Journal of Hydrology 492 (2013) 102–116

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

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

Can satellite based pattern-oriented memory improve the interpolation of sparse historical rainfall records?

J.P. Matos [a,b,*], T. Cohen Liechti [a], D. Juízo [c], M.M. Portela [b], A.J. Schleiss [a]

[a] École Polytechnique Fédérale de Lausanne (EPFL), Laboratoire de Constructions Hydrauliques (LCH), Lausanne, Switzerland
[b] Technical University of Lisbon (UTL), Instituto Superior Técnico (IST), Lisbon, Portugal
[c] Eduardo Mondlane University (EMU), Faculty of Engineering, Maputo, Mozambique

ARTICLE INFO

Article history:
Received 19 November 2012
Received in revised form 13 March 2013
Accepted 6 April 2013
Available online 18 April 2013
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Hervé Andrieu, Associate Editor

Keywords:
Historical rainfall data
Kriging
Least-Squares Support Vector Regression
Pattern-oriented memory
Rainfall interpolation

SUMMARY

There is a standing challenge in obtaining long localized records of rainfall data in many large river basins of the developing world. Recent spaceborne instrumentation offers a consistent source of rainfall information, but this information covers only a relatively limited time period. In this context, and given its consistence, a question rises on the potential offered by this new wealth of information to improve our understanding of the rainfall patterns and how to use them in order to alleviate the historical problems of scarcity of observed historical records.

The present research focuses on the interpolation of historical rainfall records over large spatial scales and low availability of observed point data, with distances between measurement points in the order of tenths to hundreds of kilometers and temporal scales ranging from daily to monthly. The main goals of the work are twofold: firstly, to evaluate the potential of using a novel pattern-oriented interpolation technique to learn complex spatial rainfall patterns from satellite data and applying this knowledge in the interpolation of historical rainfall maps; secondly, to assess the performance of the proposed methodology by comparing its results to those of other interpolation techniques suitable for spatially sparse datasets.

The proposed pattern-oriented interpolation technique uses modern data sources to enhance the reliability of the interpolation of historical rainfall areal distributions. Results show that, under given conditions, the pattern-oriented memory class of models can considerably reduce the errors traditionally associated with historical rainfall interpolation at large spatial scales and under low availability of spatial data.

1. Introduction

The ability to reproduce areal rainfall maps from point observations is a demanding challenge and, quite often, one of vital importance in hydrological studies. In the past decades the high variability of rainfall over space and time and how to mathematically cope with it has been a subject of major concern of hydrologists (Lanza et al., 2001).

The task of producing interpolated rainfall maps is a far-reaching problem whose solution must take into account, firstly, the spatial and temporal scales at which the rainfall estimates are available and, secondly, the location and amount of the existing data. In this multivariate problem, the performance of different interpolation methods varies greatly.

Rainfall areal distributions are a valuable input for hydrological models. The calibration and validation of these models become particularly demanding tasks which usually require long series of data (rainfall, discharge, evapotranspiration, etc.). Today, satellite derived rainfall products can provide distributed rainfall estimates that cover the equivalent of a dense measurement grid, in some cases, much denser than the one comprised by the existing rain gauge networks. In such areas, where the traditional rain gauge networks are sparse, it is convenient to resort to satellite derived rainfall estimates directly as input data or, as an alternative, to a combination of satellite, gauge, and eventually, radar data (Pegram et al., 2004; Sinclair and Pegram, 2005).

The African DAms ProjecT (ADAPT) is an interdisciplinary research project aiming to develop a set of methods that help assessing the ecological and socio-economic effects of dams in the Zambezi River Basin in Southern Africa (Fig. 1). Within the scope of ADAPT, robust hydraulic–hydrologic models are valuable tools

* Corresponding author. Address: SHRHA, DECivil, Instituto Superior Técnico, Avenida Rovisco Pais, 1049-001 Lisbon, Portugal. Tel.: +351 96 366 0799; fax: +351 21 849 7650.

E-mail addresses: jose.matos@epfl.ch (J.P. Matos), theodora.cohen@epfl.ch (T. Cohen Liechti), juizo@uem.mz (D. Juízo), mps@civil.ist.utl.pt (M.M. Portela), anton.schleiss@epfl.ch (A.J. Schleiss).

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.04.014

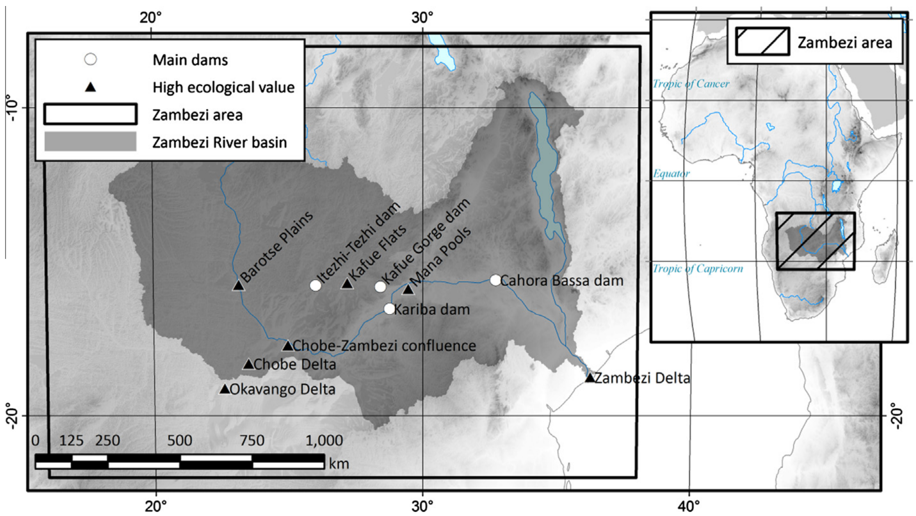

Fig. 1. Zambezi River Basin and study area. Main dams and sites of high ecological value highlighted.

upon which a wide range of water management related analyses, studies and recommendations can be assessed. In the past, several hydrological models aiming at predicting river discharges have been proposed for the Zambezi Basin (e.g. Landert, 2008; Michailovsky, 2008; Vörösmarty and Moore, 1991), but the difficulties related to the use of temporally and geographically sparse hydrologic data, notably rainfall, have not been overcome. As such and as a part of ADAPT, the present research addresses the lack of precise historical rainfall data within the basin, one of the major difficulties modelers and hydrological technicians face as it can negatively affect the quality and accuracy of hydraulic–hydrologic models.

When satellite derived rainfall estimates or the associated combined products are used as rainfall input data the performance and robustness of the hydrological models are constrained by the length of the satellite derived rainfall datasets (near-worldwide reliable estimates have become available from 1998 onwards, with the Tropical Rainfall Measuring Mission, TRMM). This all too often implies that the modeler will not be able to utilize precious historical discharge data that is useful for calibration and validation purposes. The ability to produce accurate areal distributions of historical or "older" rainfall events (denoting the events that predate reliable satellite derived rainfall estimates) would greatly contribute to alleviate this situation.

Today, a plethora of interpolation techniques is able to produce areal rainfall estimates from point measurements. Most of these methods, however, perform the interpolation for a given period in time, generally disregarding the full set of available records or, at best, doing so implicitly, and yield suboptimal results when distances between point measurements are large and short time steps are used. Explicitly using the full set of available records to perform the rainfall interpolation task could lead to improved performances, particularly so when considering the large amount of information that is contained in satellite derived rainfall estimates.

In the context of this research, pattern-oriented memory refers to the ability to recall (identify and reproduce) complex non-linear patterns from related historical rainfall events; models that can display pattern-oriented memory are, therefore, an interesting approach as they can use the full set of satellite derived rainfall estimates in order to perform each interpolation task. This newly proposed method employs asynchronous information derived from areal rainfall estimates in addition to information from the data points that serve as the basis for the interpolation, applying regression techniques to produce its estimates. Its potential to achieve better performances than established interpolation methods is mainly due to its feature of explicitly making use of more data, the non-linear behavior and the capacity to recall recorded observations.

In order to produce enhanced historical rainfall areal maps one can make use of pattern-oriented models in two phases. In the first phase the model learns the rainfall patterns that are identifiable in the satellite rainfall estimations using only the pixels that overlap gauging stations with historical records. In the second phase, the learnt patterns are recalled by the model and used to interpolate the areal rainfall map with inputs from the historical gauged data. The process is exemplified in Fig. 2.

Once rainfall patterns are "learnt" in the first phase, in order to perform the second phase of the interpolation one must match satellite and rain gauge data. The application of patter-oriented memory models is, therefore, related to the combination of satellite and rain gauge data (Pegram et al., 2004; Sinclair and Pegram, 2005), but only indirectly, as before 1998 there were no reliable satellite estimates to merge with existing rain gauge data. Also, the pattern-oriented memory interpolation models recall patterns by performing a non-linear regression on the satellite rainfall estimates used for training, hence sharing some similarities with regression rainfall data patching techniques (e.g. Makhuvha et al., 1997).

Given the defined problematic, the present research aims to assess the possibility of using modern data sources, such as satellite derived rainfall estimates, to enhance the reliability of interpolated historical rainfall maps through the use of models that display pattern-oriented memory and to compare their results with those of alternative interpolation techniques.

Regarding structure, the manuscript is divided into six main parts. In Section 2 a brief overview of the applied techniques is presented. A summarized introduction to the study area, the Zambezi River Basin, is presented in Section 3. Section 4 is dedicated to the description of the applied data and methodologies. The results and discussion are unfolded in Section 5, where the effects of spatial data availability and time resolution, as the impact of time series lengths are analyzed. Also, considerations on the potential and limitations of the tested models are drawn. Finally, conclusions are presented in Section 6.