Journal of Hydrology 513 (2014) 179–191

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Regional prediction of flow-duration curves using a three-dimensional kriging

Attilio Castellarin *

Civil, Chemical, Environmental and Materials Engineering – DICAM, Università degli Studi di Bologna, Bologna, Italy

ARTICLE INFO

Article history:
Received 14 August 2013
Received in revised form 15 March 2014
Accepted 23 March 2014
Available online 31 March 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Uwe Haberlandt, Associate Editor

Keywords:
Surface water availability
Hydrological Predictions in Ungauged Basins (PUB)
Physiographic-Space Based Interpolation (PSBI)
Leave-one-out cross-validation (LOOCV)
South Tyrol and Central Italy

SUMMARY

The relationship between magnitude and frequency of daily streamflows over a number of years for a given basin is often termed long-term flow-duration curve (FDC). This analysis addresses the problem of predicting FDCs in ungauged basins by means of a three-dimensional (3D) kriging interpolation of empirical FDCs. A three-dimensional *xyz* space is defined to perform the interpolation, where *x* and *y* are functions of physiographic and climatic catchment descriptors, while *z* represents the streamflow duration in terms of standard-normal variate. The 3D interpolation technique is applied to several catchments located in two broad geographical regions of Northern (Alpine catchments) and Central (Apenninic catchments) Italy, for which several geomorphological and climatic descriptors are available. An extensive cross-validation procedure is used to quantify the accuracy of the proposed technique in both case studies, and to compare it to traditional regionalization procedures. The cross-validation points out that 3D kriging is a reliable and robust approach, which performs as well as or better than traditional regional models. In particular the approach significantly outperforms conventional approaches for the prediction of low-flows (i.e. streamflows associated with high durations) in ungauged basins.

1. Introduction

Empirical long-term flow-duration curves (FDCs) represent the percentage of time (or duration) during which a given streamflow was equalled or exceeded over the historical period of record. Deterministic and statistical hydrological viewpoints coexist in an FDC: a deterministic hydrologist would probably refer to an FDC as a key signature of the hydrologic behavior of a given basin, as it results from the interplay of climatic regime, size, morphology, and permeability of the basin; a statistical hydrologist would refer to an FDC as the exceedance probability function of streamflows. An FDC provides a simple, yet comprehensive, graphical view of the overall historical variability associated with streamflow in a river basin.

FDCs have a long tradition in hydrology. Vogel and Fennessey (1995) provide a detailed discussion of possible utilization of FDCs to graphically address a number of water-related problems, such as hydropower generation, water supply, irrigation planning, waste-load allocation, sedimentation studies, and habitat suitability. This is merely an initial list as the possibilities to practically utilize FDCs are uncountable, because of their ability to condense a wealth of hydrologic information into a single graphic image, and to convey complex hydrologic information to people who may or may not have a background in hydrology (see e.g. Vogel and Fennessey, 1995).

Several countries and geographical areas worldwide lack streamflow observations. This condition led to the formulation of a number of procedures for regionalising FDCs (see e.g. Furness, 1959; Fennessey and Vogel, 1990; Castellarin et al., 2004a; Mendicino and Senatore, 2013). The common objective of these studies is the estimation of FDCs at ungauged river basins, or the enhancement of empirical FDCs constructed for streamgauges where only a limited number of years of streamflow observation is available.

A rough classification of the available regionalisation procedures into two broad categories distinguishes between parametric and non-parametric procedures. The former adopt a parametric representation of the curves and relates the parameter values to a suitable set of physiographic and climatic descriptors of the catchment, the latter identify the curves as non-parametric objects (numerical or graphical) and pool the empirical curves together according to the similarity among catchments (see e.g. Castellarin et al., 2004a; Smakhtin et al., 1997; Ganora et al., 2009). The main

* Tel.: +39 051 209 3365; fax: +39 051 209 3140.
E-mail address: attilio.castellarin@unibo.it

http://dx.doi.org/10.1016/j.jhydrol.2014.03.050
0022-1694/

practical limitation of parametric approaches is the need for mathematical expressions with several parameters to accurately represent the highly non-linear relationship between daily streamflows and duration (i.e. expressions with 4 or more parameters are frequently adopted, see e.g. LeBoutillier and Waylen, 1993; Castellarin et al., 2004b). Regressing these parameters against catchment descriptors sometimes results in regional models characterized by low robustness and high uncertainty (see e.g. Castellarin et al., 2007). This is not a matter of concern when non-parametric regionalisation approaches are used, since the non-linear relationship between streamflow and duration is predicted on the basis of empirical relationships observed at hydrologically similar catchments.

Concerning statistical regionalisation of hydrological information, regional frequency analysis of hydrological extremes (e.g., floods; low-flow indices; rainstorms; etc.) has been a topical issue in hydrology since the 1960s, and the international scientific community actively and fruitfully worked in particular on the problem of delineation of homogeneous regions or pooling-groups of sites. During the evolution of the regionalisation techniques the definition of homogeneous pooling-groups of sites evolved from fixed and geographically identifiable regions to pooling-groups of sites delineated in terms of hydrological affinity with the site of interest (i.e. focused pooling or Region-of-Influence, RoI, approach, see e.g. Burn, 1990; Reed et al., 1999; Ouarda et al., 2001). The most recent approaches dispense with the identification of homogeneous pooling-group of sites altogether, in fact they statistically interpolate the hydrometric index of interest on the stream-network (see Topological kriging or Top-kriging, Skøien et al., 2006; Castiglioni et al., 2011) or on a two-dimensional space of catchment descriptors (i.e., Canonical Kriging or CK, Chokmani and Ouarda, 2004; Castiglioni et al., 2009, 2011). As the literature reported on improvements associated with the evolution of the regionalisation techniques, the state-of-the-art moved forward.

Considering the regionalisation of FDCs the situation is rather different. Studies adopting focused-pooling approaches (see e.g. Holmes et al., 2002) predate recent applications based on the identification of fixed and contiguous regions (see e.g. the studies performed by Mohamoud (2008) for a Mid-Atlantic Region of US, by Viola et al. (2011) for Sicily, or by Niadas (2005) for a Western-Northwestern region in Greece), or application of clustering algorithms (e.g. Sanborn and Bledsoe (2006) in Colorado, USA, or Lin and Wang (2006) in Southern Taiwan, or Ganora et al. (2009), in Swiss and Italian Alps, who propose an original clustering metric designed to express similarity between empirical FDCs through similarities between catchment descriptors).

The scientific literature does not report an ultimate recommendation on how to select the most suitable approach, which remains therefore an open problem in FDC regionalisation. Moreover, regionalisation of FDCs appears to be less advanced than regional flood frequency analysis in terms of conceptualization and formalization of hydrological homogeneity. This paper tackles this specific problem by developing a kriging-based technique to predict FDCs in ungauged basins. As in Ganora et al. (2009), the attention is focussed on the prediction of FDCs standardized by Mean Annual Flow (MAF), which represent a topical issue in hydrology and one of the main steps of all regionalisation procedures based on the index-flow concept (e.g. 1st step: prediction of a standardized curve; 2nd step: prediction of the index streamflow, such as MAF, see Castellarin et al., 2004b, and references therein). In particular, the study proposes an original non-parametric regional approach that predicts dimensionless FDCs in ungauged basins by performing a smooth statistical interpolation of empirical dimensionless FDCs constructed for hydrologically similar gauged basins. The approach dispenses with a preliminary identification of hydrologically homogeneous sub-regions and extends the main idea behind Canonical Kriging, CK, which was originally proposed by Chokmani and Ouarda (2004) for predicting flood quantiles in ungauged basins. CK differs from a standard application of kriging because instead of interpolating the quantity of interest in the geographical coordinates it refers to a two-dimensional physiographic space, in which x and y are functions of morphological and climatic characteristics of the catchment. Hundecha et al. (2008) show a successful application of CK for a regional prediction of rainfall–runoff model parameters at ungauged watersheds, showing through a comprehensive split-sample validation in space and time that the performance of the procedure is comparable with the standard calibration of rainfall–runoff models. Also, CK has been proven to be suitable for predicting low-flow indices in ungauged basins (Castiglioni et al., 2009, 2011), outperforming regional multiregression models.

The study addresses two main science questions. (a) Can Canonical Kriging be adapted for predicting FDCs in ungauged basins? (b) How does such an adaptation of CK perform compared to other regional approaches for predicting FDCs in ungauged sites? An extension of CK termed three-dimensional Canonical Kriging (3*DCK*) is presented here which predicts FDCs in ungauged sites by performing an interpolation of empirical curves in a 3D space xyz, where x and y are functions of the available catchment descriptors, while z is a function of the streamflow duration. Also, the study presents a regional application of the proposed approach for two rather wide geographical Italian regions. In particular, the approach is first implemented for a set of Apenninic catchments located in Central Italy, and then compared against a benchmark regional model developed in a previous study for the same area. Then, a second application of the approach that adopts the implementation scheme developed for the Apenninic case study is presented for an Alpine region in Northern Italy, and the performances of the approach are again assessed through a comprehensive cross-validation procedure.

Without loss of generality, the analysis focuses on the regionalisation of standardized FDCs of daily streamflows, obtained by dividing empirical long-term flow-duration curves by the corresponding MAF, but the procedure is suitable for predicting other kinds of standardized flow-duration curves, such as annual flow-duration curves of daily (as well as other time intervals) streamflow (see e.g. Vogel and Fennessey, 1994; Castellarin et al., 2004b). As previously highlighted, such standardization of FDCs is a preliminary step for several regionalisation approaches proposed in the scientific literature (see e.g. Fennessey and Vogel, 1990; Castellarin et al., 2004a; Ganora et al., 2009, where an interested reader can also find indications on possible approaches to predict MAF in ungauged catchments). In the remainder of the study standardized FDCs, or simply FDCs, will therefore indicate dimensionless FDCs of daily streamflow, unless otherwise specified.

2. Predictions in ungauged basins through statistical interpolation

2.1. Canonical Kriging

Canonical Kriging (CK, also referred to as Physiographic-Space Based Interpolation, PSBI, see Chokmani and Ouarda, 2004) is a Best Linear Unbiased Estimator (BLUE) that spatially interpolates the hydrometric index of interest in a two-dimensional space of physiographic and climatic descriptors, the so-called physiographic space (see Chokmani and Ouarda, 2004). The two orthogonal coordinates x and y of the physiographic space are derived from a set of $n > 1$ geomorphoclimatic catchment descriptors through the application of multivariate techniques such as the Principal