Journal of Hydrology 420–421 (2012) 216–227

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

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

SAC-SMA *a priori* parameter differences and their impact on distributed hydrologic model simulations

Ziya Zhang [a,*], Victor Koren [a], Seann Reed [a], Michael Smith [a], Yu Zhang [a], Fekadu Moreda [b], Brian Cosgrove [a]

[a] *Office of Hydrologic Development, NOAA/NWS, Silver Spring, MD 20910, USA*
[b] *Water and Ecosystems Management, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, USA*

ARTICLE INFO

Article history:
Received 29 April 2011
Received in revised form 18 November 2011
Accepted 2 December 2011
Available online 13 December 2011
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Vazken Andréassian, Associate Editor

Keywords:
SSURGO
STATSGO
SAC-SMA
A priori parameters
Distributed modeling

SUMMARY

Deriving *a priori* gridded parameters is an important step in the development and deployment of an operational distributed hydrologic model. Accurate *a priori* parameters can reduce the manual calibration effort and/or speed up the automatic calibration process, reduce calibration uncertainty, and provide valuable information at ungauged locations. Underpinned by reasonable parameter data sets, distributed hydrologic modeling can help improve water resource and flood and flash flood forecasting capabilities. Initial efforts at the National Weather Service Office of Hydrologic Development (NWS OHD) to derive *a priori* gridded Sacramento Soil Moisture Accounting (SAC-SMA) model parameters for the conterminous United States (CONUS) were based on a relatively coarse resolution soils property database, the State Soil Geographic Database (STATSGO) (Soil Survey Staff, 2011) and on the assumption of uniform land use and land cover. In an effort to improve the parameters, subsequent work was performed to fully incorporate spatially variable land cover information into the parameter derivation process. Following that, finer-scale soils data (the county-level Soil Survey Geographic Database (SSURGO) (Soil Survey Staff, 2011a,b), together with the use of variable land cover data, were used to derive a third set of CONUS, *a priori* gridded parameters. It is anticipated that the second and third parameter sets, which incorporate more physical data, will be more realistic and consistent. Here, we evaluate whether this is actually the case by intercomparing these three sets of *a priori* parameters along with their associated hydrologic simulations which were generated by applying the National Weather Service Hydrology Laboratory's Research Distributed Hydrologic Model (HL-RDHM) (Koren et al., 2004) in a continuous fashion with an hourly time step. This model adopts a well-tested conceptual water balance model, SAC-SMA, applied on a regular spatial grid, and links to physically-based kinematic hillslope and channel routing models. Discharge and soil moisture simulated using the different set of parameters are presented to show how the parameters affect the results and under what conditions one set of parameters works better than another. In total, 63 basins ranging in size from 30 km^2 to 5224 km^2 were selected for this study. Sixteen of them were used to study the effects of different *a priori* parameters on simulated flow. Simulated hourly flow time series from three cases were compared to hourly observed data to compute statistics. Although the overall statistics are similar for the three different sets of parameters, improvements in simulated flow are observed for small basins when SSURGO-based parameters are used. Fifty-seven basins covering different climate regimes were used to analyze differences in the modeled soil moisture. Results again showed that the use of SSURGO-based parameters generate better soil moisture results when compared to STATSGO-based results, especially for the upper soil layer of smaller basins and wet basins.

Published by Elsevier B.V.

1. Introduction

Hydrologic models typically need to be calibrated in order to achieve the simulation accuracy acceptable for operational river forecasting. Often, different calibrators may derive slightly different parameter data sets due to various factors including data sets and objective functions used in calibration, level of experience and personal approach to calibration. Furthermore, the overall simulation statistics can be similar from different parameter sets in the same basin, reflecting the equifinality concept discussed by many (e.g., Beven, 2006)—yet one set of parameters may be superior and more robust due to greater spatial consistency and more realistic representation of hydrologic processes. Lack of attention to the

* Corresponding author. Address: Office of Hydrologic Development, NOAA National Weather Service, 1325 East West Highway, Room 8376, Silver Spring, MD 20910, USA. Tel.: +1 301 713 0640x158; fax: +1 301 713 0963.
E-mail address: Ziya.Zhang@noaa.gov (Z. Zhang).

0022-1694/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.jhydrol.2011.12.004

physical properties of basins and regional variations can limit the transferability of parameters and the consistency of model performance across basins in a region. This problem is due, in part, to the high levels of uncertainty in the initial parameter values used at the start of the calibration process. Because there are dependences between parameters, if initial parameters are highly uncertain, the calibration results could vary a lot depending on who does the manual calibration. As Kuzmin et al. (2009) indicated in a study of automatic calibration algorithm, with an informative and spatial variability of priori estimated parameters, one can speed up calibration process using one of filtering, i.e., improving the *a priori* estimates based on observed data (typically precipitation and streamflow), rather than one of bounded global optimization as in traditional automatic model calibration. While problematic in a lumped modeling environment, the issue will be of even greater concern with distributed modeling where spatially varying gridded parameter sets are required. With this in mind, better initial parameter estimation for hydrological modeling is important. It can either speed up the calibration process or improve simulations for ungauged basins (Koren et al., 2000; Carpenter and Georgakakos, 2004). By reducing the subjectivity in the calibration process, the resulting model parameters will be more reliable and consistent and will exhibit a reasonable variation of value over a large region or different regions (e.g., Koren et al., 2006). Physically derived initial parameters can help constrain the calibration process and mitigate the issues of data sets and personal approach mentioned above.

With the increased availability of spatially detailed data and computer processing power, and the ever increasing demand for localized information, more and more distributed hydrological models are being developed and applied for research and operational use (Leavesley et al., 1983; Abbott et al., 1986; Wigmosta et al., 1994; Bell and Moore, 1998; Koren et al., 2004; to name a few). Such is the case in the National Weather Service (NWS), where, historically, lumped implementations of the Sacramento Soil Moisture Accounting model (SAC-SMA) have been used for river forecasting. Recently, NWS hydrologists have started using a finer scale, distributed hydrologic model for improved river and flash flood forecasting, as well as for producing prototype gridded soil moisture and temperature products. The system used is the National Weather Service Hydrology Laboratory's Research Distributed Hydrologic Model (HL-RDHM) (Koren et al., 2004). HL-RDHM in this study uses the heat transfer version of SAC-SMA (SAC-HT; Koren et al., 2006) to model rainfall-runoff processes including soil moisture, and kinematic routing for hillslope and channel routing in an hourly, continuous mode for several years.

One of the challenges facing distributed modeling efforts is to have a set of initial parameters that is based on a basin's physical properties, so that either a smaller number of parameters will require calibration, or minimum manual or automatic calibration will be required. Addressing this challenge, Koren et al. (2000) developed a systematic approach to derive eleven SAC-SMA parameters from soil and land use properties. In the initial implementation of the method, they used the State Soil Geographic Database (STATSGO) to derive the parameters for the conterminous United States (CONUS). The STATSGO data are available at a scale of 1:250,000. The soil polygons defined in the STATSGO data set typically range in the size from about 100 to 200 km^2. Although the method of Koren et al. (2000) allows one to account for different land use types, they derived initial CONUS parameters assuming that the land cover/land use across the United States is "pasture or range land use" under "fair" hydrologic conditions. The only spatially variable inputs were soil texture and hydrologic soil group. Subsequent work has shown that when spatially variable land cover data are incorporated into the process, more physically meaningful parameters can be derived (Anderson et al., 2006), although their results were based on lumped simulations on a selected few basins.

While the STATSGO-based gridded parameters provide a good estimate of initial values for distributed modeling as shown in the Distributed Model Intercomparison Project (DMIP) (Smith et al., 2004; Reed et al., 2004; Koren et al., 2004), there are a few shortcomings that limit their application. In addition to the constant land cover and land use assumption in the STATSGO based gridded parameters estimation used in the DMIP, the STATSGO data offer less detailed soil information. A map unit in STATSGO can contain a large number of components. When a distributed model is applied to basins less than 100 km^2 (the case for most flash flood scenarios), the parameters based on 100–200 km^2 soil polygon texture information may not resolve spatial variations within the basin and therefore may not accurately depict runoff process. Serving as a solution to this resolution problem, the Natural Resources Conservation Service (NRCS) also develops and maintains the Soil Survey Geographic Database (SSURGO) data in which the data resolution is approximately 10 times higher than that of STATSGO. The digitization of SSURGO data is nearly complete for most of the CONUS. By using this high-resolution soil data, a new set of gridded SAC-SMA parameters can be derived (Zhang et al., 2011). Based on STATSGO and SSURGO soil data and different land cover assumptions, we can derive three different sets of 11 of the 16 gridded SAC-SMA model parameters. The three different parameter sets are based on (1) STATSGO soil data plus "uniform land cover" assumption (STATSGO ONLY case), (2) STATSGO soil data plus use of variable land cover (STATSGO + LULC case), and (3) SSURGO soil data plus use of variable land cover (SSURGO + LULC case). Because the STATSGO ONLY and STATSGO + LULC cases differ only in their use of land cover data, it is expected that the main differences would be in those parameters associated with the upper zone. In this paper, parameter comparisons between these three sets are presented for the CONUS and selected basins. We will concentrate on the impacts of these different *a priori* parameter sets on hydrologic simulations.

Several published papers, described below, feature comparisons between STATSGO- and SSURGO-based parameters and detail how use of the parameter data sets affects simulated discharge and soil moisture. In comparing outlet stream flow simulations using STATSGO-based and SSURGO-based parameters for the Little Washita watershed (600 km^2) in Oklahoma, Reed (1998) found that there was not much difference between the two cases. Using soils data, Reed (1998) estimated runoff model parameters for the Green and Ampt infiltration equation and a simple percolation model. Part of the reason for the small simulation differences was that the overall surface soil texture distribution, and hence the model parameters defined by the STATSGO and SSURGO data, were similar for this basin. In related research, Anderson et al. (2006) derived basin-averaged STATSGO-based and SSURGO-based SAC-SMA parameters for use by the lumped SAC-SMA model in simulations over several basins within the National Weather Service's (NWS) Ohio River Forecast Center and the West Gulf RFC domains. They found that use of SSURGO-based parameters improved the simulation of basin-outlet flow for basins where there was a noticeable difference in soil texture distributions between STATSGO and SSURGO data sets. The TOPMODEL has also been used to investigate the impact of parameter estimates on simulated streamflow. In particular, Williamson and Odom (2007) used the TOPMODEL for the prediction of streamflow in the South Fork of the Kentucky River near Booneville, Kentucky (area of 1938 km^2) using soil properties from STATSGO and SSURGO data sets. Results show that use of SSURGO-based data produced more accurate streamflow output as compared to the use of STATSGO-based data.