

OPINION ARTICLES

10.1002/2014WR015820

Key Points:

- Complex process-based models have strong a priori constraints
- We provide an example demonstrating strong sensitivity of fixed parameters
- Relaxing strong a priori constraints can help improve hydrology simulations

Correspondence to:

P. A. Mendoza,
pmendoza@ucar.edu

Citation:

Mendoza, P. A., M. P. Clark, M. Barlage, B. Rajagopalan, L. Samaniego, G. Abramowitz, and H. Gupta (2015), Are we unnecessarily constraining the agility of complex process-based models?, *Water Resour. Res.*, 51, 716–728, doi:10.1002/2014WR015820.

Received 7 MAY 2014

Accepted 20 NOV 2014

Accepted article online 27 NOV 2014

Published online 7 JAN 2015

Are we unnecessarily constraining the agility of complex process-based models?

Pablo A. Mendoza^{1,2,3}, Martyn P. Clark³, Michael Barlage³, Balaji Rajagopalan^{1,2}, Luis Samaniego⁴, Gab Abramowitz⁵, and Hoshin Gupta⁶

¹Department of Civil, Environmental, and Architectural Engineering, University of Colorado at Boulder, Boulder, Colorado, USA, ²Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, Colorado, USA, ³Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA, ⁴UFZ Helmholtz Centre for Environmental Research, Leipzig, Germany, ⁵Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, Australia, ⁶Department of Hydrology and Water Resources, The University of Arizona, Tucson, Arizona, USA

Abstract In this commentary we suggest that hydrologists and land-surface modelers may be unnecessarily constraining the behavioral agility of very complex physics-based models. We argue that the relatively poor performance of such models can occur due to restrictions on their ability to refine their portrayal of physical processes, in part because of strong a priori constraints in: (i) the representation of spatial variability and hydrologic connectivity, (ii) the choice of model parameterizations, and (iii) the choice of model parameter values. We provide a specific example of problems associated with strong a priori constraints on parameters in a land surface model. Moving forward, we assert that improving hydrological models requires integrating the strengths of the “physics-based” modeling philosophy (which relies on prior knowledge of hydrologic processes) with the strengths of the “conceptual” modeling philosophy (which relies on data driven inference). Such integration will accelerate progress on methods to define and discriminate among competing modeling options, which should be ideally incorporated in agile modeling frameworks and tested through a diagnostic evaluation approach.

1. Introduction

The hydrologic community has made substantial investments in the development of complex physics-based models that provide detailed representations of the dominant physical processes and their interactions [e.g., Abbott *et al.*, 1986; Wigmosta *et al.*, 1994; VanderKwaak and Loague, 2001; Ivanov *et al.*, 2004; Maxwell and Miller, 2005; Rigon *et al.*, 2006; Qu and Duffy, 2007; Lawrence *et al.*, 2011; Niu *et al.*, 2011]. In spite of their complexity and physical realism, distributed process-based models perform similarly to, or only slightly better than, traditional bucket-style rainfall-runoff models [e.g., Reed *et al.*, 2004; Smith *et al.*, 2012]. In this commentary we discuss some issues that can result in relatively low performance of complex models, illustrate some of these shortcomings through an example application, and make practical recommendations that should lead to improved physics-based model simulations.

2. On the Need for Model Agility

Over the last four decades, a number of important issues related to process representation and model performance have been widely discussed [Freeze and Harlan, 1969; Bergstrom, 1991; Blöschl and Sivapalan, 1995; Beven, 2000, 2002, 2006; Sivapalan *et al.*, 2003; Kirchner, 2006; Clark *et al.*, 2008, 2011; Gupta *et al.*, 2008, 2012; Wagener *et al.*, 2009a; Beven and Cloke, 2012]. A fundamental challenge is developing models that represent how the spatial variability in hydro-meteorological fields, topography, vegetation and soils combines to produce fluxes of energy and water at catchment, regional and global scales [e.g., Reggiani *et al.*, 1999; Beven, 2002; Kollet *et al.*, 2010]. Meeting this challenge requires extensive evaluation and refinement of model representations of hydrological processes [Beven, 2002; Clark *et al.*, 2011], in particular those related to representing fluxes of water and energy at the spatial scale of the model discretization [Samaniego *et al.*, 2010].

However, a factor that complicates the problem of evaluating and refining the behavior of process-based models is that many of them have fixed representations of spatial variability (e.g., a single spatial resolution and configuration, parameter look-up tables with limited number of soil and vegetation classes), fixed representations of model physics (e.g., a single set of process representations), and fixed (hard-coded) model parameter values. Such strong a priori constraints arguably reflect overconfidence in the spatiotemporal representation of physics-based equations describing complex systems, which are heterogeneous across different spatial scales and often poorly characterized by direct measurement [Kirchner, 2006], resulting in models with insufficient ability to adequately simulate the heterogeneity of biophysical and hydrological processes.

In view of such problems, we believe that hydrologic and land surface modeling systems should be *agile* (i.e., have the capability to adjust model equations and parameters to faithfully represent observed processes), in order to enable testing multiple hypotheses of hydrologic behavior [Clark *et al.*, 2011]. Specifically, modeling frameworks should be agile enough to support at least the following key aspects: (i) the capability to modify the representation of spatial variability and hydrologic connectivity (e.g., support different spatial resolutions, grid cell versus hydrologic response units, mosaic versus semitile approach to represent subgrid heterogeneity), (ii) the capability to modify model parameterizations for individual processes (e.g., different soil stress functions for evapotranspiration, nonlinear reservoir versus multiple parallel reservoirs for base flow), and (iii) the capability to modify model parameter values. Furthermore, these features should be extensible to facilitate iterative improvements in the representation of complex systems (i.e., model reconfiguration) as new data that might support new hypotheses becomes available [Son and Sivapalan, 2007; Fencia *et al.*, 2008; Clark *et al.*, 2011].

The need for model agility is increasingly recognized, and many modeling frameworks are now available that facilitate experimenting with competing modeling alternatives [Clark *et al.*, 2011]. For instance, Pomeroy *et al.* [2007] developed the Cold Regions Hydrologic Model (CRHM) to experiment with different alternative representations of cold region processes; Clark *et al.* [2008] developed the Framework for Understanding Structural Errors (FUSE) to test different parameterizations of soil hydrology used in traditional bucket-style rainfall-runoff models; Niu *et al.* [2011] developed the Noah-MP model with the aim to experiment with several model parameterizations of biophysical and hydrological processes used in land-surface models; Essery *et al.* [2013] developed the Joint UK Land Environment Simulator (JULES) Investigation Model (JIM) to test different options to simulate snow processes. Nevertheless, these modeling frameworks lack an integrated supporting system for experimenting with different representations of spatial variability, a broad range of physics parameterizations (i.e., they are *all* somewhat limited in scope), and different model parameter values. For example, FUSE includes only simple parameterizations of soil hydrology, and is focused on spatially lumped structures; JIM is restricted to snowpack processes; and Noah-MP is limited to a semitile grid structure.

We believe that modeling systems addressing at least the three requirements proposed above (capability to modify spatial variability and hydrologic connectivity, capability to modify model parameterizations of individual processes and capability to modify model parameter values) will provide a robust framework for the assessment of differences among process representations in existing hydrological models, and to accelerate future model development and improvement.

3. An Example of Unnecessary Constraints in a Complex Process-Based Model: Treating Uncertain Model Parameters as Physical Constants

Somewhat paradoxically, many physics-based models set uncertain model parameters to fixed values. For example, transfer functions that link measurable properties of the landscape (e.g., clay and sand contents, percent of organic matter) with model parameters (e.g., soil porosity, saturated soil hydraulic conductivity) typically include fixed coefficients. These coefficients should ideally be described by a sampling distribution, as they are commonly obtained through statistical analysis of data samples taken for a given region and spatial domain. Similarly, 'observable' model parameters (e.g., saturated hydraulic conductivity, soil porosity, vegetation height) are defined as single values for each model element. This is problematic because such parameters are difficult to define precisely given large within-element spatial heterogeneity and errors associated with direct and indirect measurement techniques. More worrisome, many of the functional 'free'

```

! runoff parameters used for SIMTOP and SIMGM:
REAL, PARAMETER :: TIMEAN = 10.5 !gridcell mean topographic index (global mean)
REAL, PARAMETER :: FSATMX = 0.38 !maximum surface saturated fraction (global mean)

! adjustable parameters for snow processes

REAL, PARAMETER :: M = 1.0 ! 2.50 !melting factor (-)
REAL, PARAMETER :: ZOSNO = 0.002 !snow surface roughness length (m) (0.002)
REAL, PARAMETER :: SSI = 0.03 !liquid water holding capacity for snowpack (m3/m3) (0.03)
REAL, PARAMETER :: SWEMX = 1.00 !new snow mass to fully cover old snow (mm)
!equivalent to 10mm depth (density = 100 kg/m3)

```

Figure 1. Section of the source code of Noah-MP containing some fixed runoff and snow model parameters.

parameters (e.g., coefficients in conceptual base flow and surface runoff parameterizations) are hard-coded as spatially constant values. Setting model parameters to fixed values effectively treats them as physical constants, neglecting the large uncertainty in their estimates and the large impact that they have on model predictions.

In this section we provide an example of the impact of fixed model parameters through analysis of the Noah Land Surface Model with Multiple Parameterization Options (Noah-MP) [Niu et al., 2011]. We discuss the existence of hard-coded model parameters, and demonstrate how hydrological simulations can be very sensitive to their values.

3.1. Model Description

Noah-MP has been developed as a general representation of large-scale hydrologic and biophysical processes, applicable for the full range of hydroclimatic environments worldwide. The model domain is discretized using a one-layer vegetation model, three-layer snow model, and four-layer soil model. The vegetation module includes snow interception, including loading/unloading, melt/refreeze capabilities, and sublimation of canopy-intercepted snow, along with detailed representations of transmission and attenuation of radiation through the canopy, and within and below-canopy turbulence. The snow module uses a multilayer snowpack with a thin surface layer to simulate liquid water retention, refreezing and snowpack densification. The soil moisture module uses the one-dimensional unsaturated form of Richards' equation for storage and transmission of liquid water in soil, with sink terms for transpiration. Finally, the model includes a representation of permeable frozen soil and an unconfined aquifer that interacts with the soil module.

In this study we use a single suite of physics options for Noah-MP, including a Ball-Berry type model for canopy stomatal resistance, the Community Land Model (CLM) [Oleson et al., 2010] soil stress function to control stomatal resistance, the SIMTOP model for runoff and groundwater [Niu et al., 2005], a Monin-Obukhov similarity theory-based drag coefficient, supercooled liquid water and frozen soil permeability based on Niu and Yang [2006], a two-stream radiation transfer scheme applied only to the vegetated fraction, a snow surface albedo parameterization based on the Canadian Land Surface Scheme (CLASS) [Verseghy, 1991], partitioning of precipitation into snowfall and rainfall based on Jordan [1991] and a Noah-type lower boundary of soil temperature. Readers are referred to Niu et al. [2011] for a full description of each model component.

3.2. Hard-Coded Model Parameters

Although Noah-MP has look-up tables to define soil and vegetation parameter values for different soil and land cover types, it incorporates several hard-coded parameters for snow and runoff processes. This is a typical problem in many complex physics based models, such as the Variable Infiltration Capacity model (VIC) [Wood et al., 1992; Liang et al., 1994, 1996] and the Community Land Model (CLM) [Oleson et al., 2010]. Many other examples can be found in the literature, including models that have hard-wired constants based on limited experimental data.

As an example, Figure 1 displays a section of code wherein Noah-MP developers have commented that several snow and runoff parameters could be treated as "adjustable"; however, adjusting these parameters requires manual alteration of the appropriate lines of code and subsequent recompiling of the model subroutine before a new parameter trial can be conducted. This severely constrains the ability to conduct extensive sensitivity analysis and/or parameter estimation. Similarly, hard-coded parameters can be found in the CLASS snow albedo parameterization (Figure 2), where minimum and maximum snow albedo have

```

----- local variables -----
INTEGER :: IB           !waveband class
-----
! zero albedos for all points
ALBSND(1: NBAND) = 0.
ALBSNI(1: NBAND) = 0.

! when cosz > 0
ALB = 0.55 + (ALBOLD-0.55) * EXP(-0.01*DT/3600.)

! 1 mm fresh snow(SWE) -- 10mm snow depth, assumed the fresh snow density 100kg/m3
! here assume 1cm snow depth will fully cover the old snow
IF (QSNOW > 0.) then
  ALB = ALB + MIN(QSNOW*DT, SWEMX) * (0.84-ALB)/(SWEMX)
ENDIF

ALBSNI(1)= ALB      ! vis diffuse
ALBSNI(2)= ALB      ! nir diffuse
ALBSND(1)= ALB      ! vis direct
ALBSND(2)= ALB      ! nir direct

```

Figure 2. Section of the source code of Noah-MP containing the snow albedo CLASS parameterization.

been set to 0.55 and 0.84 respectively (dimensionless units), and time decay in snow albedo has been set to 0.01 (units of h^{-1}). One would expect these hard-coded parameters to vary regionally and seasonally, and there is no apparent justification for setting the parameter values to globally fixed constants when, in fact, they are subject to large estimation and scaling uncertainties, and therefore more appropriately described by probability density functions.

3.3. Model Performance and Parameter Sensitivity

In this example, we configure Noah-MP to simulate runoff in three headwater catchments in the Colorado River basin: the Yampa River at Steamboat Springs (1468 km²), the East River at Almont (748 km²) and the Animas River at Durango (1819 km²). The predominant land surface cover of these basins is deciduous and evergreen forest, while the hydrology is mainly dominated by snow processes. All hydrologic model simulations were forced using hourly reanalysis outputs from the Weather Research and Forecasting (WRF) model [Skamarock et al., 2008], using the 4 km simulations described by Rasmussen et al. [2014]. The initial and 3 hourly lateral boundary conditions for the WRF runs were taken from the North American Regional Reanalysis (NARR) [Mesinger et al., 2006], whose spatial resolution is 0.3° (~32 km). WRF simulations have been previously validated against SNOTEL sites, and precipitation spatial variability, timing, and intensities are well represented by the model [Ikeda et al., 2010; Prein et al., 2013].

The period for hydrologic model simulations is October 2000 through September 2008 (hourly time steps), and the first 2 years are used as a warm-up period (so that analysis is restricted to October 2002 through September 2008). No horizontal routing of surface overland flow, subsurface flow, or channel flow is performed; instead, basin-average runoff is computed as the average of the 1D (vertical) 4 km model grid cells. Finally, model outputs are aggregated to daily time steps to compute evaluation metrics.

Figure 3a displays scatter plots with runoff model simulations versus observations (period October 2002 to September 2008) at the three basins of interest, using default parameter values. The RMSE values and Nash-Sutcliffe efficiencies indicate that model performance is quite poor, especially at the East River and Animas River basins, and that parameter calibration is a necessary step to improve model fidelity. This point seems obvious for the hydrologic community (especially for applied hydrologists relying on traditional bucket-style rainfall-runoff models), where tremendous advances have been achieved in terms of parameter estimation methods [e.g., Duan et al., 1992; Gupta et al., 1998; Yapo et al., 1998; Vrugt et al., 2003a, 2003b, 2006; Pokhrel et al., 2012], sensitivity analysis [e.g., Tang et al., 2007; van Werkhoven et al., 2008; Foglia et al., 2009; Wagener et al., 2009b; Göhler et al., 2013; Rakovec et al., 2014], ensemble simulation and verification [e.g., Carpenter and Georgakakos, 2004; De Lannoy et al., 2006; Pauwels and De Lannoy, 2009] and parameter uncertainty quantification [e.g., Beven and Binley, 1992; Uhlenbrook et al., 1999; Vrugt et al., 2005; Kavetski et al., 2006a, 2006b; Kuczera et al., 2006; Thyer et al., 2009]. However, use of these techniques is less common in the land surface community, where most attention has been focused on improving process parameterizations, typically using fixed parameter values obtained from the literature.

With the aim to identify which parameters have the largest impact on model predictions, we use the Distributed Evaluation of Local Sensitivity Analysis (DELSA) method [Rakovec et al., 2014] to evaluate the sensitivity of a suite of metrics (Table 2) to variations in the model parameters (Table 1). All parameters listed in Table 1 can be considered 'observable' (i.e., a priori values can be specified by direct measurement or using indirect procedures), except the following (i.e., "free" model parameters): empirical canopy wind parameter (w_{rp}), runoff decay factor (f), base flow coefficient ($R_{sb,max}$), maximum surface saturated fraction (F_{sat}), exponent used in the curves for the melting season (m_s) and the exponent in snow decay albedo relationship

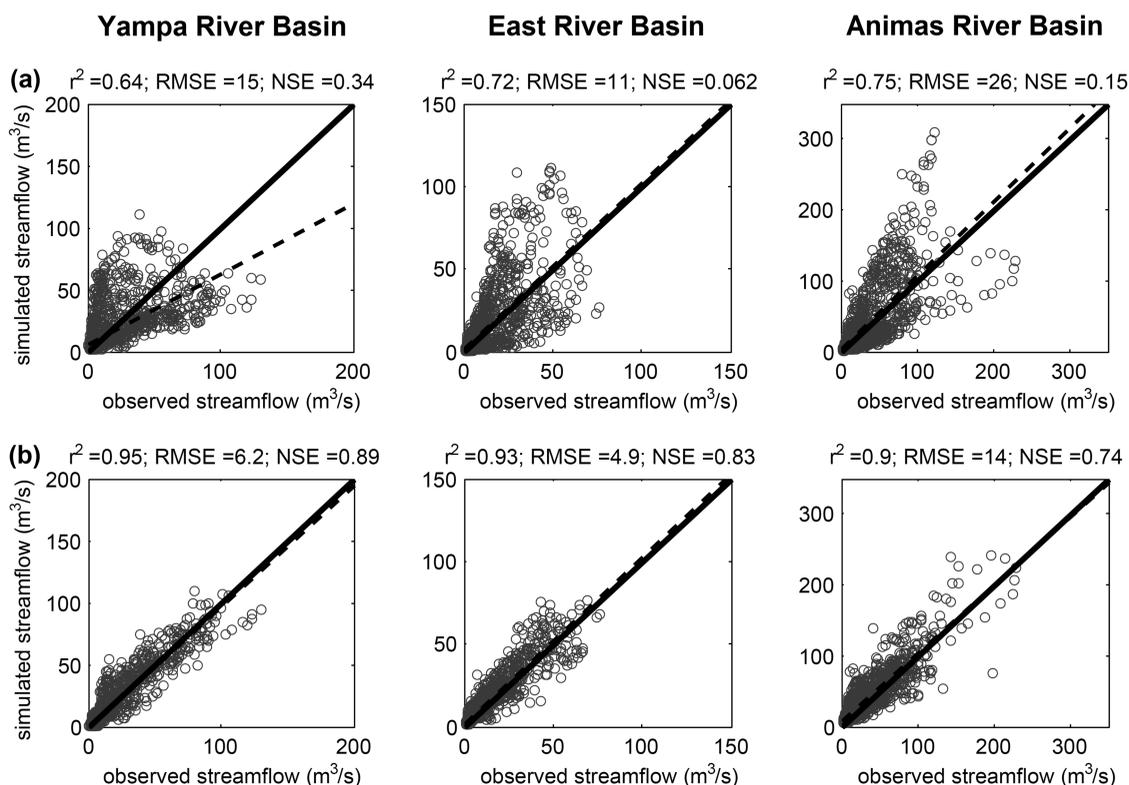


Figure 3. Model streamflow simulations versus observations for the period October 2002 to September 2008 using (a) default parameter values (top row), and (b) calibrated values for six originally hard-coded parameters: f , $R_{sb,max}$, λ_m , m_s , α_{min} and κ (bottom row). The solid line is the 1:1 line, and the dashed line is the linear regression. In all plots, r^2 , RMSE and NSE denote coefficient of determination, root mean squared error and Nash-Sutcliffe efficiency, respectively.

(κ). From these free parameters, five of them were originally hard-coded (f , $R_{sb,max}$, F_{sat} , m_s and κ). It is noteworthy that this DELSA application required modification of the source code in order to ‘uncover’ all runoff and snow parameters (i.e., increase model agility), whose values were originally hard-coded.

The results in Figure 4 demonstrate very high sensitivity for model parameters that were originally hard-coded. Specifically, RMSE is most sensitive to the monthly leaf area index for spring/summer (LAI_{ss}), the runoff decay factor (f), the exponent used in the snow depletion curves for the melting season (m_s) and the exponent in the snow decay albedo relationship (κ). In the case of the runoff ratio (%BiasRR), the most sensitive parameter is f , followed by the Clapp-Hornberger b parameter, the saturated hydraulic conductivity (K_{sat}), the slope of conductance-to-photosynthesis relationship (m_{ps}) and LAI_{ss} . When looking at variations in flashiness of runoff (%BiasFMS), the most sensitive parameters are f and κ , followed by the Clapp-Hornberger b parameter and the empirical canopy wind parameter (w_{rp}). Finally, the sensitivity in runoff seasonality (%BiasCTR) is mostly explained by variations of m_s , κ and the minimum snow albedo, α_{min} (i.e., snow parameters). The reader can also note that, among free parameters, relative differences in sensitivities between formerly hard-coded parameters and those exposed depend on the metric examined. For instance – when looking at RMSE – f , m_s and κ (originally hard-coded) are the most sensitive free parameters, followed by w_{rp} and $R_{sb,max}$ which have similar sensitivity. Nevertheless, w_{rp} becomes the second most sensitive free parameter (after f) when the objective criterion is %BiasRR. Overall, the most sensitive parameters are those that were formerly hard-coded—this result holds for all basins and all objective functions.

In summary, the response to high-intensity precipitation events, flashiness of runoff and seasonality are highly sensitive to variations of snow and runoff parameters (all of them originally hard-coded), while soil and vegetation parameters become more relevant when evaluating model behavior in terms of evapotranspiration processes. This suggests that calibration efforts aimed to improve model fidelity should include some of the hard-coded parameters in Table 1. To test this idea, we perform a simple calibration experiment aimed to adjust six runoff and snow parameters (f , $R_{sb,max}$, λ_m , m_s , α_{min} and κ) using the Shuffled Complex

Table 1. Parameters of Noah-MP Considered in This Example Application^a

Parameter	Description	Units	Distributed ^b	Range		Comment
				Min	Max	
<i>Soil Parameters^c</i>						
<i>b</i>	Clapp-Hornberger <i>b</i> parameter		yes	0.42	1.84	Multipliers obtained from <i>b</i> exponent values in the range 2–15 [Cosby <i>et al.</i> , 1984].
θ_{sat}	porosity	m ³ m ⁻³	yes	0.88	1.14	Multipliers obtained from porosity values in the range 0.35–0.53 [Cosby <i>et al.</i> , 1984], porosity constrained to be larger than field capacity.
Ψ_{sat}	saturated soil matric potential	m m ⁻¹	yes	0.15	2.20	Multipliers obtained from saturated matric potential ranging from 0.02 to 0.78 [Cosby <i>et al.</i> , 1984].
K_{sat}	saturated soil hydraulic conductivity	m s ⁻¹	yes	0.20	9.56	Multipliers obtained from range 5 × 10 ⁻⁷ – 5 × 10 ⁻⁵ for <i>ksat</i> [Cosby <i>et al.</i> , 1984].
k_{qtz}	soil quartz content		yes	0.29	1.37	Multipliers obtained from range 0.1–0.82 for quartz content [Hogue <i>et al.</i> , 2005; Rosero <i>et al.</i> , 2010].
<i>Vegetation Parameters^c</i>						
$Z_{0,veg}$	momentum roughness length	m	yes	0.17	2.39	Multipliers obtained from range 0.01–2.6 m [Dorman and Sellers, 1989; Xia <i>et al.</i> , 2012].
ρ_L	leaf reflectance		yes	0.90	1.10	Multipliers based on average standard deviations reported by Asner <i>et al.</i> [1998].
ρ_S	stem reflectance		yes	0.90	1.10	Multipliers based on average standard deviations reported by Asner <i>et al.</i> [1998].
τ_L	leaf transmittance		yes	0.90	1.10	Arbitrary ±10 % multipliers, constrained by variations in leaf reflectance ($\rho + \tau \leq 1$).
τ_S	stem transmittance		yes	0.90	1.10	Arbitrary ±10 % multipliers, constrained by variations in leaf reflectance ($\rho + \tau \leq 1$).
χ_L	leaf/stem orientation index		yes	0.50	1.67	Multipliers defined such that max. absolute orientation index is 0.5 [Prihodko <i>et al.</i> , 2008].
w_{rp}	empirical canopy wind parameter	m ⁻¹	no	0.18	10	Obtained from Goudriaan [1977].
T_{min}	minimum temperature for photosynthesis	K	yes	1.00	1.03	Multipliers obtained from range 265–281°K [Sacks <i>et al.</i> , 2007].
$V_{max,25}$	maximum rate of carboxylation at 25° C	μmol(CO ₂) m ⁻² s ⁻¹	yes	0.65	1.35	Multipliers obtained from standard deviations reported by [Kattge <i>et al.</i> , 2009].
m_{ps}	slope of conductance-to-photosynthesis relationship		yes	0.67	1.33	Multipliers obtained from the slope range 4–12 [Sellers <i>et al.</i> , 1996; Wolf <i>et al.</i> , 2006].
SAI_{fw}	monthly stem area index, one sided (fall/winter)	m ² m ⁻²	yes	0.10	2.14	Multipliers obtained from stem area index range 0.01–3.0 [Otto <i>et al.</i> , 2011].
SAI_{ss}	monthly stem area index, one sided (spring/summer)	m ² m ⁻²	yes	0.10	1.88	Multipliers obtained from stem area index range 0.01–3.0 [Otto <i>et al.</i> , 2011].
LAI_{fw}	monthly leaf area index, one sided (fall/winter)	m ² m ⁻²	yes	0.10	3.18	Multipliers obtained from leaf area index range 0.01–7 [Dorman and Sellers, 1989; Hastie <i>et al.</i> , 2002; Myneni <i>et al.</i> , 1997].
LAI_{ss}	monthly leaf area index, one sided (spring/summer)	m ² m ⁻²	yes	0.10	1.27	Multipliers obtained from leaf area index range 0.01–7 [Dorman and Sellers, 1989; Hastie <i>et al.</i> , 2002; Myneni <i>et al.</i> , 1997].
<i>Runoff Parameters^d</i>						
f	runoff decay factor	m ⁻¹	no	1.0	10	Based on values reported in Beven [1997].
$R_{sb,max}$	Base flow coefficient	mm s ⁻¹	no	0.5	8	Based on Niu <i>et al.</i> [2005].
λ_m	grid cell mean topographic index		no	7.35	13.65	Variations up to 30 % from the default hard-coded value (10.35).
F_{sat}	maximum surface saturated fraction		no	0.29	0.46	Based on Niu <i>et al.</i> [2005].
<i>Snow Parameters^d</i>						
m_s	exponent used in the curves for the melting season		no	0.5	3	Based on range in Niu and Yang [2007].
$Z_{0,sno}$	snow surface roughness length	m	no	0.0001	0.01	Based on range suggested by Marks and Dozier [1992] and Reba <i>et al.</i> [2014].
θ_{lwc}	liquid water holding capacity for snowpack	m ³ m ⁻³	no	0.01	0.08	Based on ranges in Amoroch and Espildora [1966] and Anderson [1973].
SWE_{new}	new snow mass to fully cover old snow	mm	no	0.5	5	Minimum is 50 % of default value; maximum obtained from Xia <i>et al.</i> [2012].
α_{min}	minimum snow albedo		no	0.45	0.65	Based on Aguado [1985] and Dirmhirn and Eaton [1975].
α_{max}	maximum snow albedo		no	0.70	0.95	Based on Aguado [1985] and Essery and Etchevers [2004].
κ	exponent in snow decay albedo relationship	h ⁻¹	no	0.001	0.1	Based on Essery and Etchevers [2004].

^aThe parameter ranges investigated (columns 5 and 6) were selected based on literature review of the different model components. The explanation of ranges in multipliers (if the parameter is spatially distributed in the basin) or raw values (if the parameter is spatially uniform) is provided in the “Comment” column, together with the associated references.

^bIf the parameter is distributed, its sensitivity is analyzed on the basis of its multipliers. Although description and units refer to actual parameters in Noah-MP, parameter values in bold represent the multiplier values (instead of actual parameters).

^cExposed to users.

^dHard-coded parameters.

Table 2. Objective Functions Included in DELSA^a

Name	Description	Equation
RMSE	Root Mean Squared Error	$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (Q_t^{sim} - Q_t^{obs})^2}$
%BiasRR	Percent bias in runoff ratio	$\%BiasRR = \frac{\sum_{t=1}^N (Q_t^{sim} - Q_t^{obs})}{\sum_{t=1}^N Q_t^{obs}} \times 100$
%BiasFMS	Percent bias in FDC midsegment slope	$\%BiasFMS = \frac{[\log(Q_{m1}^{sim}) - \log(Q_{m2}^{sim})] - [\log(Q_{m1}^{obs}) - \log(Q_{m2}^{obs})]}{[\log(Q_{m1}^{obs}) - \log(Q_{m2}^{obs})]} \times 100$
%BiasCTR	Percent bias in runoff seasonality	$\%BiasCTR = \frac{CTR_{flow}^{sim} - CTR_{flow}^{obs}}{CTR_{flow}^{obs}} \times 100$

^a Q_t^{sim} : simulated flow for time step t , Q_t^{obs} : observed flow for time step t , Q_{m1}^{sim} : simulated flow with exceedance probability of $m_1 = 0.2$, Q_{m1}^{obs} : observed flow with exceedance probability of $m_1 = 0.2$, Q_{m2}^{sim} : simulated flow with exceedance probability of $m_2 = 0.7$, Q_{m2}^{obs} : observed flow with exceedance probability of $m_2 = 0.7$, $CTR_{flow} = \frac{\sum_{i=1}^{365} t_i Q_i}{\sum_{i=1}^{365} Q_i}$: centroid of daily hydrograph for an average water year.

Evolution (SCE-UA) algorithm [Duan et al., 1992, 1993], by minimizing the root mean squared error between observed and simulated daily streamflow (RMSE) for the period 1 October 2002 to 30 September 2008. These parameters were selected because they showed the largest sensitivities for RMSE among formerly hard-coded parameters. The results displayed in Figure 3b clearly demonstrate how the inclusion of these parameters in the calibration process improves model accuracy (e.g., higher NSE and r^2 , and lower RMSE).

3.4. The Physical Basis of Hard-Coded Parameters

Complex models represent physical processes at a fine level of granularity (i.e., detail), and, as such, it is possible to impose much stronger a priori constraints on model behavior in comparison to simpler, conceptual models. Noah-MP explicitly simulates all energy fluxes at the snow surface, as opposed to more parsimonious temperature-index models that represent snowmelt as an empirical function of temperature. Process-based models can therefore simulate accelerated snow melt during rain-on-snow events when turbulent heat fluxes are an important component of the snow-surface energy balance [e.g., Marks et al., 1999], whereas such processes are poorly represented in the temperature-index snow models as the empirical relationships between temperature and snow melt are applied consistently to all snow melt events. While the temperature-index snow models may have fewer parameters, their values can be difficult to constrain correctly because the empirical functions implicitly represent a wide range of physical processes.

Although the strong physical basis of more complex process-based models provides powerful justification for their widespread use, it is important to recognize that empirical functions are widely used in such models at a much finer level of granularity. For example, consider the albedo decay parameterization from the source code illustrated in Figure 2, noting that the albedo decay parameter (κ) is one of the most sensitive model parameters for the criteria examined here (Figure 4). The physical processes affecting decreases in snow albedo over time include rounding and growth of the snow grains, deposition of dust on the snow surface, among others. These physical processes are included in some models [e.g., Jordan, 1991; Flanner et al., 2007], but in Noah-MP the albedo decay rate is set to be constant over time. The lumping of multiple physical processes into a single albedo decay parameter is hence very similar to the lumping of all snow-surface energy fluxes into a single empirical temperature-melt expression, and there is no physical basis to treat time decay in snow albedo as a fixed constant in both space and time. This is a common problem, since other more flexible albedo parameterizations reported in the literature [e.g., Yang et al., 1997] also lack a proper justification for fixing parameters defining the albedo decay rate.

This specific and compelling example underscores a fundamental issue in process-based modeling: it is important to carefully specify the uncertainty of the different model parameters and process parameterizations [Montanari and Koutsoyiannis, 2012], and retain the flexibility to adjust the model parameters to suit different hydroclimatic regimes. Because most physical processes are parameterized to some extent, treating uncertain model parameters as fixed physical constants can unnecessarily constrain the agility of process-based models and severely limit their applicability to scales and locations for which these “parameters” have not been tuned. Furthermore, although the reasons for imposing hard wired parameters may be obvious to the original model developers (e.g., related to the lack of measurements at the spatial scales of

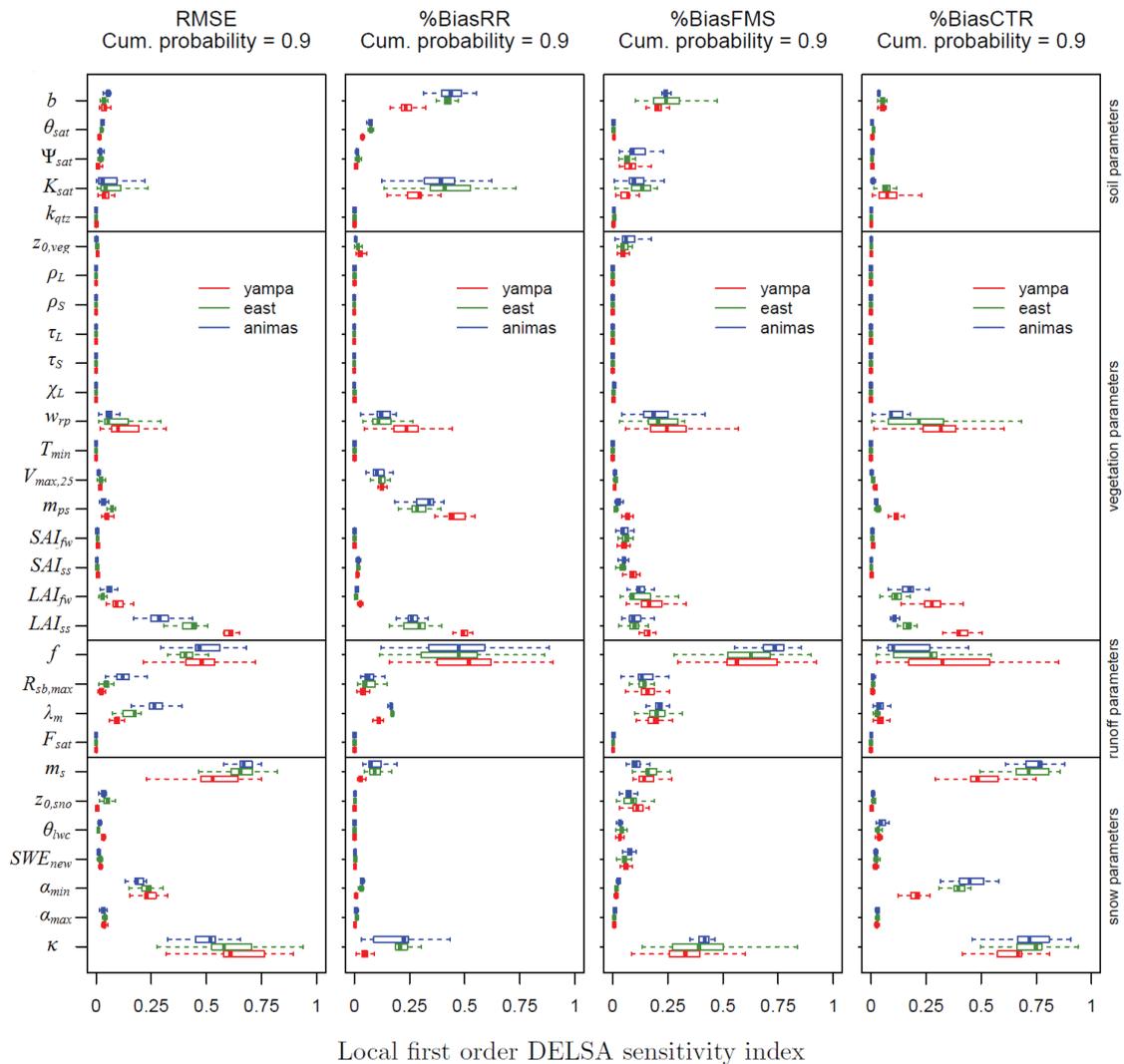


Figure 4. 90% quantiles of the full frequency distribution of local first order sensitivity indices for several objective functions: root mean squared error (RMSE), percent bias in runoff ratio (%BiasRR), percent bias in flashiness of runoff (%BiasFMS) and percent bias in runoff seasonality (%BiasCTR). The uncertainty estimates are obtained by bootstrapping (resampled 1000 times). The vertical bold line in the boxplot is the median, the body of a boxplot shows the interquartile range (Q75–Q25) and the whiskers represent the sample minima and sample maxima. In DELSA, the assessment of parameter sensitivity is based on local gradients of the model performance index with respect to model parameters at multiple points throughout the parameter space. DELSA indices scale between 0 and 1, and larger values are associated with very sensitive parameters.

application, or the need to impose boundaries for some coefficients), these reasons may be less obvious to future model developers and users. Exposing parameter values to users is, in our opinion, a transparent and informative practice that supports future model development and improvement. Further, it is naive to believe that these hard-coded numerical values (e.g., the value 0.55 in the equation of Figure 2) can be denoted by precise values instead of probability density functions, considering that they have been either specified based on order-of-magnitude considerations or estimated via statistical analysis.

Moreover, ignoring the spatial scales for which physically based equations describing fluxes of water and energy were derived (e.g., Richards' equation) and the spatial scale at which the empirical parameterizations were originally estimated (e.g., the Clapp-Hornberger pedo-transfer-functions or the saturated soil hydraulic conductivity in Table 1) will induce large uncertainties due to inappropriate scaling or averaging procedures, which in turn will propagate into model states and fluxes. In other words, hydrologic theory (e.g., Darcy's law) developed at the scale of laboratory experiments (0.01–0.1 m) may be appropriate for predictions at the point scale, but may need to be modified for applications at larger scales (e.g., hillslope, catchment and beyond) due to effects of nonlinearities, heterogeneities of landscape properties (e.g., vegetation, soils) and preferential flow of water through the soil matrix [Beven, 2002]. For instance, although the Clapp-

Hornberger b parameter (as defined in Table 1) appears to be valid for a grid whose area is either 1 m^2 or 100 km^2 (i.e., it is implied to be quasi-scale invariant), it depends on the soil texture [Clapp and Hornberger, 1978], implying that the equations at which this parameter appears should be estimated at a scale for which the soil texture can be assumed quasi-homogeneous but still with some degree of uncertainty. Since a cell of 10 km^2 will contain many kinds of soil types, a scaling procedure should be performed to estimate the “effective” soil saturated hydraulic conductivity that best represents the subgrid variability of soil within the given cell [Samaniego et al., 2010].

4. Where To From Here?

Our call for increased agility of process-based models contributes to the debate on the “correct” approach to modeling [e.g., Freeze and Harlan, 1969; Beven, 2002; Reggiani and Schellekens, 2003; Loague and VanderKwaak, 2004]. “Physics-based” models reflect a high level of confidence in the spatiotemporal representativity of physics-based equations describing complex systems, encoding very strong a priori assumptions regarding individual processes [e.g., Abbott et al., 1986; Wigmosta et al., 1994; VanderKwaak and Loague, 2001; Ivanov et al., 2004; Maxwell and Miller, 2005; Rigon et al., 2006; Qu and Duffy, 2007; Lawrence et al., 2011; Niu et al., 2011], which hinder the representation of hydrologic process idiosyncrasies in specific catchments. By contrast, “conceptual” models begin with limited a priori assumptions and infer knowledge through interpretation of how catchments respond to external forcing [e.g., Burnash et al., 1973; Lindström et al., 1997; Perrin et al., 2003; Fenicia et al., 2011], but are typically highly parameterized and do not explicitly represent many of the dominant physical processes necessary to reasonably simulate hydrological processes under changing hydroclimatic and land use conditions. The relative strengths of the so-called “physics-based” and “conceptual” modeling philosophies are therefore in their respective reliance on prior knowledge and data-based inference. A key challenge in moving forward is to integrate these strengths to improve model representation of hydrological processes.

Finding a good balance between strong prior knowledge and data-based inference requires stepping back from specific model equations and examining the major decisions in the development of process-based hydrological models: (1) what schemes should we use to represent spatial variability and hydrologic connectivity throughout the model domain; (2) what parameterizations should we use to simulate the fluxes of water and energy at the spatial scale of the model discretization; and (3) what values should we use for the model parameters. When viewed from this perspective, there is no real distinction between physics-based and conceptual models: there is a continuum of modeling approaches [Gupta et al., 2012], with inter-model differences simply defined by decisions on which processes are represented explicitly, the spatial resolution used to simulate them, and the methods used to estimate model parameter values. The fundamental question follows from the key challenge just expressed: *How can we integrate our understanding of environmental physics with the available data to both define the structure of a hydrological model and define suitable values for model parameters?*

In our opinion, improving hydrological models requires developing effective methods to define and discriminate among competing modeling options, including both model structure and model parameters. This involves both (1) increasing the physical realism of traditional rainfall-runoff models and reducing the reliance on traditional model calibration methods that are plagued by compensatory errors and unrealistic hydrologic process simulations; and (2) increasing the agility of physically motivated modeling systems to better suit local conditions. Modeling advances require explicitly simulating all dominant biophysical and hydrological processes, and focusing attention on detailed process-based evaluation of the suitability of different methods to represent spatial variability and hydrological connectivity, different scale-appropriate flux parameterizations, and different approaches to estimate model parameter values. Implementing this vision requires effective methods for a controlled and systematic approach to model development and improvement [Clark et al., 2011; Gupta et al., 2012], obtained by incorporating multiple modeling options into agile physics-based modeling frameworks and by applying a process-based philosophy for model evaluation and diagnosis.

Further, achieving this vision requires reconciling more agile models with the available data in order to identify suitable model structures and model parameter values. A useful solution to this problem can be found in the ‘diagnostic approach’ for model evaluation, based on confronting information contained in the data with the information provided by models [Gupta et al., 2008], and in the use of probabilistic representation of process parameterization equations [Bulygina and Gupta, 2011]. The diagnostic approach has

proved to be useful for finding optimal parameter sets that provide a more realistic representation of catchment processes [e.g., Pokhrel and Gupta, 2009; van Werkhoven et al., 2009; Kollat et al., 2012; Pokhrel et al., 2012]. The combined use of a diagnostic evaluation approach with inverse estimation and data assimilation methods (see Liu and Gupta [2007] and Gupta et al. [2012] for an overview of techniques) can reduce the dimensionality of the model evaluation problem (e.g., focus on a subset of processes), and facilitate the reconfiguration of agile models (e.g., refinement of model equations, state and parameter updating) using information extracted from new data sets.

5. Concluding Remarks

In this commentary we argue that the relatively poor performance of very complex physics-based hydrologic models can originate from unnecessary constraints that make it difficult to experiment with different kinds of spatial variability and process parameterizations. As in the example presented here, it is typical for parameters in complex models to be specified using values reported in the literature, often based on limited data or order-of-magnitude considerations. This practice constrains our abilities to conduct extensive analysis and limits our opportunities to improve model fidelity and characterizing model uncertainty.

In view of this, we encourage an expanded and more comprehensive evaluation of critical modeling assumptions, building on the advances in multiple hypothesis modeling methodologies [e.g., Pomeroy et al., 2007; Clark et al., 2008, 2011; Fenicia et al., 2011; Niu et al., 2011; Essery et al., 2013]. Future modeling systems should incorporate the capability to modify representations of spatial variability and hydrologic connectivity, individual process representations, numerical schemes, and couplings with other model components (e.g., atmosphere, sediment transport). Ongoing development of more agile versions of Noah-MP is just one example of active research in this area. Moreover, model reconfiguration capabilities should be able to cater to variable data availability (e.g., more complex model structures and meaningful specification of parameter values as more information is available) and integrate mechanisms for uncertainty quantification and analysis (e.g., ensemble generation, data assimilation, statistical postprocessing and visualization). Such capabilities are necessary to facilitate diagnosis of model adequacy problems, refine model representations of natural processes, understand the major sources of uncertainty in model simulations, and identify critical areas for future research.

Finally, future research should also investigate robust physically based scaling theories that can explain, and hence simulate, the heterogeneity of biophysical and hydrologic processes across multiple spatial scales. Progress in this direction will facilitate improved predictions of water and energy fluxes across different scales and locations, while constituting a necessary step toward addressing the grand challenge of hyper-resolution large-scale modeling proposed by Wood et al. [2011].

Acknowledgments

We acknowledge the United States Geological Survey for making available daily streamflow records. The WRF data sets used in this article can be obtained through the Hydrometeorological Applications Program (HAP) at the National Center of Atmospheric Research. We thank Mark Raleigh, Naoki Mizukami, Ethan Gutmann, Andy Newman, Andy Wood, Thorsten Wagener, Alberto Montanari, Paul Smith, Uwe Ehret and two anonymous reviewers for highly constructive comments, which contributed significantly to improve this manuscript. We also thank Oldrich Rakovec and Mary Hill for their advice in the implementation of DELSA. The first author received support from the CIRES Graduate Fellowship Award, a research grant from the Bureau of Reclamation, and a contract from the U.S. Army Corps of Engineers. The National Center for Atmospheric Research is sponsored by the National Science Foundation. We also acknowledge support from the Australian Research Council Centre of Excellence for Climate System Science (CE110001028) and from the EU-funded SWAN project (grant 294947) under the EU 7th Framework Programme.

References

- Abbott, M., J. Bathurst, J. Cunge, P. O'Connell, and J. Rasmussen (1986), An introduction to the European Hydrological System—Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system, *J. Hydrol.*, *87*(1), 61–77.
- Aguado, E. (1985), Radiation balances of melting snow covers at an open site in the Central Sierra Nevada, California, *Water Resour. Res.*, *21*(11), 1649–1654, doi:10.1029/WR021i011p01649.
- Amorocho, J., and B. Espildora (1966), Mathematical simulation of the snow melting processes, *Water Sci. Eng. Pap. 3001*, Dep. of Water Sci. Eng., Univ. of Calif., Davis.
- Anderson, E. (1973), National Weather Service River Forecast system—Snow accumulation and ablation model, *NOAA Tech. Memo. NWS HYDRO-17*, 217 pp., U.S. Dep. of Commer., Silver Spring, Md.
- Asner, G. P., C. A. Wessman, D. S. Schimel, and S. Archer (1998), Variability in leaf and litter optical properties: Implications for BRDF model inversions using AVHRR, MODIS, and MISR, *Remote Sens. Environ.*, *257*, 243–257.
- Bergstrom, S. (1991), Principles and confidence in hydrological modelling, *Nord. Hydrol.*, *22*, 123–136.
- Beven, K. (1997), TOPMODEL: A critique, *Hydrol. Processes*, *11*, 1069–1085.
- Beven, K. (2000), Uniqueness of place and process representations in hydrological modelling, *Hydrol. Earth Syst. Sci.*, *4*(2), 203–213.
- Beven, K. (2002), Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, *Hydrol. Processes*, *16*(2), 189–206, doi:10.1002/hyp.343.
- Beven, K. (2006), Searching for the Holy Grail of scientific hydrology: $Qt=(S, R, \Delta t)A$ as closure, *Hydrol. Earth Syst. Sci.*, *10*(5), 609–618, doi:10.5194/hess-10-609-2006.
- Beven, K., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298.
- Beven, K. J., and H. L. Cloke (2012), Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al., *Water Resour. Res.*, *48*, W01801, doi:10.1029/2011WR010982.
- Blöschl, G., and M. Sivapalan (1995), Scale issues in hydrological modelling: A review, *Hydrol. Processes*, *9*, 251–290.
- Bulygina, N., and H. Gupta (2011), Correcting the mathematical structure of a hydrological model via Bayesian data assimilation, *Water Resour. Res.*, *47*(5), W05514, doi:10.1029/2010WR009614.

- Burnash, R., R. Ferral, and R. McGuire (1973), *A Generalized Streamflow Simulation System—Conceptual Modeling for Digital Computers*, Joint Fed.-State River Forecast Center, Natl. Weather Serv. and State of Calif. Dep. of Water Resour., Sacramento, Calif.
- Carpenter, T. M., and K. P. Georgakakos (2004), Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model, *J. Hydrol.*, *298*(1-4), 202–221, doi:10.1016/j.jhydrol.2004.03.036.
- Clapp, R., and G. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, *14*(4), 601–604.
- Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.
- Clark, M. P., D. Kavetski, and F. Fenicia (2011), Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resour. Res.*, *47*, W09301, doi:10.1029/2010WR009827.
- Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn (1984), A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils, *Water Resour. Res.*, *20*(6), 682–690, doi:10.1029/WR020i006p0682.
- De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2006), Assessment of model uncertainty for soil moisture through ensemble verification, *J. Geophys. Res.*, *111*, D110101, doi:10.1029/2005JD006367.
- Dirmhirn, I., and F. Eaton (1975), Some characteristics of the Albedo of Snow, *J. Appl. Meteorol.*, *14*, 375–379.
- Dorman, J., and P. Sellers (1989), A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB), *J. Appl. Meteorol.*, *28*, 833–855.
- Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031.
- Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, *J. Optim. Theory Appl.*, *76*(3), 501–521, doi:10.1007/BF00939380.
- Essery, R., and P. Etchevers (2004), Parameter sensitivity in simulations of snowmelt, *J. Geophys. Res.*, *109*, D20111, doi:10.1029/2004JD005036.
- Essery, R., S. Morin, Y. Lejeune, and C. B. Ménard (2013), A comparison of 1701 snow models using observations from an alpine site, *Adv. Water Resour.*, *55*, 131–148, doi:10.1016/j.advwatres.2012.07.013.
- Fenicia, F., J. J. McDonnell, and H. H. G. Savenije (2008), Learning from model improvement: On the contribution of complementary data to process understanding, *Water Resour. Res.*, *44*, W06419, doi:10.1029/2007WR006386.
- Fenicia, F., D. Kavetski, and H. H. G. Savenije (2011), Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, *Water Resour. Res.*, *47*, W11510, doi:10.1029/2010WR010174.
- Flanner, M. G., C. S. Zender, J. T. Randerson, and P. J. Rasch (2007), Present-day climate forcing and response from black carbon in snow, *J. Geophys. Res.*, *112*, D11202, doi:10.1029/2006JD008003.
- Foglia, L., M. C. Hill, S. W. Mehl, and P. Burlando (2009), Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function, *Water Resour. Res.*, *45*, W06427, doi:10.1029/2008WR007255.
- Freeze, R. A., and R. Harlan (1969), Blueprint for a physically-based, digitally simulated hydrologic response model, *J. Hydrol.*, *9*, 237–258.
- Göhler, M., J. Mai, and M. Cuntz (2013), Use of eigendecomposition in a parameter sensitivity analysis of the Community Land Model, *J. Geophys. Res. Biogeosciences*, *118*, 904–921, doi:10.1002/jgrg.20072.
- Goudriaan, J. (1977), *Crop Micrometeorology: A Simulation Study*, Pudoc, Cent. for Agric. Publ. and Doc, Wageningen, Netherlands.
- Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*(4), 751–763, doi:10.1029/97WR03495.
- Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, *Hydrol. Processes*, *22*, 3802–3813, doi:10.1002/hyp.
- Gupta, H. V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye (2012), Towards a comprehensive assessment of model structural adequacy, *Water Resour. Res.*, *48*, W08301, doi:10.1029/2011WR011044.
- Hastie, T., R. Tibshirani, and J. Friedman (2002), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, N. Y.
- Hogue, T. S., L. Bastidas, H. Gupta, S. Sorooshian, K. Mitchell, and W. Emmerich (2005), Evaluation and transferability of the Noah land surface model in semiarid environments, *J. Hydrometeorol.*, *6*(1), 68–84.
- Ikeda, K., et al. (2010), Simulation of seasonal snowfall over Colorado, *Atmos. Res.*, *97*(4), 462–477, doi:10.1016/j.atmosres.2010.04.010.
- Ivanov, V. Y., E. R. Vivoni, R. L. Bras, and D. Entekhabi (2004), Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: A fully-distributed physically-based approach, *J. Hydrol.*, *298*(1-4), 80–111, doi:10.1016/j.jhydrol.2004.03.041.
- Jordan, R. (1991), *A One-Dimensional Temperature Model for a Snow Cover*, Cold Regions Research and Engineering Lab, U.S. Army Corps of Engineers, Hanover, NH.
- Kattge, J., W. Knorr, T. Raddatz, and C. Wirth (2009), Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, *Global Change Biol.*, *15*(4), 976–991, doi:10.1111/j.1365-2486.2008.01744.x.
- Kavetski, D., G. Kuczera, and S. W. Franks (2006a), Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts, *J. Hydrol.*, *320*(1-2), 173–186, doi:10.1016/j.jhydrol.2005.07.012.
- Kavetski, D., G. Kuczera, and S. W. Franks (2006b), Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis, *J. Hydrol.*, *320*(1-2), 187–201, doi:10.1016/j.jhydrol.2005.07.013.
- Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, *42*, W03504, doi:10.1029/2005WR004362.
- Kollat, J. B., P. M. Reed, and T. Wagener (2012), When are multiobjective calibration trade-offs in hydrologic models meaningful?, *Water Resour. Res.*, *48*, W03520, doi:10.1029/2011WR011534.
- Kollet, S. J., R. M. Maxwell, C. S. Woodward, S. Smith, J. Vanderborght, H. Vereecken, and C. Simmer (2010), Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources, *Water Resour. Res.*, *46*, W04201, doi:10.1029/2009WR008730.
- Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters, *J. Hydrol.*, *331*(1-2), 161–177, doi:10.1016/j.jhydrol.2006.05.010.
- Lawrence, D. M., et al. (2011), Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, *J. Adv. Model. Earth Syst.*, *3*, M03001, doi:10.1029/2011MS000045.
- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, *99*(D7), 14,415–14,428.

- Liang, X., E. F. Wood, and D. P. Lettenmaier (1996), Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, *Global Planet. Change*, *13*(1-4), 195–206, doi:10.1016/0921-8181(95)00046-1.
- Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström (1997), Development and test of the distributed HBV-96 hydrological model, *J. Hydrol.*, *201*(1-4), 272–288, doi:10.1016/S0022-1694(97)00041-3.
- Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, *43*, W07401, doi:10.1029/2006WR005756.
- Loague, K., and J. E. VanderKwaak (2004), Physics-based hydrologic response simulation: Platinum bridge, 1958 Edsel, or useful tool, *Hydrol. Processes*, *18*(15), 2949–2956, doi:10.1002/hyp.5737.
- Marks, D., and J. Dozier (1992), Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada 2. Snow cover energy balance, *Water Resour. Res.*, *28*(11), 3043–3054.
- Marks, D., J. Domingo, D. Susong, T. Link, and D. Garen (1999), A spatially distributed energy balance snowmelt model for application in mountain basins, *Hydrol. Processes*, *13*, 1935–1959.
- Maxwell, R., and N. Miller (2005), Development of a coupled land surface and groundwater model, *J. Hydrometeorol.*, *6*, 233–247.
- Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360, doi:10.1175/BAMS-87-3-343.
- Montanari, A., and D. Koutsoyiannis (2012), A blueprint for process-based modeling of uncertain hydrological systems, *Water Resour. Res.*, *48*, W09555, doi:10.1029/2011WR011412.
- Myneni, R. B., R. Ramakrishna, R. Nemani, and S. W. Running (1997), Estimation of global leaf area index and absorbed par using radiative transfer models, *IEEE Trans. Geosci. Remote Sens.*, *35*(6), 1380–1393, doi:10.1109/36.649788.
- Niu, G., and Z. Yang (2006), Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale, *J. Hydrometeorol.*, *9*, 937–953.
- Niu, G.-Y., and Z.-L. Yang (2007), An observation-based formulation of snow cover fraction and its evaluation over large North American river basins, *J. Geophys. Res.*, *112*, D21101, doi:10.1029/2007JD008674.
- Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, *J. Geophys. Res.*, *110*, D21106, doi:10.1029/2005JD006111.
- Niu, G.-Y., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, *116*, D12109, doi:10.1029/2010JD015139.
- Oleson, K. W., X. Zeng, and X. Zeng (2010), *Technical description of version 4.0 of the Community Land Model (CLM)*, NCAR Tech. Note NCAR/TN-478+STR, 257 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.
- Otto, J., T. Raddatz, and M. Claussen (2011), Strength of forest-albedo feedback in mid-Holocene climate simulations, *Clim. Past*, *7*(3), 1027–1039, doi:10.5194/cp-7-1027-2011.
- Pauwels, V. R. N., and G. J. M. De Lannoy (2009), Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space, *Water Resour. Res.*, *45*, W08428, doi:10.1029/2008WR007590.
- Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, *279*(1-4), 275–289, doi:10.1016/S0022-1694(03)00225-7.
- Pokhrel, P., and H. V. Gupta (2009), Regularized calibration of a distributed hydrological model using available information about watershed properties and signature measures, in *Proceedings of Symposium HS.2 at the Joint IAHS and IAH Convention on New Approaches to Hydrological Prediction in Data-sparse Regions*, pp. 20–25, Hyderabad, India.
- Pokhrel, P., K. K. Yilmaz, and H. V. Gupta (2012), Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures, *J. Hydrol.*, *418–419*, 49–60, doi:10.1016/j.jhydrol.2008.12.004.
- Pomeroy, J. W., D. M. Gray, T. Brown, N. R. Hedstrom, W. L. Quinton, R. J. Granger, and S. K. Carey (2007), The cold regions hydrological model: A platform for basing process representation and model structure on physical evidence, *Hydrol. Processes*, *21*, 2650–2667, doi:10.1002/hyp.
- Prein, A. F., G. J. Holland, R. M. Rasmussen, J. Done, K. Ikeda, M. P. Clark, and C. H. Liu (2013), Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado headwaters, *J. Clim.*, *26*(13), 4848–4857, doi:10.1175/JCLI-D-12-00727.1.
- Prihodko, L., A. S. Denning, N. P. Hanan, I. Baker, and K. Davis (2008), Sensitivity, uncertainty and time dependence of parameters in a complex land surface model, *Agric. For. Meteorol.*, *148*(2), 268–287, doi:10.1016/j.agrformet.2007.08.006.
- Qu, Y., and C. J. Duffy (2007), A semidiscrete finite volume formulation for multiprocess watershed simulation, *Water Resour. Res.*, *43*, W08419, doi:10.1029/2006WR005752.
- Rakovec, O., M. C. Hill, M. P. Clark, A. H. Weerts, A. J. Teuling, and R. Uijlenhoet (2014), Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models, *Water Resour. Res.*, *50*, 1–18, doi:10.1002/2013WR014063.
- Rasmussen, R., et al. (2014), Climate change impacts on the water balance of the Colorado Headwaters: High-resolution regional climate model simulations, *J. Hydrometeorol.*, *15*(3), 1091–1116, doi:10.1175/JHM-D-13-0118.1.
- Reba, M. L., D. Marks, T. E. Link, J. Pomeroy, and A. Winstral (2014), Sensitivity of model parameterizations for simulated latent heat flux at the snow surface for complex mountain sites, *Hydrol. Processes*, *28*(3), 868–881, doi:10.1002/hyp.9619.
- Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreta, D.-J. Seo, and A. DMIP Participants (2004), Overall distributed model intercomparison project results, *J. Hydrol.*, *298*(1-4), 27–60, doi:10.1016/j.jhydrol.2004.03.031.
- Reggiani, P., and J. Schellekens (2003), Modelling of hydrological responses: The representative elementary watershed approach as an alternative blueprint for watershed modelling, *Hydrol. Process.*, *17*(18), 3785–3789, doi:10.1002/hyp.5167.
- Reggiani, P., S. M. Hassanizadeh, M. Sivapalan, and W. G. Gray (1999), A unifying framework for watershed thermodynamics: Constitutive relationships, *Adv. Water Resour.*, *23*(1), 15–39, doi:10.1016/S0309-1708(99)00005-6.
- Rigon, R., G. Bertoldi, and T. M. Over (2006), GEOTop: A Distributed Hydrological Model with Coupled Water and Energy Budgets, *J. Hydrometeorol.*, *7*(3), 371–388, doi:10.1175/JHM497.1.
- Rosero, E., Z.-L. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G.-Y. Niu (2010), Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, *J. Geophys. Res.*, *115*(D3), 1–21, doi:10.1029/2009JD012035.
- Sacks, W. J., D. S. Schimel, and R. K. Monson (2007), Coupling between carbon cycling and climate in a high-elevation, subalpine forest: A model-data fusion analysis, *Oecologia*, *151*(1), 54–68, doi:10.1007/s00442-006-0565-2.
- Samaniego, L., R. Kumar, and S. Attinger (2010), Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, *Water Resour. Res.*, *46*(5), W05523, doi:10.1029/2008WR007327.
- Sellers, P., D. Randall, and G. Collatz (1996), A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation, *J. Clim.*, *9*, 676–705.

- Sivapalan, M., G. Blöschl, L. Zhang, and R. Vertessy (2003), Downward approach to hydrological prediction, *Hydrol. Process.*, *17*(11), 2101–2111, doi:10.1002/hyp.1425.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers (2008), *A Description of the Advanced Research WRF Version 3*, Boulder, Colo.
- Smith, M. B., V. Koren, Z. Zhang, Y. Zhang, S. M. Reed, Z. Cui, F. Moreda, B. A. Cosgrove, N. Mizukami, and E. A. Anderson (2012), Results of the DMIP 2 Oklahoma experiments, *J. Hydrol.*, *418–419*, 17–48, doi:10.1016/j.jhydrol.2011.08.056.
- Son, K., and M. Sivapalan (2007), Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, *Water Resour. Res.*, *43*(1), W01415, doi:10.1029/2006WR005032.
- Tang, Y., P. Reed, T. Wagener, and K. van Werkhoven (2007), Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, *Hydrol. Earth Syst. Sci.*, *11*(2), 793–817, doi:10.5194/hess-11-793-2007.
- Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, *Water Resour. Res.*, *45*(12), W00B14, doi:10.1029/2008WR006825.
- Uhlenbrook, S., J. Seibert, C. Leibundgut, and A. Rodhe (1999), Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, *Hydrol. Sci. J.*, *44*(5), 779–797, doi:10.1080/02626669909492273.
- VanderKwaak, J., and K. Loague (2001), Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model, *Water Resour. Res.*, *37*(4), 999–1013.
- van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2008), Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, *44*(1), 1–16, doi:10.1029/2007WR006271.
- van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2009), Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models, *Adv. Water Resour.*, *32*(8), 1154–1169, doi:10.1016/j.advwatres.2009.03.002.
- Verseghy, D. (1991), CLASS—A Canadian land surface scheme for GCMs. I. Soil model, *Int. J. Climatol.*, *11*(2), 111–133.
- Vrugt, J. A., H. V. Gupta, L. a. Bastidas, W. Bouten, and S. Sorooshian (2003a), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, *39*(8), 1214, doi:10.1029/2002WR001746.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003b), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.
- Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*(1), W01017, doi:10.1029/2004WR003059.
- Vrugt, J. A., M. P. Clark, C. G. H. Diks, Q. Duan, and B. A. Robinson (2006), Multi-objective calibration of forecast ensembles using Bayesian model averaging, *Geophys. Res. Lett.*, *33*(19), L19817, doi:10.1029/2006GL027126.
- Wagener, T., P. Reed, K. van Werkhoven, Y. Tang, and Z. Zhang (2009a), Advances in the identification and evaluation of complex environmental systems models, *J. Hydroinformatics*, *11*(3–4), 266–281, doi:10.2166/hydro.2009.040.
- Wagener, T., K. van Werkhoven, P. Reed, and Y. Tang (2009b), Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling, *Water Resour. Res.*, *45*(2), 1–5, doi:10.1029/2008WR007347.
- Wigmosta, M., L. Vail, and D. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, *Water Resour. Res.*, *30*(6), 1665–1679.
- Wolf, A., K. Akshalov, N. Saliendra, D. A. Johnson, and E. A. Laca (2006), Inverse estimation of V_c max, leaf area index, and the Ball-Berry parameter from carbon and energy fluxes, *J. Geophys. Res.*, *111*(D8), D08S08, doi:10.1029/2005JD005927.
- Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian (1992), A Land-Surface Hydrology Parameterization With Subgrid Variability for General Circulation Models, *J. Geophys. Res.*, *97*(D3), 2717–2728, doi:10.1029/91JD01786.
- Wood, E. F., et al. (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resour. Res.*, *47*(5), W05301, doi:10.1029/2010WR010090.
- Xia, Y., et al. (2012), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, *117*(D3), D03109, doi:10.1029/2011JD016048.
- Yang, Z.-L., R. E. Dickinson, A. Robock, and K. Y. Vinnikov (1997), Validation of the snow submodel of the biosphere-atmosphere transfer scheme with Russian Snow Cover and Meteorological Observational Data, *J. Clim.*, *10*(2), 353–373, doi:10.1175/1520-0442(1997)010<0353:VOTSSO>2.0.CO;2.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *J. Hydrol.*, *204*(1–4), 83–97, doi:10.1016/S0022-1694(97)00107-8.