or combinations of observations that contribute to identifying the occurrence, continuation and magnitude of a drought event (Hisdal & Tallaksen, 2000). Drought indicators can include measures of streamflow, precipitation, reservoir storage, or the evaluation of meteorological indices as a function of precipitation, temperature, the available water content of the soil, and other variables. The effectiveness of drought indicators depends on the specific region and the characteristics of the system. No single indicator can work for all regions (Tallaksen *et al.*, 2004).

The beginning and the persistency of droughts can be recognized with meteorological indices. With indices derived from low flows it is possible to recognize hydrological droughts that mainly affect water supply systems (Cancelliere *et al.*, 1998; Garrote *et al.*, 2009). Different methods to derive hydrological drought characteristics are needed in order to describe the variety of streamflow droughts. The selection of an appropriate method can be even more difficult when drought events of several streams within one region are to be analysed (Menedez, 1995; Tallaksen *et al.*, 1997).

Low flow characteristics are estimated from observed streamflow data, identifying duration curves, indices and percentiles characteristics. Two main groups of low flow indices are usually used to identify the drought. The first group is derived from the Flow Duration Curve (FDC). The flow duration curve identifies, for all observed discharge values, the percentage of time when higher discharge values are observed. It can be used as a low flow index once it is normalized by its median value (Castellarin *et al.*, 2006). Q_{90} and Q_{70} are in the second group and belong to the 90 and 70 percentiles that are frequently chosen to evaluate threshold levels in drought event definition (Pyrce, 2004).

Another common index of low flow is the annual minimum *n*-day discharge. It is the smallest average discharge of *n* consecutive days within one year. The common averaging interval (values of *n*), are 1, 7, 10, and 30 days (Gustard *et al.*, 1992). In the United States, the most widely used low flow index is the 10-year annual minimum 7-day discharge. In Tuscany, the Arno Authority River Basin, refers to the 2-year annual minimum 7-day discharge to define a low streamflow requirement.

In the real world observed streamflow data are scarce and the lack of data is a diffuse problem. For ungauged sites or sites where few data are available, alternative techniques are necessary to infer low flow characteristics. The regional regression approach is the most widely used technique to face this problem (Riggs, 1973). The regional analysis improves the capability of predicting the water flow regime at gauged sites with short time series, reducing the uncertainties, and moreover allows the estimation of the discharge properties at ungauged sites (Chokmani & Ouarda, 2004).

The first step in regionalization studies is the delineation of hydrologically and statistically homogeneous regions. In some cases it is clear how to group a domain into regions of approximately uniform hydrological and statistical behaviour but, more often, the choice is far from obvious. The delineation of a region may be accomplished using convenient boundaries based on geographic or physiographic considerations. Geographically contiguous regions may be established on the basis of residuals analysis from a regional regression model developed to estimate flow characteristics at ungauged catchments. To employ geographically contiguous regions is easier than non contiguous regions, especially in the context of scarcity of data. On the other hand, even two adjacent river catchments may have different topography, soils or other local anomalies (Laaha & Bloeschl, 2005).

Classification of catchments into groups may also be based on standardized flow characteristics estimated from available observed or simulated streamflow records (Hayes, 1991; Modarres, 2008). Otherwise the regions are delineated using catchment physiographic and climatic parameters obtained from maps and hydro-meteorological data (e.g. rainfall, evaporation) (Riggs, 1973; Chockmani & Ouarda, 2004; Laaha & Bloeschl, 2007). The second group of methods although needing more data, gives better results.

Useful statistics for regional frequency analysis, which measure regional homogeneity and goodness-of-fit have been proposed by Hosking & Wallis (1993), based on L-moments method defined by Hosking (1990).

In this work, attention is focused on low flow indices that are estimated from streamflow for the period 1949–2008. In particular, the low flow events are represented by two indices: the 7-day annual minimum series and by the annual Q_{70} series. The discordancy and heterogeneity parameters are used to test different sub-division hypotheses to evaluate the regional homogeneity.

STUDY AREA AND DATASET DESCRIPTION

The analysis is carried out on the discharge data recorded in several rivers in the Tuscany Region central Italy (Fig. 1). Tuscany is a region with an area of 23 000 km^2 and 3 600 000 inhabitants (Regione Toscana *et al.*, 2008). The minimum altitude is –3 m a.s.l. in the area of Massaciuccoli Lake, while the maximum altitude is 2037 m a.s.l. in Monte Prado. The main rivers of the region are: Arno, Serchio, and Ombrone Grossetano. The Arno basin occupies one third of Tuscany's area. Moreover there are small basins of coastal rivers near the Tyrrhenian Sea and the upstream part of Tevere, Fiora and Magra watersheds. For these last inter-regional basins, no streamflow data are available for this study.

The data used in the study were from Servizio Idrologico Regionale Toscano (Regional Hydrologic Service of Tuscany) using data of the network previously managed by Ufficio Idrografico e Mareografico (Hydrographic and Mareographic Office) integrated with a new network.

The dataset has more than 500 gauged stations that measure different hydrometric variables. There are 326 stations with recorded hydrometric data (Fig. 2(a)). Some of the records are only a few months long, or have totally discontinuous data. Only 121 stations, that had at least two years of data, were first selected for the analysis: 47 of these having only stage data, while 74 had

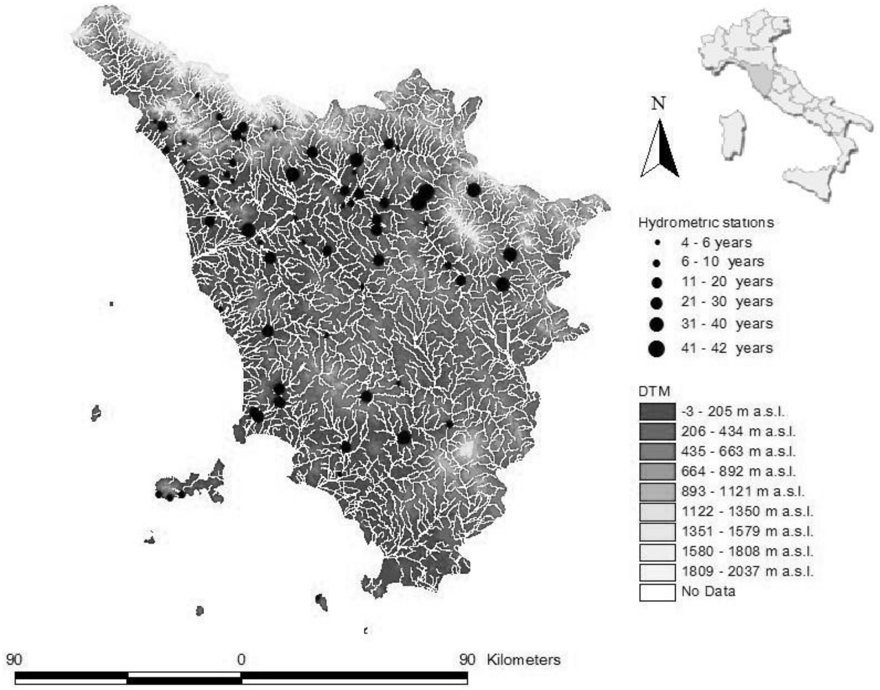

Fig. 1 The Tuscany Region and the considered hydrometric stations with the years of registrations. In white the hydrographic network.

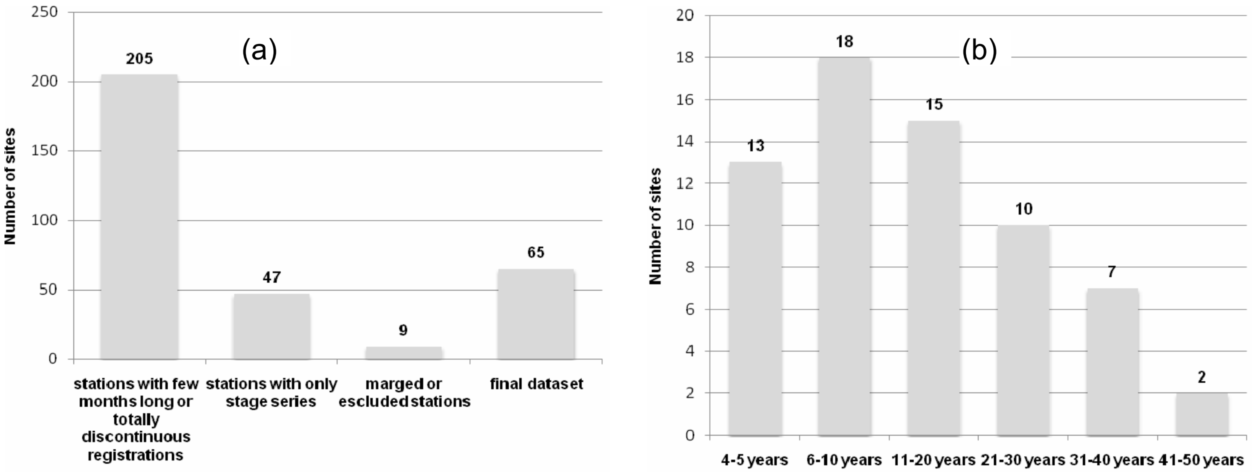

Fig. 2 (a) Dataset consistency; (b) length of time series of considered hydrometric stations.

discharge data or stage data with a related stage–discharge rating curve. A dataset of 65 stations was finally obtained by not using stations with long periods of inactivity and merging the data of traditional analogical and digital automatic stations if in the same location. Several stations had data from the 1930s, but the series were discontinuous and data were only collected during extreme high discharge events. The final dataset was from 1949 to 2008. The classification based on the years of recording is shown in Fig. 2(b).

METHODOLOGY

The aim was to find hydrologically and statistically homogeneous regions in the area of interest, using standardized low flow characteristics from available observed streamflow records

(1949–2008) for the Tuscany Region, central Italy. Following this, a low flow event regional frequency analysis, based on L-moments was carried out.

Low flow events are represented here by two indices: the 7-day annual minimum series and the annual Q_{70} series. The L-moments approach was used to assign these data to the different regions, according to homogeneity measures and properties. After testing and arranging the data, various indices were calculated. The method was performed firstly with the 7-day annual minimum series.

The L-moments were defined by Hosking (1990) as linear combinations of probability weighted moments (PWMs), previously introduced by Greenwood *et al.* (1979) and estimated for the Generalized Extreme Value distribution by Hosking *et al.* (1985). Hosking & Wallis (1993) extended the use of L-moments and developed useful statistics for regional frequency analysis, in particular the discordancy and heterogeneity parameters to evaluate the regional homogeneity. The three heterogeneity parameters that measure the homogeneity of a region, are the first, first and second, first and third L-moments, while the discordancy parameter is used to identify those sites that are grossly discordant with the group as a whole taking into account all the L-moments. Following Hosking's (1990) definition, X is a real-value random variable with cumulative distribution function $F(x)$ and quantile function $x(F)$, then the L-moments of X are:

$$\lambda_r = r^{-1}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}\,EX_{r-k:r} \text{ with } r = 1,2,\ldots \tag{1}$$

Once L-moments are calculated it is possible to derive two statistics with the aim of testing the homogeneity of a region (Hosking & Wallis, 1993). The L-moments representing coefficient of variation, skewness and kurtosis (L-cv, L-sk and L-ku) of a site can be considered as a point in a three dimensional space. A group of homogeneous sites gives a cloud of points. Any point that is far from the centre of the cloud is discordant. The first statistic to evaluate the distance of a point is the discordancy measure. Let $\boldsymbol{u} = [\mathrm{L} - \mathrm{cv}^{(i)}\ \mathrm{L} - \mathrm{sk}^{(i)}\ \mathrm{L} - \mathrm{ku}^{(i)}]$ be a vector with the values for the i-site. Let:

$$\bar{u} = N^{-1}\sum_{i=1}^{N}u_i \tag{2}$$

and $\boldsymbol{S}$ be the covariance matrix, the discordancy for the site i is defined as:

$$D_i = 1/3\ (u_i - \bar{u})^T S^{-1}(u_i - \bar{u}) \tag{3}$$

Large values of D_i indicate sites that are most discordant from the group. A site is considered to be unusual if the discordancy measure (D) is larger than 3.

Other statistics applied for homogeneity test are three heterogeneity measures (*Hi*), namely, *H1*, *H2* and *H3* with respect to L-cv scatter, Lcv–Lsk and Lcv–Lku, respectively. *Hi* for a specific site is defined as:

$$H_i = \frac{value\ at\ site - mean\ of\ i\ values}{s\tan dard\ deviation\ of\ i\ values} \tag{4}$$

Large values of *Hi* indicate sites that are most discordant from the group. A region is homogenous if any of the *Hi* values is less than 1, possibly heterogeneous if *Hi* is between 1 and 2, and definitely heterogeneous if *Hi* is more than 2 (Hosking & Wallis, 1993).

Using the above method the area of interest is divided into different regions, geographically contiguous, and the homogeneity measures are calculated to test each subdivision. Proceeding by trial-and-error some sub-basins were moved from one region to another, and some regions were split into sub-regions to reach the best possible homogeneity.

DISCUSSION OF RESULTS

Different methods to derive streamflow characteristics are needed in order to characterize the whole range of hydrological droughts. Data from Servizio Idrologico Regionale Toscano were tested and various hydrological droughts indices were calculated. Two kinds of low flow indices were chosen: the Q_{70}, derived from the Flow Duration Curve – FDC (Fig. 3(a)) and the $Q(7,2)$, the smallest average discharge of 7 consecutive days within 2 years (Fig. 3(b)). The indices were calculated only for the stations that had at least 6 years of data.

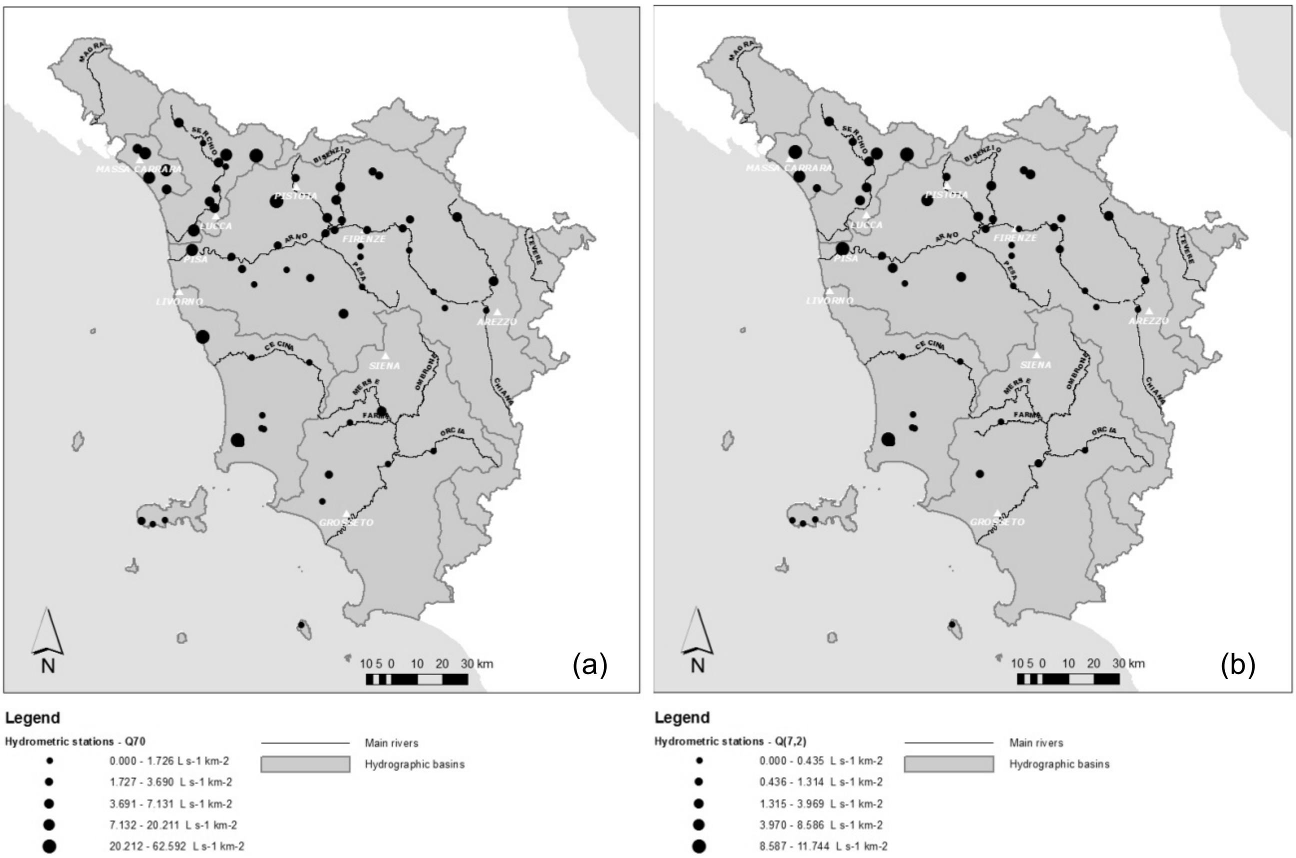

Fig. 3 (a) Q_{70} values in L s^{-1} km^{-2} at considered hydrometric stations; (b) *Q(7,2)* values in L s^{-1} km^{-2} at considered hydrometric stations.

Table 1 Heterogeneity (*H*1, *H*2, *H*3) and discordancy (*D*) parameters.

Regions	Number of stations	$H1 > 2$	$H2 > 2$	$H3 > 2$	$D > 2$	$D > 3$
Unique	48	2	2	4	9	5
North	21	1	1	1	4	2
Centre	21	0	1	1	4	2
South	6	0	0	0	1	0
Northeast	11	0	0	0	1	0
Northwest	9	0	0	0	1	1
Centre East	11	0	0	0	0	0
Centre West	9	0	0	0	3	0
South	7	0	0	0	0	0

The L-moments of the 7-day annual minimum series are also calculated for all the stations with at least 6 years of registration. The discordancy and heterogeneity parameters to evaluate the regional homogeneity are used on the $Q(7)$ to test different sub-divisions hypotheses. The discordancy (*D*) and the heterogeneity (*H*1, *H*2, *H*3) are calculated first for the whole area as a unique region. Several stations have values that suggested that this approximation was not correct. In particular 5 stations have values of discordancy higher than 3 (Table 1), the threshold value of the discordancy measure.

The area was successively split into three different sub-regions, following previous studies on rainfall extreme values (Tartaglia *et al.*, 2006; Caporali *et al.*, 2008). With this subdivision there was some homogeneity, but some stations still presented high values of discordancy. Finally a new subdivision was proposed with 5 sub-regions (Fig. 4), splitting the central and the northern regions of the three previous subdivisions, as well as following the main hydrological watersheds. The station of Colonna is not included in the subdivision, due to non-homogeneity of its data. With this subdivision the regions are more homogeneous, and the subdivision follows hydrological and precipitation features.

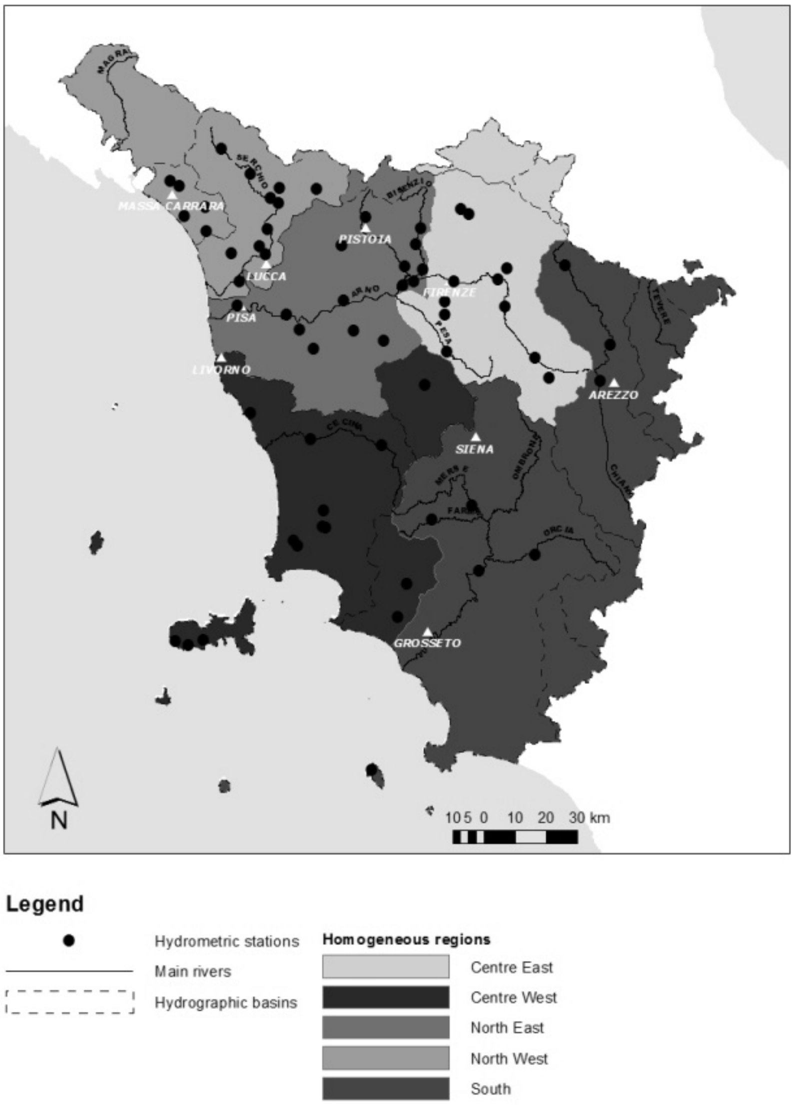

Fig. 4 Subdivision in hydrologically and statistically homogeneous regions.

CONCLUSION AND DEVELOPMENTS

The Tuscany Region rivers low flows were analysed and a subdivision in homogeneous regions was evaluated with the L-moments method and with hydrological characteristics of the studied area. As a result, the region was subdivided into five sub-regions with good properties of homogeneity.

The subdivision will be employed in further analyses, with the aim to investigate the applicability of physiographical space-based interpolation techniques for the prediction of low-flow characteristics in ungauged rivers basins (Chokmani & Ouarda, 2004). In particular, a proper set of catchment physiographic (i.e. slope, aspect, land use and soil properties) and climatic characteristics will be defined and a physiographical space-based method used to relate the low flow indices to the various areas. The new physiographical space is to be built as a linear combination of the physiographic and climatic catchment characteristics.

Different interpolation techniques, either geostatistical or deterministic, such as Kriging, Inverse Distance Weight and Thiessen polygons, of the low flow indices of the physiographical space will be applied. Uncertainty measurements will be implemented using jack-knife and bootstrap methods. Different error measurement (mean square error, mean relative error) will also be assessed to compare the results, to quantify the accuracy of the different techniques, and to define the most suitable procedure.

REFERENCES

Cancelliere, A., Ancarani, A. & Rossi, G. (1998) Susceptibility of Water supply reservoirs to drought conditions. *J. Hydrol. Engng* **3**, 140–148.

Caporali, E., Cavigli, E. & Petrucci A. (2008) The index rainfall in the regional frequency analysis of extreme events in Tuscany (Italy). *Environmetrics* **19**, 714–724.

Castellarin, A., Camorani, G. & Brath, A. (2006) Predicting annual and long-term flow-duration curves in ungauged basins. *Adv. Water Resour.* **30**, 937–953.

Chokmani, K. & Ouarda, T. B. M. J. (2004) Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. *Water Resour. Res.* **40**, 1–12.

Garrote, L., Iglesias, A. & Flores, F. (2008) Development of drought management plans in Spain. In: *Coping with Drought Risk in Agriculture and Water Supply Systems* (ed. by A. Iglesias, A. Cancelliere, D. A. Wilhite, Garrote L. & Cubillo F.), 175–184. Advances in natural and technological hazards research 26, Springer, The Netherlands.

Garrote, L., Martin-Carrasco, F., Flores-Montoya, F. & Iglesias, A. (2006) Linking drought indicators to policy actions in the Tagus basin drought management plan. *Water Res. Manage.* **21**, 873–882.

Greenwood, J. A., Landwahr, J. M., Matalas, N. C. & Wallis, J. R. (1979) Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour. Res.* **15**, 1049–1054.

Gustard, A., Bullock, A. & Dixon, J. M. (1992) Low flow estimation in the United Kingdom. *Institute of Hydrology Report 40 no. 108, Wallingford, UK.*

Hisdal, H. & Tallaksen L. (2000) Drought Event Definition. *ARIDE Technical report no. 6*. http://www.hydrology.uni-freiburg.de/forsch/aride/ (June 2009).

Hayes, D. C. (1991) Low-flow characteristics of streams in Virginia. *US Geol. Survey Water-Supply Paper*, *File Report 89*, 586–671.

Hosking, J. R. M. (1990) L-moments: analyzing and estimation of distributions using linear combinations of order statistics. *Royal Statistical Soc. J., Series B* **52**, 105–124.

Hosking, J. R. M. & Wallis, J. R. (1993) Some statistical useful in regional frequency analysis. *Water Resour. Res.* **29**, 271–281.

Hosking, J. R. M., Wallis, J. R. & Wood, E. F. (1985) Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. *Technometrics* **27**(3), 251-261.

Laaha, G &, Bloschl, G. (2005) A comparison of low flow regionalization methods – catchment grouping. *J. Hydrol.* **323**, 193–214.

Laaha, G. & Bloschl, G. (2007) A national low flow estimation procedure for Austria. *Hydrol. Sci. J.* **52**, 625–644.

Menedez, M. (1995) Aspectos Hidrologicos de las Sequias. Las sequias en España. Centro de Estudios Hidrograficos del Cedex. Madrid, Spain.

Modarres, R. (2008) Regional frequency distribution type of low flow in north of Iran by L-moments. *Water Resour. Manage.* **22**, 823–841.

Pyrce, R. (2004) Hydrological Low Flow Indices and their uses. *Watershed Science Centre Report 04-2004.*

Riggs, H. C. (1973) Regional analysis of streamflow characteristics. US Geological Survey Techniques of Water Resources, United States Government Printing office, Washington, 1982, USA.

Regione Toscana, Unioncamere Toscana, ISTAT – Istituto Nazionale di statistica (2009) Annuario Statistico Regionale, Toscana 2008. *Pubblicazioni del Sistema Statistico Regionale*. http://www.regione.toscana.it (June 2009)

Smakhtin, V. U. (2001) Low flow hydrology: a review. *J. Hydrol.* **240**, 147–186.

Tallaksen, L. M. & Van Lanen H. A. J. (2004) *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater*. Developments in Water Science 48, Elsevier Science, The Netherlands.

Tallaksen, L. M., Madsen, H. & Clausen B. (1997) On the definition and modeling of streamflow drought duration and deficit volume. *Hydrol. Sci.* **1**, 15–34.

Tartaglia, V., Caporali, E., Cavigli, E. & Moro, A. (2006) L-moments based assessment of a mixture model for frequency analysis of rainfall extremes. *Adv. Geosci.* **2**, 331–334.