Environmental Modelling & Software 25 (2010) 691–701

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Exploring uncertainty and model predictive performance concepts via a modular snowmelt-runoff modeling framework

Tyler Jon Smith*, Lucy Amanda Marshall

Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT 59717, United States

ARTICLE INFO

Article history:
Received 29 April 2009
Received in revised form
18 August 2009
Accepted 11 November 2009
Available online 6 January 2010

Keywords:
Modeling framework
Model comparison
Bayesian inference
Markov chain Monte Carlo simulation
Calibration
Uncertainty analysis
Snowmelt
Predictive performance

ABSTRACT

Model selection is an extremely important aspect of many hydrologic modeling studies because of the complexity, variability, and uncertainty that surrounds the current understanding of watershed-scale systems. However, development and implementation of a complete precipitation-runoff modeling framework, from model selection to calibration and uncertainty analysis, are rarely confronted. This paper introduces a modular precipitation-runoff modeling framework that has been developed and applied to a research site in Central Montana, USA. The case study focuses on an approach to hydrologic modeling that considers model development, selection, calibration, uncertainty analysis, and overall assessment. The results of this case study suggest that a modular framework is useful in identifying the interactions between and among different process representations and their resultant predictions of stream discharge. Such an approach can strengthen model building and address an oft ignored aspect of predictive uncertainty; namely, model structural uncertainty.

1. Introduction

Hydrologic model development covers an extensive range from black-box and empirical approaches to physics-based methods. However, neither black-box nor physics-based models are ideal due to the difficulties that surround each; black-box models lack the realism necessary to be made useful and reliable across settings (Donnelly-Makowecki and Moore, 1999), while physics-based models require theoretical understanding of a complex, changing system (Grayson et al., 1992). Because of these complications, a class of models that attempts to blend the simplicity of black-box models with some of the realism of physics-based models has become more prominent (for example, Atkinson et al., 2002; Farmer et al., 2003; Jothityangkoon et al., 2001; Littlewood et al., 2003). These conceptual models often follow a downward approach to model development (Sivapalan et al., 2003), starting with inclusion of the (perceived) most dominant drivers of the system, while lumping the less important processes into effective parameters requiring estimation via calibration (Klemeš, 1983).

Model selection is perhaps the single most important and often least scrutinized component to the precipitation-runoff modeling problem, especially in conceptual models where parameters are calibrated and exact process descriptions are not assumed. All too often a model structure is selected *a priori* with the objective of proving its appropriateness to the study location, with little or no consideration given to alternate structural formulations (Neuman, 2003). This type of approach often results in a model structure being deemed as either acceptable or unacceptable, but it does not address where the model lies on the spectrum (i.e., how acceptable or unacceptable it is).

Although commonly overlooked, some recent studies are beginning to incorporate this critical component; Fenicia et al. (2008), for example, introduced a modeling study that aimed to improve a conceptual model structure through iterative improvement of process representation (i.e., top-down model improvement) under a multi-objective scheme. While this work represents an important first step in model evaluation, it is limited by the use of a single model conceptualization under which alternatives are examined. Extending this further, Clark et al. (2008) considered 79 unique structures generated by mixing the individual flux equations of four commonly used rainfall-runoff models. The present study offers a compromise of these two prior studies by considering the interactions between alternate model components, but not individual

* Corresponding author. Department of Land Resources and Environmental Sciences, Montana State University, 334 Leon Johnson Hall, P.O. Box 173120, Bozeman, MT 59717, United States.
E-mail address: tyler.smith@msu.montana.edu (T.J. Smith).

1364-8152/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2009.11.010

flux representations. However, this study hopes to offer more insights into the relationship that exists between model components that extend beyond the soil moisture accounting context presented by Fenicia et al. (2008) and Clark et al. (2008).

Beyond structural model improvement/assessment strategies, comprehensive precipitation-runoff modeling requires the modeler to consider the additional tasks of calibration, uncertainty analysis, and assessment. Many methodologies exist and have been applied to these essential components, none of which have become universally accepted as the best approach (e.g., Beven et al., 2007; Mantovan and Todini, 2006). Commensurate with these modeling duties, the objective of this study was to examine the relative performance of multiple, alternate model structures featured within a modular snowmelt-runoff modeling framework under a formal Bayesian inferential approach to parameter estimation and uncertainty analysis. The models conceptualize major watershed processes assumed to be important to the hydrologic system. We consider several elements important to the assessment of models within the modeling suite: their individual predictive performance, the uncertainty of individual models and the entire modeling suite, and the ability of the models to simulate variables not considered in the calibration process.

This paper is divided into the following sections: Section 2 presents an overview of the methods employed in this study, Section 3 describes the study location where the methods were applied, Section 4 introduces the modeling framework and each modular component, Section 5 details the assessment of the framework, Section 6 provides a discussion of the results, and Section 7 offers relevant conclusions.

2. Overview of methods

A multi-model approach to precipitation-runoff modeling problems is favorable, according to Neuman (2003), in that it reduces the potential of committing Type I errors (rejecting valid alternative model structures, by omission) and Type II errors (accepting an invalid model structure). The use of a multi-model selection framework not only reduces the potential Type I and II errors, it allows for the inclusion of multiple alternative representations and a means of comparison across the different formulations, in line with a top-down model development philosophy (e.g., Bai et al., 2009). In this study, we consider multiple conceptualizations of elements of the hydrologic system, including snowmelt rates, soil water storage, and runoff routing. Specific details of the modeling framework applied herein are described in Section 4.

The selection of simple conceptual modeling components within the framework necessitates a calibration technique to estimate the effective model parameters. More recently, it has become important in hydrologic studies to recognize that model parameters and simulations are inherently uncertain, as errors arise in using an imperfect model structure and in using inexact data. For the purposes of this study, all of the modular structures associated with the modeling framework were implemented under a Bayesian inferential framework. The use of Bayesian inference in complex modeling problems (like those encountered in hydrology) has gained attention because of the ease with which uncertainties can be considered via posterior parameter distributions (Gilks et al., 1996). However, because the posterior distribution is generally (and in this specific case) analytically intractable, a numerical/statistical estimation approach must be implemented.

Markov chain Monte Carlo (MCMC) simulation is a commonly used statistical method for approximating the posterior distribution, dating to the early 1950s and the development of the Metropolis algorithm (Metropolis et al., 1953). Such methods are now gaining particular popularity in hydrologic studies (e.g., Bates and Campbell, 2001; Kuczera and Parent, 1998; Marshall et al., 2004; Smith and Marshall, 2008). While a variety of different algorithms exist for constructing a random walk Markov chain that converges to the posterior distribution, the Delayed Rejection Adaptive Metropolis (DRAM) algorithm (Haario et al., 2006) was selected for use in this study. The selection of the DRAM algorithm was motivated by past findings of its superiority in hydrologic settings over other MCMC methods by Smith and Marshall (2008), building on the results of Marshall et al. (2004).

3. Study location

Snowpack accumulation and subsequent melt plays a critical role in water resource availability in many mountainous regions. Modeling these processes requires an understanding of snowmelt-runoff dynamics in such areas. For this study, models were implemented for the Stringer Creek watershed located within the Tenderfoot Creek Experimental Forest (TCEF, 46°55′ N, 110°52′ W) of Central Montana, USA (see Fig. 1). TCEF was established by the United States Department of Agriculture (USDA) in 1961 and is committed to the research of subalpine forests influenced by considerable expanses of lodgepole pine (*Pinus contorta*) typically found on the east slope of the northern Rocky Mountains. Average annual precipitation across TCEF is approximately 850 mm per year, largely in the form of snow. For further details on vegetation, climate, and geology of TCEF please refer to Ahl et al. (2008).

The Stringer Creek watershed encompasses 555 ha of the total 3700 ha of TCEF and covers a range in elevations from approximately 1990 to 2430 m. Beyond the natural factors that make Stinger Creek a useful study location, data useful for implementing conceptual-type models are available for this watershed as well. Time series data were selected for the period of study spanning from March 2002 to September 2007, using a 12-h time step. The data used for this study included streamflow observations from the flume at the outlet of Stringer Creek (managed by the Rocky Mountain Research Station, USFS) and meteorological observations (including snow water equivalents) from the Onion Park SNOTEL (2258 m, managed by the NRCS). Additionally, all topographic data was obtained from a 10 m grid digital elevation model (DEM). For additional information on available instrumentation within Stringer Creek and the greater TCEF, please refer to Jencso et al. (2009).

4. Modular hydrologic modeling framework: development

In acknowledgement of the potential shortcomings of a single model approach to precipitation-runoff modeling, a modular hydrologic modeling framework was developed and applied to the Stringer Creek watershed. In order to account for the dominant processes three primary modules (each consisting of alternate conceptualizations) were built into the framework: soil moisture accounting and runoff routing, snowmelt accounting, and the method of semi-distributing precipitation inputs across the watershed area. The framework modules are then combined in all possible manners to create a total of thirty possible "modular structures". Given that each structure is a conceptualization of reality developed under a top-down modeling approach, the modules attempt to characterize the perceived dominant processes while lumping all others into several effective parameters.

4.1. Soil moisture accounting and runoff routing

Representing the water accounting processes of the physical system, three competing model base structures for soil moisture accounting and runoff routing were selected (refer to Fig. 2). The three approaches chosen for this study were a simple bucket model