Journal of Hydrology 525 (2015) 165–177

Contents lists available at ScienceDirect

Journal of Hydrology

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

A multi-criteria penalty function approach for evaluating a priori model parameter estimates

Koray K. Yilmaz [a,*], Hoshin V. Gupta [b,1], Thorsten Wagener [c,2]

[a] Department of Geological Engineering, Middle East Technical University, Ankara 06800, Turkey
[b] Department of Hydrology & Water Resources, University of Arizona, Tucson, AZ 85721, USA
[c] Department of Civil Engineering, Queens Building, University of Bristol, Bristol, UK

ARTICLE INFO

Article history:
Received 22 July 2014
Received in revised form 2 March 2015
Accepted 4 March 2015
Available online 23 March 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Sheng Yue, Associate Editor

Keywords:
Hydrologic model
A priori parameters
Penalty function
Evaluation
Calibration

SUMMARY

A priori parameterization approaches that improve our ability to provide reliable hydrologic predictions in ungauged and poorly gauged basins, as well as in basins undergoing change are currently receiving considerable attention. However, such methods are typically based on local-scale process understanding and simplifying assumptions and an increasing body of evidence suggests that hydrologic models that utilize parameters estimated via such approaches may not always perform well. This paper proposes a Maximum Likelihood multi-criteria penalty function strategy for evaluating a priori parameter estimation approaches. We demonstrate the method by examining the extent to which a priori parameter estimates specified for the Hydrology Laboratory's Research Distributed Hydrologic Model (via a set of pedotransfer functions) are consistent with the optimal model parameters required to simulate the dynamic input–output response of the Blue River basin. Our results indicated that whereas simulations using the a priori parameter estimates give consistently positive flow bias, unconstrained optimization to the response data results in parameter values that are very different from the a priori parameter set. Moreover, although unconstrained optimization performed best (as measured by the calibration criteria), poor hydrograph simulation performance was evident when evaluated in terms of multiple performance statistics not used in the calibration. On the other hand, the multi-criteria compromise solutions provided improved input–output performance in terms of measures not used in calibration, with generally more consistent behavior across calibration and evaluation years, while maintaining physically realistic a priori values for most of the model parameter estimates; adjustments were found to be necessary for only a few key model parameters.

1. Introduction

The ability to provide reliable (accurate and precise) predictions in ungaged and poorly gauged basins, as well as in basins undergoing change, is a major challenge currently receiving considerable attention from the hydrological community (Wagener et al., 2010). To provide meaningful support for decision making in such locations, improvements are needed in (a) collection of data about the system driving forces (i.e. precipitation, potential evapotranspiration), (b) development/selection of the model structure, and (c) prescription of the parameter values. The modeling community has responded in various ways to these challenges, including: (1) to improve model identifiability by building either more parsimonious models (Beck, 1987; Young, 1992; Clark et al., 2008) or more physically-based models (Refsgaard and Storm, 1995; Smith et al., 1995; Qu and Duffy, 2007); (2) to explore novel data sources (Seibert and McDonnell, 2002; Yilmaz et al., 2005; Ruhoff et al., 2013) and information extraction techniques (Wagener et al., 2003; Yilmaz et al., 2008; Pechlivanidis et al., 2012); and (3) to formulate a priori parameterization approaches to specify parameter values from observable watershed characteristics and data (e.g. Koren et al., 2000; Leavesley et al., 2003; Yadav et al., 2007; Pokhrel et al., 2008).

Each of these approaches is important to our understanding of hydrologic processes, and thereby to guiding further improvements in our capacity to predict hydrologic variables. In this paper we focus on the third of the approaches mentioned above, and propose a Maximum Likelihood multi-criteria penalty function strategy for evaluating a priori parameter estimation approaches. We

* Corresponding author. Tel.: +90 312 2102689; fax: +90 312 210 5750.
E-mail addresses: yilmazk@metu.edu.tr (K.K. Yilmaz), hoshin@email.arizona.edu (H.V. Gupta), thorsten.wagener@bristol.ac.uk (T. Wagener).
[1] Tel.: +1 520 626 9712; fax: +1 520 621 1422.
[2] Tel.: +44 (0)117 331 5737.

http://dx.doi.org/10.1016/j.jhydrol.2015.03.012
0022-1694/

demonstrate the method by examining the extent to which a priori parameter estimates specified via a set of pedotransfer functions (Koren et al., 2000, 2003, 2004) for the Hydrology Laboratory's Research Distributed Hydrological Modeling System (here after called the HL model) developed by the National Weather Service (NWS) are consistent with the *'optimal'* model parameters required to simulate the dynamic input–output response of the Blue River basin. Note that the HL model is a spatially distributed version of the Sacramento Soil Moisture Accounting Model (SAC-SMA, Burnash 1995).

Poorly gauged and ungauged basins are (by definition) data sparse, and reliability of hydrologic predictions therefore relies heavily on our ability to specify the values of the model parameters using a priori methods, without recourse to local model calibration. Additionally, changes in the dynamics of the hydrologic system (e.g. climate, land use/cover, urbanization) necessitate a clear understanding of the linkages between watershed characteristics – and changes within – and the dynamic response of the system, so as to improve predictions under change. A priori methods are those that enable us to derive parameter estimates directly from observable static watershed characteristics, either from the watershed in question – in which case we refer to them as "local" a priori estimates, or through the use of regionalized parameter-to-watershed relationships derived from model calibrations conducted on hydrologically similar gauged watersheds – in which case we refer to them as "regionalized" a priori estimates.

It is often claimed that the "local" a priori approach to model parameter specification is primarily useful for "physically-based" hydrological models having spatially distributed components. Because such models are derived from a theoretical understanding of hydrological processes, their parameters can (in principle) be inferred from observable watershed characteristics such as soil type and distribution, topography and land cover (Refsgaard and Storm, 1995; Woolhiser et al., 1990; Leavesley et al., 2003; Qu and Duffy, 2007). However, because the underlying theoretical developments are typically based on experiments conducted at scales far smaller (several centimeter-squares to meter-squares) than that of the model unit (e.g. a several kilometer-square grid cell), difficulties can arise when parameter values inferred in this way are embedded in the larger-scale model grids without proper account for scaling and emergent processes (Bloschl and Sivapalan, 1995; Sivapalan, 2005). In practice it may therefore be necessary to modify/refine parameter estimates inferred using the local a priori approach so that their "effective" values reflect relevant characteristics of the physical phenomena arising from heterogeneous structure and organization of the landscape at the scale of the model grid (Duan et al., 2001; Beven, 1989; Mertens et al., 2004; Pokhrel et. al., 2012).

Another difficulty in implementation of spatially distributed hydrologic models is the high dimensionality of their parameter space, which often result in an ill-posed/poorly conditioned optimization problem. Regularization strategies are often utilized to improve conditioning of the optimization problem by inclusion of additional information (Tikhonov and Arsenin, 1977; Doherty and Skahill, 2006; Pokhrel et al., 2008). By using prior information related to the parameters, regularization is able to better condition the objective function response surface, either through formulation of a penalty function – while retaining the original parameter search space – (Tikhonov and Arsenin, 1977; Carrera and Neuman, 1986; Doherty and Skahill, 2006) or by imposing constraints that reduce the dimensionality of the parameter search space, for example through equations that relate parameters to each other or to preference values (Tonkin and Doherty, 2005; Pokhrel et al., 2008; Pokhrel and Gupta, 2010; Samaniego et al., 2010). A very simple form of the latter regularization approach is the multiplier approach in which spatially distributed a priori values within a specific parameter grid are adjusted through the use of a single multiplier, which is then subject to calibration (Yatheendradas et al., 2008; Yilmaz et al., 2008). The multi-criteria penalty function methodology presented here utilizes both of the regularization strategies discussed above.

The "regionalized" approach to model parameter specification (Abdulla and Lettenmaier, 1997; Sefton and Howarth, 1998; Wagener and Wheater, 2006; Oudin et al., 2008; Samaniego et al., 2010; He et al., 2011) was established mainly in the context of conceptual hydrological models and involves the development of regional regression relationships between the model parameter values estimated for a large number of gauged basins (via calibration) and observable watershed characteristics (i.e. landcover and soil properties) at those locations. The idea is that these relationships can be used to infer parameter estimates for "hydrologically similar" ungauged basins, given knowledge of their observable watershed characteristics. A major assumption of the regional a priori approach is that the calibrated model parameters are uniquely and clearly related to observable watershed properties. This assumption can be difficult to justify when many combinations of parameters are found to produce similar model responses due to the existence of parameter interaction (Duan et al., 1992; McIntyre et al., 2005), measurement uncertainty (Kavetski et al., 2003) and model structure uncertainty (Beven, 2005; Renard et al., 2010), and can result in ambiguous and biased relationships between the parameters and the watershed characteristics (Wagener and Wheater, 2006). Of course, "local" and "regionalized" a priori parameter specification approaches each have their own characteristic strengths and limitations (Wagener and Montanari, 2011). The former represents a mechanistic approach to parameterization by incorporating information about the hydrological properties of static watershed physical characteristics, while the later represents an empirical approach to parameterization by incorporating information inferred from the dynamical input-state-output response characteristics of the watershed (via calibration).

The objective of this study is to exploit information from both mechanistic and empirical approaches in the specification of model parameters, and to utilize a multi-criteria approach for evaluating the extent to which the information inferred from each of the two sources is consistent (agrees or disagrees). The proposed approach is specifically valuable for poorly gauged basins where long-term response data is not available. The methodology is applied to the HL model with two assumptions: (1) the basin is poorly gauged with only short term response data being available, and (2) pedotransfer functions incorporate sufficient information regarding static watershed characteristics into the estimation of a priori parameter values. Understanding gained from this analysis can be used in the development of corrective measures aimed at improving the ability of the model to make accurate hydrologic predictions in gauged and poorly gauged basins as well as basins undergoing change by maximizing the physical integrity of the model parameters.

The paper is organized as follows. Section 2 discusses the proposed multi-criteria penalty function methodology. Sections 3 and 4 apply the method to the HL model for the Blue River basin. Section 5 discusses the conclusions and presents suggestions for future work. Note that the purpose of this work is not to criticize either the HL model or the a priori parameter estimation methodology implemented, but to use these as a vehicle for exploring the multi-criteria penalty function methodology proposed in this paper.

2. Multiple-criteria penalty function approach

The first step in implementation of a hydrologic model for a specific watershed is to evaluate the performance of the model