Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters

Martyn P. Clark¹ and Jasper A. Vrugt²

Received 23 December 2005; revised 3 February 2006; accepted 9 February 2006; published 25 March 2006.

[1] In a previous paper, Vrugt et al. (2005) presented a combined parameter and state estimation method, entitled SODA, to improve the treatment of input, output, parameter and model structural error during model calibration. The argument for using SODA is that explicit treatment of all sources of uncertainty should result in parameter estimates that closer represent system properties, instead of parameter values that are compensating for input, output and model structural errors. In this study we provide further support for this claim by applying the SODA method to the calibration of a simple 2-parameter snow model, using data from the Lake Eldora SNOwpack TELemetry (SNOTEL) site in Colorado, USA. Citation: Clark, M. P., and J. A. Vrugt (2006), Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters, Geophys. Res. Lett., 33, L06406, doi:10.1029/2005GL025604.

1. Introduction and Scope

[2] The difficulty in identifying the parameter values for hydrologic models occurs, in large part, because of model deficiencies [Moradkhani et al., 2005]. Classical model calibration methods identify parameter sets that minimize summary measure(s) of model error. These summary error metrics combine all sources of model uncertainty into a single error term (i.e., they assume the model is perfect), which can mean that model parameters are assigned unrealistic values to compensate for biases in model inputs or weaknesses in model structure.

[3] In a previous paper, Vrugt et al. (2005) presented a combined parameter and state estimation method, entitled SODA (Simultaneous Optimization and Data Assimilation), for improved treatment of input, output, parameter and model structural errors during model calibration. The basic idea of SODA is that model errors accumulate and persist in model state variables, and updating (or correcting) model state variables during the optimization process will improve the identifiability of model parameters.

[4] The purpose of this letter is twofold. First, we wish to illustrate how the perfect model assumption used in traditional model calibration strategies can lead to unrealistic parameter values. Second, we wish to demonstrate that explicitly accounting for errors in model inputs and model structure as part of the model calibration exercise can produce parameter values that are more reasonable.

[5] Our analyses are based on simulations with a simple 2-parameter snow model, applied to the Lake Eldora SNOwpack TELemetry (SNOTEL) site in Colorado, USA. Although this model has considerably less complexity than other snow models, the advantages for the current study are that the parameters are physically meaningful and the low dimensionality of the parameter space (2-D) allows us to visualize model performance for all possible parameter combinations. The transparency offered by a simple model helps to illuminate the general validity of the SODA approach.

2. Model Description and Forcing Data

2.1. The Snow Model

Snow simulations in this study are produced using a simple 2-parameter snow model:

\[
\frac{dSWE}{dt} = A - M \tag{1a}
\]

where \( SWE \) is snow water equivalent (mm), \( A \) is snow accumulation (mm day \(^{-1}\)), and \( M \) is snow melt (mm day \(^{-1}\)). Snow accumulation and melt are both parameterized based on temperature. Snow accumulation is parameterized as:

\[
A = \begin{cases} 
0 & \text{if } T > T_{\text{thresh}} \\
P & \text{if } T \leq T_{\text{thresh}}
\end{cases} \tag{1b}
\]

where \( P \) is the precipitation rate (mm day \(^{-1}\)), \( T \) is the mean daily temperature (°C), and \( T_{\text{thresh}} \) is the temperature threshold (°C) below which snow begins to accumulate. Melt is parameterized as:

\[
M = \begin{cases} 
0 & \text{if } T \leq T_{\text{thresh}} \\
\min[\alpha(T - T_{\text{thresh}}), SWE] & \text{if } T > T_{\text{thresh}}
\end{cases} \tag{1c}
\]

where \( \alpha \) is a melt factor (mm °C\(^{-1}\) day \(^{-1}\)), and, again, \( T \) is the mean daily temperature (°C) and \( T_{\text{thresh}} \) is the temperature threshold above which snow melts. Note that the same temperature threshold is used for snow accumulation and melt. In this simple parameterization snowmelt is uniform over a time step and increases linearly as the temperature exceeds the melting point, \( T_{\text{thresh}} \).

2.2. The Input Data Ensemble

[6] The probabilistic interpolation method introduced by Clark and Slater [2006] was used to produce ensemble input forcing data for the coordinates of the Lake Eldora SNOTEL site. The method is based on locally weighted regression, in which spatial attributes from station locations...
The latitude, longitude, and elevation are used as explanatory variables to predict spatial variability in precipitation and temperature. A detailed description of the interpolation method, including the algorithmic details and verification is given by Clark and Slater [2006] and will not be repeated here.

3. Result of Model Calibration

In this section we contrast the results of two different model calibration cases. In section 3.1 we ignore output and model structural errors and assign all the uncertainty in the model process to uncertainty in the parameter estimates. This corresponds to a classical model calibration. In the second case study we use the SODA method to reduce the effects of input, output, and model structural errors on the calibrated model parameters.

3.1. No State Updating Case

We used two classic model calibration tactics to provide an initial assessment of parameter sensitivity and parameter identifiability. The first was a deterministic calibration, in which we identified the parameter set that minimized the root mean squared error between the simulated SWE and the observed SWE at Lake Eldora. Here, each of the 100 input ensemble members was used for model forcing. (bottom) Contours of the objective function surface when the RMSE was computed between the ensemble mean SWE and observations.

The so-called “optimal” parameter sets in Figure 1 are clearly quite absurd. With these parameter sets, snow can accumulate up to temperatures of 5°C, and snow will not start melting until temperatures are above 5°C. Moreover, the melt rate varies by a factor of 2 among parameter sets.

The “optimal” parameter values in Figure 1 are compensating for model errors. Figure 2 illustrates simulated and observed SWE for an example water year (1994–1995). SWE was simulated using a default parameter set of (0.0, 2.0) and “optimal” parameter sets of (4.0, 5.0) and (6.0, 15.0), where the parameters are respectively the temperature threshold and the melt factor (see equation (1)). The simulations with the default parameter set produce significantly less SWE (thin dark line) than what is observed at Lake Eldora (thick dark line). It is only by raising the temperature threshold to artificially high levels that the SWE simulations match observations (see the thin light lines).

Readers could argue that since we know the appropriate range of our model parameters, then we should constrain the parameters to be within that range. However, in many hydrologic models parameters often represent effective conceptual properties that lack physical interpretability. We deliberately select wide parameter ranges to illustrate how model errors can affect the predictive accuracy of different parameter combinations.

3.2. SODA—The State Updating Case

To successfully implement the SODA method we need estimates of the model error and the output measurement error. These errors are prescribed using the hyperparameters $\lambda$ and $\zeta$ (defined later in equations (3) and (5)). Moradkhani et al. [2005] suggests methods to identify these hyper parameters. The size of the errors directly determines the spread and predictive capabilities of the ensemble, and controls the size of the state update.
y
t
l
z
in equation (6) is
l
y
is a vector of
y
is an adjustable parameter (set at 0.2). The auto-
day
y
is the observation corresponding to the
is the observation error variance,
Clark et al.
is the vector of simulated model states (SWE) for
y
¼/C0
S
3
y
is sampled from a Normal
distribution. The stochastic forcing error is parameterized
auto-correlated random numbers from a standard Normal
distribution with mean equal to the SWE observation (y')
and variance equal to R (equation (5)). K in equation (6) is
the Kalman gain (see Vrugt et al. [2005] or Clark et al.
[2006] for computational details). The recursive state
updating steps are repeated until the end of the time series
is reached, and the likelihood of the assumed parameter set
computed as the root mean square error between the
ensemble mean and SWE observations before the state
update. The best parameter set is thus defined as the one
that minimizes the differences between one-observation-ahead
model forecasts and observations [Vrugt et al., 2005].

3.2.3. Implementation of SODA

[17] The SODA method implements an inner Ensemble
Kalman Filter (EnKF) loop for recursive state estimation
(conditioned on an assumed parameter set), within an outer
stochastic global optimization loop for batch estimation of
the posterior density of the parameters. The EnKF uses a
Monte Carlo based ensemble of state trajectories to propa-
gate and update estimates of the mean and covariance of the
uncertain state variables from one time step to the next. When
an output measurement is available, each forecasted ensem-
ble state is updated by means of a linear updating rule:

\[ \psi_{i,t}^{u} = \psi_{i,t} + K \left( y_{t}^{o} - \psi_{i,t}^{a} \right) \] (6)

where \( \psi_{i,t}^{a} \) is the simulated SWE for the \( i \)th ensemble member
before the update, \( y_{t}^{o} \) is the observation corresponding to the
\( i \)th ensemble member \( y_{t}^{o} \) is sampled from a Normal
distribution with mean equal to the SWE observation (y')
and variance equal to R (equation (5)). K in equation (6) is
the Kalman gain (see Vrugt et al. [2005] or Clark et al.
[2006] for computational details). The recursive state
updating steps are repeated until the end of the time series
is reached, and the likelihood of the assumed parameter set
computed as the root mean square error between the
ensemble mean and SWE observations before the state
update. The best parameter set is thus defined as the one
that minimizes the differences between one-observation-ahead
model forecasts and observations [Vrugt et al., 2005].

3.2.4. Results With SODA

[18] Figure 3 illustrates results of the SODA optimiza-
tion. We first note that the errors are much lower than those
in Figure 1. This is expected from the state-updating step, as
we use the same observations for the model updates that are
used for model calibration. The most interesting feature in
Figure 3 is that the parameter values with the minimum
objective function include values that we expect from
physical reasoning, that is, temperature thresholds between
0°C and +2°C, and melt factors less than 5 mm °C⁻¹ day⁻¹.
This result provides strong support that the data assimilation
step in SODA can reduce the effects of model structural
ersors and produce model parameter values that are more
identifiable and physically meaningful.

[19] The effectiveness of the state updating in SODA is
naturally model dependent. The size of the state updates
depends on the relative magnitude of model and observation
errors as well as the covariance between the model pre-
dictions of the observations and model states—if the obser-
vation errors are large or if the covariance is low, then the
state updates will be small. Moreover, the effectiveness of
state updates in removing model errors depends on the
extent to which the errors accumulate and persist in model
state variables. These limitations should be considered
carefully when using SODA in other models.

3.3. Improving the Snow Model: Interpretation of the
State Updates

[20] Diagnosing the time series of state updates to iden-
tify model structural flaws is occasionally used as part of the
model development process [e.g., Beck, 1987]. However,
most previous approaches do not examine the state updates in the context of model calibration, and there is thus some ambiguity as to whether state updates reflect a model structural weakness or an unsuitable choice of model parameters. An enticing facet of the SODA algorithm is that it separates the effects of parameter uncertainty and model weaknesses such that the time series of state updates directly provides information on model errors.

Almost all updates in the simple 2-parameter snow model are positive (not shown), indicating that the model is underestimating snow accumulation. A variety of reasons might explain this phenomenon. Here we assume that these model errors occur due to under-catch of precipitation in the COOP and SNOTEL gauge network, and accordingly modify the accumulation part of the model in equation (1b) with a multiplicative parameter to remove the bias in the precipitation measurements:

\[
A = \begin{cases} 
0 & \text{if } T > T_{\text{thresh}} \\
\kappa & \text{if } T \leq T_{\text{thresh}} 
\end{cases}
\]  

(7)

where \(\kappa\) is a dimensionless precipitation correction factor.

To evaluate the effects of including the additional parameter, Figure 4 shows the topology of the objective function for different parameter combinations in the 3-parameter snow model. PM is the precipitation multiplier (dimensionless), MF is the melt factor (mm °C\(^{-1}\) day\(^{-1}\)), and TT is the temperature threshold (°C). The contour shading is the same as in Figure 1b, where the dark, medium, and light shades depict objective function values less than 60 mm, 80 mm, and 100 mm respectively.

The discerning reader may point out that we didn’t really need SODA to identify the model structural errors, as they were blatantly apparent in Figure 2. We won’t quibble with this argument—the purpose of this paper is to use a very simple model to illustrate the general validity of the SODA approach. The real power of SODA will be in more complex models, in which model state variables are either not routinely measured, or lack physical interpretability (or both). In these more complex situations model state variables will be updated based on the modeled covariance across model states [see, e.g., Clark et al., 2006], and the update time series from SODA will provide information on model structural weaknesses that has hitherto been unavailable.

4. Summary and Discussion

This letter used a simple snow model to illustrate how the SODA method reduces the impact of model errors on model parameter estimates. Our classical model calibrations that lump input, output, and model structural errors into a single white noise term resulted in optimized parameter values that are physically unrealistic. Our model calibration with SODA resulted in model forecasts that not only closer represent the observed snow accumulations, but perhaps more importantly led to optimized values of the model parameters that more closely represent processes they are meant to represent.
In this case study we only “know” that the parameters from SODA are more realistic because the parameters are physically meaningful. We will rarely have this luxury. This prompts the following hypothesis: In more complex models when parameters often represent effective conceptual properties that lack physical interpretability, the step of reducing persistent model errors should produce calibrated parameter values that are better suited to the processes they are meant to represent. We clearly cannot test this hypothesis, but the results from this study provide strong support for this claim. The quest for more realistic model parameters is a noble one—more realistic parameters will both boost confidence in model predictions and increase the transferability of parameters to ungauged basins.

References


M. P. Clark, CIRES, University of Colorado at Boulder, Boulder, CO 80309, USA. (clark@vorticity.colorado.edu)
J. A. Vrugt, Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87544, USA.