Journal of Hydrology 528 (2015) 177–191

Contents lists available at ScienceDirect

Journal of Hydrology

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

Interpolation of monthly runoff along rivers applying empirical orthogonal functions: Application to the Upper Magdalena River, Colombia

Lars Gottschalk[a,*], Irina Krasovskaia[b], Efrain Dominguez[c], Fabian Caicedo[c], Andres Velasco[c]

[a] Department of Geosciences, University of Oslo, P.O. Box 1022, Blindern, N-0315 Oslo, Norway
[b] IGC, Lergravsvägen 33, SE-26352 Lerberget, Sweden
[c] Pontificia Universidad Javeriana, Departamento de Ecología y Territorio, Colombia

ARTICLE INFO

Article history:
Received 10 January 2015
Received in revised form 10 June 2015
Accepted 12 June 2015
Available online 18 June 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Niko Verhoest, Associate Editor

Keywords:
Interpolation of discharge
Karhunen–Loève expansion
Empirical orthogonal functions
Spatial variance–covariance

SUMMARY

An approach for interpolation of discharge along rivers is presented applying empirical orthogonal functions. From a theoretical point of view this approach for hydrological applications should be looked upon as a generalisation of a Karhunen–Loève expansion. The eigenvalue problem is then represented by a Fredholm integral equation of the second kind over the spatial domain, which in a hydrological application is defined by the principal drainage area. Most methods for numerical solution of this equation entail reduction to an approximately equivalent algebraic problem. The problem is to correctly account for drainage areas and drainage patterns, as the irregularity in sizes of drainage areas might violate the orthogonality if not accounted for correctly. A basic solution to the problem is developed and demonstrated on discharge observations from the Upper Magdalena drainage basin, Colombia. Seven discharge time series are used to determine the empirical orthogonal functions and principal components and four series are used for validation of the application for gap filling and estimation at ungauged sites. The results show a high accuracy for larger basins while those for mountain headwater stations are moderately good. For these latter stations the gain of a short period of observations compared to no observations at all is 5% increase in the coefficient of determination. The results confirm the plausibility of the theoretical approach.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The principal objective of this paper is to develop a model for estimation of monthly streamflow at ungauged sites and for gap-filling based on Karhunen–Loève expansion, and specifically empirical orthogonal functions (eof), with due consideration that observations represent values with spatial support structured along stream networks. This paper builds on an earlier sequence of publications where the focus is the "mapping" of streamflow characteristics along rivers by what we prefer to name a hydrostochastic approach (Gottschalk, 1993a,b; Sauquet et al., 2000; Yan et al.,2012a,b; Gottschalk et al., 2006, 2011, 2013). The aim of this approach is to adapt the theory of random processes so that basic process oriented hydrological laws are accounted for. The correct consideration of the spatial support of runoff data was already mentioned. Furthermore, the water balance equation along rivers should be satisfied, and the variance – covariance structure should correctly account for processes related to the drainage basin dynamics. The novelty is to apply this approach for regional interpolation of runoff time series to ungauged sites and for filling gaps in existing series.

The theory of random processes is the background for the characterisation of variability of hydrological observation series across space and time (Gottschalk, 2005). The characterisation of a random process is commonly done partially in accordance with three different schemes:

(i) Characterisation by distribution function (one dimensional).
(ii) Second moment characterisation.
(iii) Karhunen–Loève expansion, i.e. a series representation in terms of random variables and orthogonal deterministic functions.

The third scheme is in focus in this study. It is a tool for analysing spatio-temporal variability relative to the first and second order moments in terms of new sets of common orthogonal

---

* Corresponding author. Tel.: +47 (0)42331006.
E-mail address: igc_lg@telia.com (L. Gottschalk).

http://dx.doi.org/10.1016/j.jhydrol.2015.06.029
0022-1694/© 2015 Elsevier B.V. All rights reserved.

random functions (time series). The analysis starts by characterising the spatial pattern of the mean and variance and of the spatial covariance of available hydrological observation series over a common period of time within a larger drainage basin. Having completed these first steps the orthogonal functions are determined and can be applied for interpolation and gap filling.

The type of analysis performed in this study is sometimes named principal component analysis (pca). The basic mathematical operation extracting the eigenvalues and eigenvectors from a variance–covariance or correlation matrix was invented by Karl Pearson in 1901 (Pearson, 1901). The generality of the method might be a strength for many applications but with caution. The distinction made here is that when referring to a Karhunen–Loève expansion and eof the eigenvalue problem is derived as integral equation across space. Numerical approximation of this integral equation may lead to the same discrete matrix formulation as for pca but only under certain circumstances and not for the general case of discharge data with spatial support. In the case of application of pca to spatial data in meteorology Buell (1971) pointed to the very strong geometrical elements that can be used advantageously but which are missing in the pca matrix formulation. His recommendation was therefore an integral equation formulation of the problem. The fact that in meteorology values are obtained from measurements at discrete points in space (irregularly and perhaps sparsely located) might be a practical limitation to the numerical solution of the problem. The details for application to discharge observations are further developed below.

The "proper orthogonal decomposition theorem" (Loève, 1945; Karhunen, 1946) embraces predetermined deterministic orthogonal functions like harmonic functions and wavelets, and functions determined from the data set itself i.e. eof. The original theorem is developed for a random function in one dimension. Eof might be seen as a generalisation of this theorem that applies to a spatio-temporal random process (Lorenz, 1956; Obukhov, 1960; Holmström, 1963, 1970). Since the introduction of eof in atmospheric research it has been widely used to analyse atmospheric data. The major reason for this is that it allows a space display and a time display relevant to climate researchers. The objectives have been twofold: to reduce the dimensionality of the system by retaining a much smaller set of domain patterns and to obtain a few, and in some cases only one, leading patterns of variability that are physically relevant (Hannachi et al., 2006). By construction, eofs of climate data yield sets of orthogonal spatial patterns and sets of uncorrelated time series.

Eof/pca is also a well-established method in hydrological research and application. It has been used for several purposes: analysis and classification of patterns (flow regimes) (Bartlein, 1982; Gottschalk, 1985; Krasovskaia et al. 1999, 2003a,b), interpolation of precipitation fields (Creutin and Obled, 1982; Obled and Creutin, 1987), interpolation and gap filling of runoff (Rao and Hsieh, 1991; Hisdal and Tveito, 1992, 1993; Sauquet et al., 2000, 2008), simulation of spatial fields (Braud 1990; Braud and Obled, 1991; Braud et al., 1993; Kitterød and Gottschalk, 1997), and analysis of spatial drought (Krasovskaia and Gottschalk, 1995). The specific nature of discharge data having a support and being ordered along rivers is mostly neglected in publications related to the interpolation of runoff.

Hydrological networks are commonly sparse and do not fully satisfy the demands in discharge data. Furthermore, discharge records are commonly incomplete time series, containing gaps of random and/or systematic character. The data set used in the present study for the upper part of the Magdalena River, the major river of Colombia, offers a typical example. The total basin area of the upper part is 11,998 km^2 and the number of gauging stations is 11. The size of the gauged basins varies between 321 and 11,998 km^2. Sporadic recordings are available since 1971 but more complete ones from 1980 until 2007. This latter sample of theoretically 3696 monthly values contains 712 gaps (19%). Interpolation of runoff to ungauged sites and filling in gaps in existing records are therefore fundamental tasks for this case and in applied hydrology in general.

Section 2 of the paper describes the basic data set from the Upper Magdalena River. Specifically, a denested set of streamflow data is estimated, which will allow summation and averaging along river branches in later steps. The basic theoretical expressions for eof to apply to discharge data and satisfy hydrological laws are reformulated in Section 3. Pcs are estimated for the common period based on the denested sample spatial variance–covariance matrix in Section 4. Interpolation of discharge along rivers is elaborated in Section 5. Finally, in Section 6, a validation of the interpolation scheme is performed.

2. Upper Magdalena River basin

The Magdalena River (*Río Magdalena*) is the principal river of Colombia, flowing northward about 1528 km through the western half of the country. Its drainage basin covers a surface of 273,000 km^2, which is 24% of the country's area, and where 66% of its population lives. The Magdalena River has its origin the Colombian Andes mountain ranges in the valley located between the Cordillera Central and Cordillera Oriental, and it is this upper part of the river that is the target for the present study (Fig. 1).

There are eleven gauging stations in the Upper Magdalena drainage basin (Table 1, Fig. 1). As mentioned earlier, the total area of the basin is 11,998 km^2. The five gauged headwater basins vary in size between 269 and 564 km^2. This defines a reasonable resolution at which we are able to resolve the spatial variability of streamflow across space. The total basin area was divided into 26 basic units having an average size of 460 km^2 in accordance with this scale of resolution (Fig 1, map b). The starting point for this selection was the average area of all small headwaters. The eleven drainage basins were then all divided into an integer number of units having approximately this average size. A simplified river network adapted to this set of basin units, also shown in the figure, will be the background for mapping the streamflow characteristics herein. In the diagram in Fig. 2, graph a), the drainage area along this simplified river network with its nine basic tributaries is shown as a function of the distance from the basin outlet. These tributaries are numbered from left to right. The first four join with the Magdalen River, while the rest flow first into the Paez River that later joins with Upper Magdalena near the outlet station for this study. In the diagram the position of streamflow gauging sites (rings) and outlets of the basic units (crosses) are marked. This principle for displaying information in a diagram against distance to the outlet will be used throughout this study. It is a convenient way to display different characteristics and it allows checking the consistency in the information along rivers.

The digital elevation model used to identify rivers and water dividers has a resolution of 250 m. In principle it is possible to map discharge characteristics to this scale but it would give a totally false picture of the accuracy of the map. A proper balance must exist between the resolution of the map and the scale that the available hydrological observations may resolve.

Most of the eleven gauging stations have more or less complete discharge records since 1980. The number of missing monthly values is indicated in the table. For the period 1980–2007 seven of the stations have complete records, except for some very few missing values, and this common period will be used in the study to test the methodology. The four series with longer periods of missing values will be used for validation, as they will not be used for parameter estimation. Standard statistical parameters are

Download English Version:

https://daneshyari.com/en/article/6410884

Download Persian Version:

https://daneshyari.com/article/6410884

Daneshyari.com