Towards improved treatment of parameter uncertainty in hydrologic modeling
Vrugt, J.A.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
CHAPTER 7

Improved Treatment of Uncertainty in Hydrologic Modeling: Combining the Strengths of Global Optimization and Data Assimilation

Abstract

Hydrologic models use relatively simple mathematical equations to conceptualize and aggregate the complex, spatially distributed and highly interrelated water, energy, and vegetation processes in a watershed. A consequence of process aggregation is that the model parameters often do not represent directly measurable entities, and must therefore be estimated using measurements of the system inputs and outputs. During this process, known as model calibration, the parameters are adjusted so that the behavior of the model approximates, as closely and consistently as possible, the observed response of the hydrologic system over some historical period of time. In practice, however, because of errors in the model structure and the input (forcing) and output data, this has proven to be difficult, leading to considerable uncertainty in the model predictions. This Chapter surveys the limitations of current model calibration methodologies, which treat the uncertainty in the input–output relationship as being primarily attributable to uncertainty in the parameters, and presents a Simultaneous Optimization and Data Assimilation method, entitled SODA, which improves the treatment of uncertainty in hydrologic modeling. The usefulness and applicability of SODA is demonstrated by means of a pilot study using data from the Leaf River Watershed in Mississippi and a simple hydrologic model with typical conceptual components.

---

1This Chapter has been written by Jasper A. Vrugt, Cees G.H. Diks, Hoshin V. Gupta, Willem Bouten, and Jacobus M. Verstraten, and was submitted to Water Resources Research.
CHAPTER 7

7.1. Introduction and scope

Hydrologic models often contain parameters that cannot be measured directly, but can only be meaningfully inferred by calibration to a historical record of input-output data. During model calibration, the parameters are adjusted in such a way that the behavior of the model approximates, as closely and consistently as possible, the observed response of the hydrologic system over some historical period of time. Because of the time consuming nature of manual trial-and-error model calibration, there has been a great deal of research into the development of automated (computer based) calibration methods [see e.g., Gupta and Sorooshian, 1994; Yapo et al., 1998; Boyle et al., 2000]. Automatic methods for model calibration seek to take advantage of the speed and power of computers, while being relatively objective and easier to implement than manual methods.

While considerable progress has been made in the development and application of automated optimization methods, major weakness of these approaches include their underlying treatment of the uncertainty in the input-output representation of the model as being primarily (and explicitly) attributed to uncertainty in the parameter estimates, without explicit treatment of the input, output and model structural uncertainties. However, uncertainties in the modeling procedure stem not only from uncertainties in the parameter estimates, but also from measurement errors associated with the system input (forcing) and output, and from model structural errors arising from the aggregation of spatially distributed real-world processes into a mathematical model. Not properly accounting for these errors during model calibration, can result in model simulations and their associated prediction uncertainty bounds, which do not consistently represent and bracket the measured system behavior. This is usually evidenced by residuals, which exhibit considerable variations in bias (non-stationarity), variance (heteroscedasticity), and correlation structures under different hydrologic conditions. Several contributions to the hydrologic literature have therefore brought into question the continued usefulness of the classical paradigm for estimating model parameters [Beven and Binley, 1992; Gupta et al., 1998; Kavetski et al., 2003], especially in the face of the emerging generation of spatially distributed multi-input-output hydrologic models for which multiple (and often conflicting) sources of information are available for model calibration.

Some interesting methods for addressing these problems, particularly in the context of estimating reasonable confidence bands on the model simulations, have begun to appear in the literature. These methods include the use of classical Bayesian [Kuczera and Parent, 1998; Kavetski et al., 2003; Thiemann et al., 2001; Vrugt et al., in Chapter 3 of this thesis], pseudo-Bayesian [Beven and
SIMULTANEOUS PARAMETER AND STATE ESTIMATION

Binley, 1992; Freer et al., 1996], set-theoretic [Keesman, 1990; Klepper et al., 1991; van Straten and Keesman, 1991; Vrugt et al., 2001a], and multiple-criteria [Gupta et al., 1998; Yapo et al., 1998; Boyle et al., 2000; Vrugt et al., in Chapter 6 of this thesis] methods to represent model parameter and prediction uncertainty. Each of these methods, however, uses one or more aggregate statistics of model performance over a large range of hydrologic behaviors, an action that is now understood to result in considerable loss of important information that can be used to distinguish between competing parameter sets. To increase the discriminative power of model calibration strategies, interest has therefore recently begun to switch towards recursive model identification strategies [Thiemann et al., 2001; Young, 2001; Vrugt et al., in Chapter 5 of this thesis; Wagener et al., 2003], which sequentially move through a time series of discharge data and can provide estimates of parameter uncertainty. These approaches can not only be used to localize the most informative measurements for parameter estimation [see Vrugt et al., in Chapter 5 of this thesis], but perhaps more importantly, to provide a method for checking for violations of the underlying assumption that parameters are constant [Wagener, 2003; Misirli, 2003]. Despite progress made, such methods still assign the uncertainty in the input – output representation primarily to a combination of uncertainty in the parameter estimates plus a residual “model” error term [see Thiemann et al., 2001, Gupta et al., 2003], and therefore lack the conceptual rigor needed to distinguish between all the important sources of uncertainty.

In a separate line of research, considerable progress has been made in the development and application of sequential data assimilation (SDA) techniques. Such methods provide a general framework for explicitly dealing with input, output and model structural uncertainty, and for optimal merging of uncertain model predictions with observations. In contrast to classical model calibration strategies, SDA methods continuously update the states in the model when new measurements become available to improve the model forecast and evaluate the forecast accuracy. The prototype of the SDA methods, the Kalman Filter (KF), was developed in the 1960s for optimal control of systems governed by linear equations and was first introduced into hydrology during the early 80s to improve real-time forecasting of river discharges [Kitanidis and Bras, 1980a,b], and to recursively estimate model parameters [Beck, 1987]. For nonlinear dynamics, the extended Kalman Filter (EKF) can be used, which linearizes the error covariance equation using a tangent linear operator. However, because third- and higher order moments in the closure scheme are discarded, this linearization is notoriously unstable if the nonlinearities are strong [Evensen, 1992; Miller et al., 1994].

Although the KF and EKF offer a very general framework for accounting for all sources of uncertainty, the filter typically assumes that the optimal values of the model parameters are
known prior to the modeling procedure. This is particularly difficult in hydrologic modeling, where parameters often represent conceptual properties whose values can generally not be independently measured or assessed from lookup tables. This uncertainty in the correct choice of the parameter values, results in uncertainty in the estimated state values and therefore the model output prediction. To improve the usefulness and applicability of SDA methods to hydrologic modeling, it is important to explicate a formal methodology which recursively estimates the states in the model, while simultaneously also identifying the most likely values of the model parameters. The major difference between such a framework and the aforementioned classical model calibration strategies is that this approach no longer assigns the primary uncertainty in the input-output representation to uncertainty in the parameters, but explicitly accounts for input, output and model structural errors during model calibration.

This Chapter presents a Simultaneous parameter Optimization and Data Assimilation method entitled, SODA, which combines the strengths of the parameter search efficiency and explorative capabilities of the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm [Vrugt et al., in Chapter 3 of this thesis], with the power and computational efficiency of the Ensemble Kalman Filter [Evensen, 1994] to provide a better treatment of the input, output, parameter and model structural uncertainties in hydrologic modeling. The implementation of SODA requires a change in philosophy because it means that the model calibration procedure can no longer be considered as finding a single set of parameter values which generate the best possible long-term model forecasts. Instead, the optimal set of parameters identified with SODA only has meaning in combination with a SDA method, and consequently, will (per definition) not generate the best possible forecasts when no state adjustments are allowed, which is typically the case when evaluating the predictive capabilities of the model during an independent evaluation period. We hypothesize that the implementation of SODA will result in, (i) meaningful prediction uncertainty bounds on the model simulations, (ii) parameter estimates which better represent system properties that are less corrupted by modeling errors, thereby increasing the prospects of finding useful regionalization relationships, and (iii) a time series of recursive state and output adjustments, whose interpretation will generate inspiration to improve our model concepts and as such our understanding of the functioning of hydrologic systems.

This Chapter is organized as follows. Section 7.2 surveys the limitations of current model calibration strategies, which do not explicitly treat input, output and model structural uncertainty, but assign the uncertainty in the input-output representation primarily to uncertainty in the parameter estimates. In section 7.3 and 7.4 we discuss the rationale and architecture of SDA and global optimization methods, and demonstrate how the strengths of these methods can be
SIMULTANEOUS PARAMETER AND STATE ESTIMATION

combined to derive the hybrid SODA framework. In section 7.5 we illustrate the power and applicability of SODA by means of a standard mathematical study using the three-parameter highly nonlinear Lorenz model [Lorenz 1963], and a classical rainfall-runoff modeling example. In this section we also introduce a nonparametric variance estimator, which is especially designed to estimate the measurement error of output data. Finally, in section 7.6 we summarize the results.

7.2. Basic Inverse Formulation

The fundamental problem with which we are concerned is to simultaneously estimate parameter values and state variables in a hydrologic model using a historical record of input - output data. The formulation of this resulting inverse problem can be expressed in a generic form if we assemble the state variables in the nonlinear hydrologic model at time \( t \) into the state vector \( \psi_t \). The evolution of this state vector is described with:

\[
\psi_{t+1} = \eta(\psi_t, \bar{x}_t, \theta) \tag{7.1}
\]

where \( \psi \) is a vector of \( m \) unknown state variables, \( \eta(\cdot) \) represents the nonlinear hydrologic model used to simulate the state evolution, \( \bar{x} \) is an observed forcing field, \( \theta \) is a set of \( p \) model parameters, and \( t \) denotes time. We further assume that:

\[
\theta \in \Theta \subseteq \mathbb{R}^p \tag{7.2}
\]

where \( \mathbb{R}^p \) denotes the \( p \)-dimensional Euclidean space. If \( \Theta \) is not the entire domain space \( \mathbb{R}^p \), the inverse problem is said to be constrained. In hydrologic modeling, the feasible parameter space \( \Theta \) can usually be properly constrained using realistic upper and lower bounds on each of the model parameters. Let \( \bar{Y} = \{ \bar{y}_1, \ldots, \bar{y}_n \} \) denote the vector of streamflow measurement data available at time steps \( 1, \ldots, n \) and let \( Y(\theta) = \{ y_1(\theta), \ldots, y_n(\theta) \} \) represent the corresponding vector of model output predictions using the parameter values \( \theta \). These model output predictions are related to the model state according to:

\[
y_i = H(\psi_i) \tag{7.3}
\]
where $H(\cdot)$ is the measurement operator, which maps the state space into the measurement or model output space. The tilde over $X$ and $Y$ emphasizes that these quantities are measured and hence subject to sampling and measurement error. The differences between the model-simulated output and measured data can be represented by the residual vector:

$$ E(\theta) = G[Y(\theta)] - G[Y] = \{e_1(\theta), ..., e_n(\theta)\} $$

(7.4)

where the function $G(\cdot)$ allows for various user-selected linear or nonlinear transformations.

The classical approach to estimating the parameters in Eq. (7.1) is to ignore input uncertainty ($\tilde{X} = X$) and to assume that the predictive model $\eta$ is a correct, or at least accurate, representation of the underlying physical data-generating process. In line with classical statistical estimation theory, the residuals in Eq. (7.4) are then assumed to be mutually independent (uncorrelated), Gaussian distributed, with a constant variance. Under these circumstances, the traditional “best” parameter set in Eq. (7.1) can be found by minimizing the following lumped simple least square (SLS) objective function with respect to $\theta$:

$$ F_{SLS}(\theta) = \sum_{i=1}^{n} e_i^2 $$

(7.5)

The limitations of neglecting input errors and lumping observed response and model structural error into one single white noise term, becomes immediately apparent when inspecting the time series of residuals after model calibration. To illustrate this, consider Figure 7.1, which displays a 5-month portion of the measured and simulated hydrographs for the Leaf River basin in Mississippi. The HYMOD conceptual watershed model was calibrated using the Shuffled Complex Evolution (SCE-UA) global optimization algorithm developed by Duan et al. [1992] using the SLS criterion and 1 year of calibration data. Notice, that the fit to the observed data can be considered quite good, but that the residuals exhibit considerable variation in bias (non-stationarity), variance (heteroscedasticity) and correlation structure under different hydrologic conditions. Indeed, the common approach of dealing with model structural and data errors, as being “small” or somehow “absorbed” into the output error residual needs serious reconsideration. Moreover, given the presence of these errors, overconditioning of the model to a single parameter set is unreasonable, and cannot be justified.
One set of responses set forth to directly address the problem of overconditioning is to abandon the Frequentists approach of believing that the model parameters in Eq. (7.1) are fixed but unknown, and to adopt a Bayesian viewpoint which allows the identification of a plausible set of values for the parameters of the model given the available data. The Bayesian approach treats the model parameters in Eq. (7.1) as probabilistic variables having a joint posterior probability density function (pdf), which summarizes our probabilistic belief about the parameters $\theta$ in the light of the observed data $\tilde{y}$. Examples of Bayesian approaches to hydrology include the Generalized Likelihood Uncertainty Estimation (GLUE) framework of Beven and Binley [1992] and the Bayesian Recursive Estimation (BaRE) approach of Thiemann et al [2001] for representing model parameter and prediction uncertainty within the context of Monte Carlo analysis, and the Shuffled Complex Evolution Metropolis (SCEM-UA) global optimization algorithm of Vrugt et al [Chapter 3 of this thesis] for simultaneously estimating the traditional “best” parameter set along with a sample set of parameter values describing the probabilistic representation of the remaining parameter uncertainty (which is then used to generate probabilistic forecasts). We note that the SCEM-UA and BaRE methods operate within the
context of the classical Bayesian framework (which makes implicit assumptions about the correctness or adequacy of the model of the data generating process), while GLUE requires the user to make other kinds of subjective decisions including the selection of a cutoff threshold that separates behavioral from non-behavioral parameter sets.

An alternative set of responses set forth to identify model structural inadequacies, is to use Recursive Parameter Estimation algorithms, which provide better extraction of the information in the data, because the temporal aggregation associated with traditional batch processing is reduced. Examples of recursive algorithms which can be applied to hydrologic models, include the PIMLI and recursive SCEM-UA algorithms [Vrugt et al., in Chapter 3 of this thesis 2002; Vrugt et al., 2003b], the DYNIA approach [Wagener et al., 2003], the BaRE algorithm [Thiemann et al., 2001], and the application of Transfer Functions with time-varying parameters identified using Instrumental Variable techniques [Young, 2001]. While such methods introduce some much needed creative thinking into the field of hydrologic model identification, the use of such methods to help diagnose and quantify model structural errors has yet to be satisfactorily demonstrated.

A third set of responses set forth to more directly confront the problem of model structural errors is to pose the model calibration problem into a multi-objective framework, as was advocated by Gupta et al. [1998]. A major consequence of model structural imperfection is that the model is incapable of reproducing all aspects and portions of the output data with a single parameter set. By employing a number of complementary criteria in the optimization procedure, and analyzing the tradeoffs in the fitting of these criteria, the hydrologist is able to better understand the limitations of current hydrologic model structures, and gain insights into possible model improvements. The resulting Pareto solution set of parameters defines a kind of parameter uncertainty attributable to model structural errors. Yapo et al. [1998] and later Vrugt et al. [Chapter 5 of this thesis] presented the MOCOM-UA and MOSCEM-UA algorithms, respectively, capable of solving, in a single optimization run, the multi-objective problem posed by this approach. Unlike the previous methods, the multi-objective approach explicitly acknowledges the existence of model structural errors during the parameter estimation procedure. However, because of subjectivity in the selection of the measures [Gupta et al., 1998] and data subsets [see Boyle et al., 2000] to be included in the multi-objective vector, considerable fuzziness in the specification of the Pareto optimal parameter space and thus related model structural errors exists. Moreover, multi-objective calibration is typically response focused, and does not explicitly consider the influence of input errors [see Kavetski et al., 2003]. Like the
previous methods all sources of uncertainty are treated primarily as parameter uncertainty, giving rise to the concept of Pareto optimality.

Finally, the last set of responses set forth to more directly confront the problem of input errors is to define more realistic error models. One such approach is the recently proposed BAyesian Total Error Analysis (BATEA) of Kavetski et al. [2003], which imposes a hitherto missing rigor in environmental modeling by requiring the specification of realistic input uncertainty models. From the surveyed methods in this section, BATEA is the only method that explicitly takes into account input errors in the development of the likelihood function, and as such no longer assigns all the uncertainty in the input–output representation to uncertainty in the parameters. However, it is not yet clear how this likelihood function can be easily extended to include for model structural and output errors as well.

In summary, current model calibration strategies are poorly suited to the task of explicitly treating the various important sources of uncertainty associated with the application of hydrologic models. In the next two sections of this Chapter, we will discuss the challenging question of how to develop a model calibration strategy that explicitly accounts for input, output and model structural errors.

7.3. Sequential Data Assimilation

Sequential data assimilation methods, at least in principle, specify the uncertainty in the system states that arises from imperfect process representation and from input and output data uncertainty. To facilitate the description of the KF, we start by writing the model dynamics in Eq. (7.1) as a stochastic differential equation:

\[
d\psi_t = \eta(\psi_{t-1}, \hat{X}_{t-1}, \theta)dt + dq_{t-1}
\]  

(7.6)

where \( dq \in \mathbb{R}^n \) is a dynamical noise term representing errors in the model formulation. This stochastic forcing term tends to flatten the probability density function of the model states during the integration. In section 7.3.2 of this Chapter we further elaborate on the simulation of \( q \). The observation equation, Eq. (7.3), also has a random additive error \( \varepsilon \), called the measurement error:

\[
\gamma_t = H(\psi_t) + \varepsilon_t ; \quad \varepsilon_t \sim N(0, \sigma^2)
\]  

(7.7)
where \( \sigma'_t \) signifies the error deviation of the measurements. At each measurement time \( t = 1, \ldots, n \) when an observation becomes available compute the output forecast error, \( z_t \):

\[
z_t = \tilde{y}_t - H(\psi'_t)
\]

(7.8)

and update the forecasted states, \( \psi'_t \) using the standard analysis equation:

\[
\psi'_t = \psi'_t + K_t \left[ \tilde{y}_t - H(\psi'_t) \right]
\]

(7.9)

where \( \psi'_t \) is the updated or analyzed state, and \( K_t \) denotes a matrix of weights (Kalman gain), which is computed as:

\[
K_t = \Sigma_t \Sigma_t H^T \left[ H \Sigma_t H^T + \Sigma_e \right]^{-1}
\]

(7.10)

where \( \Sigma_t \) and \( \Sigma_e \) denote the covariance matrices of the stochastic model error term (\( = \sigma_q^2 \)) and observations (\( = \sigma_e^2 \)), respectively. The difference between the forecasted and updated state is commonly referred to in the literature as the state innovation, \( \psi'_t' \):

\[
\psi'_t = \psi'_t' - \psi'_t
\]

(7.11)

which mapped in the output space is called the output innovation, \( I_t \):

\[
I_t = H(\psi'_t) - H(\psi'_t)
\]

(7.12)

The analyzed state \( \psi'_t \) then recursively feeds the next state propagation step in the model:

\[
\psi'_{t+1} = \eta(\psi'_t, \tilde{X}_t, \theta)
\]

(7.13)

Since, uncertainty in the model structure and output data can be specified through the stochastic forcing term (\( q \)) and output measurement error (\( e \)), respectively, and uncertainty in the input data
can be taken into account through stochastic perturbations of the elements of the input ($X$), the KF offers a very general framework for dealing with all sources of uncertainty.

Even though the KF or EKF implement a more sensible model calibration strategy than traditional model calibration methods, as these filters explicitly specify the uncertainty in the system states that arises from imperfect process representation and from input and output data uncertainty, their widespread application has been limited by the strong non-linearity of hydrologic models, and the computational requirements associated with the storage and forward integration of the error covariance matrix. To resolve these two major problems, Evensen [1994] proposed the Ensemble Kalman Filter (EnKF), which uses a Monte Carlo (MC) method to generate an ensemble of model trajectories from which the time evolution of the probability density of the model states, and related error covariances are estimated. The EnKF avoids many of the problems associated with the traditional EKF method, e.g., there is no closure problem as is introduced in the EKF by neglecting contributions from higher order statistical moments in the error covariance evolution. Moreover, the conceptual simplicity, relative easy of implementation and computational efficiency of the EnKF make the method an attractive option for data assimilation in the meteorologic, oceanographic and hydrologic sciences [Evensen and van Leeuwen, 1996; Houtekamer and Mitchell, 1998; Lermusiaux, 1999; Madsen and Caniêres, 1999; Keppenne, 2000; Reichle et al., 2002; among others].

### 7.3.1. The Ensemble Kalman Filter

The EnKF uses an ensemble of model trajectories to solve for Eqs. (7.6) to (7.12). The description of the basic algorithm of the EnKF below follows Evensen [1994] and is further illustrated in Figure 7.2.

1. **Generate initial ensemble.** Sample $N$ combinations of $m$ model states, $\psi^j, j = 1, ..., N$ randomly from the prior distribution and store them in a matrix $A[1:m, 1:N]$:

   $A = (\psi^1, ..., \psi^N) \in \mathbb{R}^{mN}$  \hspace{1cm} (7.14)

2. **Forecast step.** Propagate each of the $N$ ensemble members of $A$ forward in time with the nonlinear model $\eta(\cdot)$, using a fixed set of model parameters $\theta$ and a corresponding ensemble of $N$ stochastic forcing fields:
Explicit details on how to simulate $q_{t|t}$ will be given in the next section.

\[
\psi'_t = \eta(\psi'_{t-1}, \tilde{X}_{t-1}, \theta) + q'_{t-1}
\]  

(7.15)

Figure 7.2. Flowchart of the Ensemble Kalman Filter (EnKF) used to recursively estimate state variables.
[3] **Compute the covariance matrix of the forecasted states.** At each time step $t$ compute the covariance matrix, $\Sigma^* \in \mathbb{R}^{m \times m}$ of the ensemble forecasted states in matrix $A$:

$$
\Sigma^* = \frac{1}{N-1} A'(A')^T
$$

(7.16)

where $A' = A - \overline{A}$ and $\overline{A}$ denotes the ensemble mean of the forecasted states.

[4] **Compute the mean ensemble forecast error.** At time $t$ when a measurement become available, compute the mean forecast error of the ensemble:

$$
\bar{z} = H(A) - \bar{y}
$$

(7.17)

[5] **Generate vector of observations.** At each measurement time, generate an ensemble of $N$ observations from a distribution with mean equal to the observation, $\bar{y}$, and covariance equal to $\Sigma^*$:

$$
\begin{align*}
\bar{y}' & = \bar{y} + \bar{e}' \\
\bar{e}' & \sim N(0, \sigma^2) \\
\Sigma^* & = \bar{e}' \bar{e}'^T 
\end{align*}
$$

(7.18)

and store them in a matrix $D[1:1:N]$:

$$
D = (\bar{y}', ..., \bar{y}^N) \in \mathbb{R}^{1 \times N}
$$

(7.19)

This particular step ensures that the spread of the updated ensemble is consistent with the true estimation error covariance [Burgers et al., 1998].

[6] **State update or analysis step.** At each measurement time, update each ensemble member of $A$ using the following analysis equation:

$$
A = A + \Sigma^* H^T (H \Sigma^* H^T + \Sigma^*)^{-1} [D - H(A)]
$$

(7.20)
CHAPTER 7

where the product $\sum^w H^T (H \sum^w H^T + \sum^u)^{-1}$ is a numerical approximation of the Kalman gain, $K$, in Eq. (7.9).

[7] Check stop criterion. If $t$ is less than the total number of time steps, $n$, return to step 2, otherwise stop.

In summary, the EnKF uses a large ensemble of model trajectories to approximate the probability density of the model states at each time step. The mean of this ensemble represents the “best” state estimate, whereas the variance provides a measure of the spread of the ensemble members (uncertainty). For more information about the derivation of the EnKF and its practical implementation, please refer to Evensen [1994].

7.3.2. Simulation of the Stochastic Model Error Term

To successfully implement the EnKF it is particularly important to specify accurate values of the output measurement error and stochastic forcing term used to simulate the time evolution of dynamical model noise. Hence, it is not difficult to see that the size of these terms directly controls the spread among the ensemble members, and as such the mean ensemble model forecasts and associated prediction uncertainty bounds. Fortunately, in most situations a reasonable a-priori estimate can be made of the measurement error, $\sigma^\alpha$, of the observations. A more delicate issue however, is how to retrieve a reasonable estimate of the stochastic forcing term. This term reflects many sources of uncertainty including model discretization errors and neglected dynamical processes.

Based on recommendations in the classic treatise of the EnKF in Evensen [1994] we used the following equation to simulate the time evolution of the stochastic forcing:

$$s_t = \phi s_{t-1} + \sqrt{1-\phi^2} w_{t-1}$$

(7.21)

where $w_t$ is an $m \times 1$ vector of draws from a normal distribution, and $\phi \in [0,1)$ represents the first-order decorrelation of the stochastic forcing. The form of Eq. (7.21) ensures that, with an increasing number of elements, $s$ will approach a distribution having mean equal to zero and variance equal to one. A value for $\phi = 0$ generates a sequence of white noise terms, while $\phi = 1$
will remove the stochastic forcing. Based on random walk theory, the nonlinear model $\eta(\cdot)$ can then be written as:

$$
\psi_t = \eta(\psi_{t-1}, X_{t-1}, \theta) + \tau s_{t-1} \sqrt{\Delta t \sigma}
$$

(7.22)

in which $\Delta t$ denotes the time step, $\sigma$ is an $m \times 1$ vector of error deviations of individual states and $\tau$ is a factor which ensures that the forcing error over time becomes independent of $\phi$ and $\Delta t$ and is computed as [Evensen, 1994]:

$$
\tau = \sqrt{\frac{1}{\Delta t} \frac{(1-\phi)^2}{\Delta t / -2\phi - \phi^2 + 2\phi^{(l)}}}
$$

(7.23)

where $l$ denotes the number of steps in each time unit $\Delta t$ ($l=1$).

While this equation is frequently used in meteorology and oceanography to simulate the time evolution of the stochastic forcing, our attempt to apply this equation to conceptual hydrologic modeling has revealed one major drawback. When the states in the model do not have a clear physical meaning, as is typically the case with conceptual hydrologic models, it is difficult to specify accurate values of the error deviations of individual states. To circumvent this problem, we decided to specify $\sigma$ in the output space of the model and to use this output perturbation to compute the corresponding corrupted states in the model:

$$
\psi_t = \eta(\psi_{t-1}, X_{t-1}, \theta) + \tau s_{t-1} \sum_{\phi} \frac{H \sum_{\phi} H^T}{\sqrt{\Delta t \sigma^2}}
$$

(7.24)

where $\sum_{\phi}$ represents the covariance matrix of the forecasted states and $\sigma^2$ denotes the error deviation of the stochastic forcing. Among various alternatives, this implementation was shown to work well for simulating the time evolution of the stochastic forcing for a variety of test problems.

### 7.4. Combining Global Optimization and Sequential Data Assimilation

Even though the EnKF provides a general framework to account for input, output, and model structural uncertainty, a defining characteristic of the filter is its focus on state estimation alone,
without recourse to parameter estimation. When applying the EnKF to recursively estimate state values, it is common practice to fix the model parameters to some predefined values, thereby explicitly ignoring the effects of parameter uncertainty and interaction. This section discusses how the strengths of global optimization and the EnKF can be merged to derive a hybrid framework, which recursively tracks the model states, while simultaneously also estimating the values of the model parameters (and their associated uncertainty).

To extend the applicability of the EnKF to simultaneous state and parameter estimation we must do two things. First, we must specify a mathematical criterion that measures the “closeness” between the EnKF derived mean ensemble model forecast and the corresponding measurement, which can be used to extract the information content in the data and transform it into estimates for the model parameters. This issue will be addressed in this paragraph. Second, we must implement a method that can effectively and efficiently minimize this measure, thereby preferably also generating an estimate of parameter uncertainty. We will tackle this problem in the next paragraph. To estimate the model parameters, we adopt a classical Bayesian approach using the following posterior density criterion, \( p(\theta | \bar{Y}) \), [Box and Tiao, 1973]:

\[
p(\theta | \bar{Y}) \propto \left( \sum_{j=1}^{m} \xi_j(\theta) \right)^{-\frac{1}{2}} \]

(7.25)

in which \( \xi(\theta) \) denotes the time series of mean ensemble forecast errors corresponding to the parameter set \( \theta \). Using this criterion, the best set of model parameters is defined as that set which is on average associated with the best one-observation-ahead model forecasts.

To generate samples from Eq. (7.25), and to summarize the posterior parameter pdf using statistical moments and histograms, we use an implementation of the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm. The SCEM-UA algorithm is a general-purpose global optimization algorithm that provides an efficient estimate of the most likely parameter set and its underlying posterior probability distribution within a single optimization run [see Vrugt et al., in Chapter 3 of this thesis]. The algorithm is a Markov Chain Monte Carlo (MCMC) sampler, which generates a sequence of parameter sets \( \{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(k+1)}\} \) that convergences to the stationary posterior distribution for a large enough number of simulations \( k \). The SCEM-UA algorithm is related to the successful SCE-UA global optimization method, but uses the Metropolis-Hastings (MH) search strategy [Metropolis et al, 1953; Hastings, 1970] instead of the Downhill Simplex method for population evolution, and is therefore able to simultaneously infer...
both the most likely parameter set and its underlying posterior probability distribution within a single optimization run. A detailed description and explanation of the method appears in Vrugt et al. [Chapter 3 of this thesis], and so will not be repeated here.

In brief, the SCEM-UA method involves the initial selection of a population of points distributed randomly throughout the $p$-dimensional feasible parameter space. In the absence of prior information about the location of the maximum likelihood value a uniform sampling distribution is used. For each point the complete system of equations of the EnKF, as outlined in steps [1] to [7] in section 7.3.1, is executed and the time series of mean ensemble forecast errors returned to the SCEM-UA algorithm. This time series is subsequently inserted in Eq. (7.25) to calculate the posterior density for each point. The population of parameter sets is subsequently partitioned into a number of complexes, and in each complex a parallel sequence is launched from the point that exhibits the highest posterior density. A new candidate point in each sequence is generated using a multivariate normal distribution either centered on the current draw of the sequence or the mean of the points in the complex augmented with the covariance structure induced between the points in the complex. The Metropolis-annealing [Metropolis et al., 1953] criterion is used to test whether the candidate point should be added to the current sequence. Finally, the new candidate point is shuffled into the original population of complexes. The evolution and shuffling procedures are repeated until the Gelman - Rubin convergence diagnostic for each of the parameters demonstrates convergence to a stationary posterior target distribution [Gelman and Rubin, 1992]. Experiments conducted using standard mathematical test problems have shown that the SCEM-UA derived posterior distribution closely approximates the target distribution [Vrugt et al., in Chapter 3 of this thesis]. Throughout the remainder of this Chapter, we refer to this Simultaneous Optimization and Data Assimilation procedure, as SODA.

The overriding characteristic of the SODA methodology is the improved treatment of input, output, parameter and model structural uncertainties, during model calibration. Theoretically, if the model structure would be correct, and the input and output data are observed without error, no state adjustments are needed during model calibration, and the set of parameters that maximizes the posterior density in Eq. (7.25), will also be associated with, on average, the smallest output residuals, when performing traditional batch calibration without state adjustments (as done in section 7.2). However, because of errors in the model structure and input (forcing) and output data this will generally not be the case. The major objective of this Chapter is to demonstrate how the hybrid SODA strategy can be used to simultaneously estimate model parameters and state variables, resulting in reliable model prediction uncertainty
CHAPTER 7

bounds and a time series of state and output innovations whose interpretation will generate inspiration to improve our model concepts and as such our understanding of the functioning of hydrologic systems.

7.5. Case Studies

We illustrate the power and applicability of SODA by means of two different case studies. The first is a classical mathematical study using the three-parameter highly nonlinear Lorenz model [Lorenz, 1963], to demonstrates the ability of SODA to simultaneously estimate state variables and parameter values when confronted with highly nonlinear model dynamics. The second case study explores the usefulness of SODA by application to hydrologic modeling using the simple HYMOD conceptual watershed model and historical streamflow data from the Leaf River Watershed in Mississippi. In this study we are especially concerned with the estimation of model structural errors and present a novel nonparametric method, which is especially designed to estimate the measurement error of streamflow data.

7.5.1. Case Study I: A Highly Nonlinear Case: The Lorenz Equations

To demonstrate that SODA is indeed able to simultaneously estimate state variables and model parameters when confronted with highly nonlinear model dynamics, the first case study considers the three-parameter Lorenz model [Lorenz, 1963]. This model consists of a system of three nonlinear and coupled ordinary differential equations:

\[
\frac{dx}{dt} = \gamma(y - x) \\
\frac{dy}{dt} = \rho x - y - xz \\
\frac{dz}{dt} = x y - \beta z
\]

(7.26)

where \(x(t), y(t),\) and \(z(t)\) denote the dependent variables and \(\gamma, \rho,\) and \(\beta\) represent model parameters. Because of its highly nonlinear nature, this model has served as a testbed in the field of data assimilation for examining the properties for various data assimilation methods. A reference solution of \((x(t), y(t), z(t))\) with output print step \(\delta t = 0.25\) was computed for \(t \in [0,40]\) by solving Eq. (7.26) starting from the initial condition \((x_0, y_0, z_0) = (1.508870, -1.531271, 25.46091)\)
and fixed values for the parameters of \(\gamma=10\), \(\rho=28\), and \(\beta=8/3\). Synthetic observations were subsequently computed by imposing normally distributed noise, with mean zero and variance equal to 2.0, to the \(x\), \(y\) and \(z\)-values in this reference solution.

The error covariance of the stochastic forcing is defined to be diagonal with variances equal to 2.00, 12.13, and 12.31 for the three equations in (7.26) respectively, and is included in the model through the stochastic forcing term in Eq. (7.22). The error variances of the initial conditions and observations were taken to be the same values as those used for generating the synthetic observations. The values for the parameters, initial conditions, and error variances used in the setup of this case study are identical to those used in similar studies reported in the literature [see e.g. Evensen, 1994; Miller et al., 1994; Evensen, 1997; Evensen and van Leeuwen, 2000; among others].

The SODA algorithm was used to simultaneously estimate the state variables and model parameters using the synthetically generated observations and the density criterion in Eq. (7.25). A search population of 50 points in combination with 5 parallel sequences was selected based on recommendations in our previous work. The feasible parameter space was taken to be a uniform hypercube between \([0 - 30]\), \([0 - 50]\) and \([0 - 10]\) for the parameters \(\gamma\), \(\rho\), and \(\beta\) respectively. Furthermore, an ensemble of 100 members was used to compute the error statistics with the EnKF, on the basis of experimental evidence that larger ensemble sizes gave only marginal improvements in the approximation of the probability density of the model states. The estimated ensemble prediction uncertainty ranges associated with the computed \(x\)-values in Eq. (7.26) and corresponding to the most likely SODA identified parameter set is shown in Figure 7.3a.
Figure 7.3. Results obtained with SODA for the three-parameter highly nonlinear Lorenz model; (a) Prediction uncertainty ranges corresponding to the best parameter set identified with SODA and 100 ensemble members. The solid dots denote the observations, (b-d) Sampled γ, β, and ρ parameters in five different sequences (coded with different symbols) during the evolution of SODA to the posterior target distribution. The asterisks at the right hand side of these figures denote the “true” parameter values used to generate the synthetic observations.

The solid circles correspond to the observed x-values, whereas the dashed black line represents the evolution of the mean ensemble prediction. Note, that the prediction uncertainty ranges generally bracket the observations very well, indicating that the EnKF does a good job in tracking the state transitions. Further, the width of the ensemble prediction uncertainty bounds is consistent with the specified error variance of the model and observations.

The transitions of the parameters γ, ρ, and β in each of the five sequences (Markov Chains) during the evolution of SODA to the stationary posterior distribution is illustrated in the Figures 7.3b-d respectively. For clarity, the three different parallel sequences are coded with different symbols. The asterisks at the right hand side indicate the “true” values of the parameters. The 1-D scatterplots of the sampled parameter space demonstrate that at early stages during the evolution, the individual sequences tend to occupy different regions of the parameter space. At a later stage during the evolution, however, all of the individual sequences have been able to fully explore the feasible parameter space, and indeed they assign the highest probability
to the “true” parameter values used for generating the synthetic observations. All in all, we can conclude that SODA is able to simultaneously estimate state variables and model parameters when confronted with highly nonlinear model dynamics, so that we can proceed with the next case study.

7.5.2. Case Study II: The HYMOD Conceptual Watershed Model

We illustrate the power and applicability of the SODA algorithm to hydrologic modeling by application to the HYMOD conceptual watershed model using historical data from the Leaf River watershed (1950 km$^2$) located North of Collins, Mississippi. The data, obtained from the Hydrologic Research Laboratory (HL), consists of mean areal precipitation (mm/day), potential evapotranspiration (mm/day), and streamflow (m$^3$/s). Forty consecutive years of data (1948-1988) are available for this watershed, representing a wide variety of hydrologic conditions. The HYMOD model, consisting of a relatively simple rainfall excess model, described in detail by Moore [1985], connected with two-series of linear reservoirs (three-identical quick and a single reservoir for the slow response). A schematic overview of HYMOD appears in Figure 7.4.

Figure 7.4. Schematic representation of the HYMOD conceptual watershed model.

The model has five parameters: the maximum storage capacity in the catchment, $C_{\text{max}}$ (L), the degree of spatial variability of the soil moisture capacity within the catchment, $b_{\text{op}}$ (-), the factor distributing the flow between the two series of reservoirs, $\text{Alpha}$ (-), and the residence time of the linear quick and slow reservoirs, $R_q$ (T) and $R_s$ (T), respectively. The upper and lower bounds that define the prior uncertainty ranges of these parameters appear in Table 7.1.
Table 7.1. Prior uncertainty ranges and description of the HYMOD model parameters.

<table>
<thead>
<tr>
<th>Par.</th>
<th>Unit</th>
<th>Description</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{\text{max}}$</td>
<td>[mm]</td>
<td>Maximum storage watershed</td>
<td>200.00</td>
<td>500.00</td>
</tr>
<tr>
<td>$b_{\text{sp}}$</td>
<td>[-]</td>
<td>Spatial variability soil moisture capacity</td>
<td>0.10</td>
<td>2.00</td>
</tr>
<tr>
<td>Alpha</td>
<td>[-]</td>
<td>Distribution factor between two reservoirs</td>
<td>0.50</td>
<td>0.99</td>
</tr>
<tr>
<td>$R_s$</td>
<td>[d]</td>
<td>Residence time slow flow reservoir</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td>$R_q$</td>
<td>[d]</td>
<td>Residence time quick flow reservoir</td>
<td>0.30</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Because the HYMOD model and the Leaf River data have been discussed extensively in previous work [e.g., Yapo et al., 1996; Vrugt et al., in Chapters 3 and 5 of this thesis; Boyle et al., 2002; Misirli, 2003], we will not describe the details of these here. To implement the SODA framework outlined above, it is further necessary to specify an appropriate value for the parameter $\phi$ in Eq. (7.21). As our knowledge about a correct value for this first-order decorrelation coefficient is quite poor, we added this parameter to the optimization problem, using a feasible range between [0-1]. Based on our experience with the model we used 3 years of streamflow data (28 July 1952 – 28 July 1955) to illustrate that the SODA method is relatively simple to implement and helps to distinguish between the various sources of uncertainty in hydrologic modeling.

7.5.2.1. Derivation of the Measurement Error of Streamflow Data

To solve the model calibration problem using SODA, we must do two things. First, a realistic estimate of the error deviation of the output (streamflow) data must be retrieved. This issue will be addressed in this section. Second, we must find a way to obtain a reasonable estimate of input and model structural uncertainty. We will tackle this problem in the next section.

Unfortunately, the problem of estimating the measurement error of streamflow data has not proved to be simple. To further elaborate on this problem, consider observations obtained through:

$$y_i = h(t) + e_i, \quad e_i \sim N(0, \sigma_i^2) \quad (t = 1, \ldots, n)$$  \hspace{1cm} (7.27)

where $h(t)$ is the actual streamflow at time $t$ and the errors $e_i$ are independent random variables with zero mean and unknown variance, $\sigma_i^2$. Without any doubt, most attention in hydrologic modeling has been given to a correct estimation of the function $h$. However, estimation of the
error deviation is nearly as important as the estimation of \( b \) itself, to obtain reasonable confidence intervals on the model predictions.

During the last two decades several nonparametric methods have been proposed which estimate the error deviation in Eq. (7.27) [see e.g. Rice, 1984; Hall et al., 1990; Seifert et al., 1993; Dette et al., 1998]. All these methods involve differencing the original time series, \( \tilde{y}_j \), and estimating the error deviation as:

\[
\hat{\sigma}_i = \sqrt{\frac{1}{2(n-1)} \sum_{j=2}^{n} (\tilde{y}_j - \tilde{y}_{j-1})^2}
\]  

(7.28)

The underlying assumptions of this approach include that (1) \( b(t) \) as a function of \( t \) is sufficiently smooth, (2) the sampling interval is high compared to the typical timescale of \( b(t) \), and (3) the error terms, \( e \), exhibit a constant variance (homoscedastic). While the first two assumptions can be considered quite reasonable for a time series of daily streamflow data, the assumption of homoscedasticity of the streamflow error terms is not very realistic. For instance, Sorooshian and Dracup [1980] commented that large flows tend to have larger error variances compared to smaller flows, partially because of the non-linear nature of the rating curve used to transform stage measurements to flow volume estimates. To enable the identification of these non-homogeneous (heteroscedastic) streamflow errors, we decided not to compute an average standard error estimate of the total time series, as suggested in Eq. (7.28), but instead, to apply the following nonparametric error deviation estimator locally in the time series:

\[
\hat{\sigma}_i = \sqrt{\frac{\sum_{j=2}^{n} (\Delta^u \tilde{y}_j)^2}{(2n-1)}}
\]  

(7.29)

where \( \Delta^u \) denotes the difference operator applied \( u \) times. It can be readily verified that the estimator in Eq. (7.29) is insensitive to polynomial trends in \( b(t) \) up to order \( u \). In the literature more sophisticated higher-order differencing procedures have been proposed [Hall et al., 1990]. However, investigations with numerically generated streamflow data showed that Eq. (7.29) with the choice \( u = 3 \) works well in practice.

To demonstrate the validity of the proposed approach for estimating the error deviation of streamflow data, we synthetically generated two different test problems. Synthetic daily streamflow data, \( \{y, \ldots, y_n\} \) were first generated for the period June 28, 1952 to September 30,
1961, by driving the HYMOD model with observed mean areal rainfall of the Leaf River watershed and values for the parameters identical to those used in Fig. 7.1. This output time series was subsequently corrupted by adding one of the two following error models to the data to yield two different time series of ‘observed’ streamflow data; (1) a homoscedastic error model: $e_t \sim N(0,10)$; and (2) a heteroscedastic error model: $e_t \sim N(y_t,0.1y_t)$. Both these time series were subsequently used in conjunction with Eq. (7.29) to estimate the error deviation of the streamflow measurements. The results of this analysis are summarized in Figure 7.5, which presents scatterplots of ‘observed’ streamflow data versus estimated error deviations for the (a) homoscedastic, and (b) heteroscedastic error case.

![Figure 7.5](image)

**Figure 7.5.** Synthetic case studies to demonstrate the validity of the nonparametric estimator to derive the error deviation of streamflow measurements; (a) The homoscedastic error case, and (b) Heteroscedastic error case. The solid and dashed black lines in each of the figures denote the estimated and actual error model, respectively.

To further facilitate graphic interpretation of the results, the dark black line in each of the Figures represents the optimal fit of a spline function through the data, whereas the dashed line corresponds to the original error model used to corrupt the streamflow data. Notice, that the agreement between the predicted and actual error model can be considered quite excellent. This is especially true for the homoscedastic error case. Although not further reported here, we performed a variety of synthetic experiments with different error distributions, and found that the nonparametric estimator in Eq. (7.29) provides a close approximation to the actual error model, when the time series of streamflow data is sufficiently long (> 2 years).

To unravel the relationship between flow level and measurement error for the Leaf River watershed, Figure 7.6, presents a similar scatterplot as Fig. 7.5 for this real world data set. Again, the black line denotes the optimal fit of a spline function through the scattered data. The pattern
of the scatter plot and slope of the regression function clearly indicate the presence of heteroscedastic errors in the streamflow data.

![Figure 7.6. Scatterplot of observed streamflow data versus estimated error deviation of the measurements using 10 years of data from the Leaf River watershed. The solid line represents the fit of a spline function through the data.](image)

In general, the size of the streamflow error almost log-linearly increases from 0.01 to about 80 m$^3$/s in the flow range of 2 to 1000 m$^3$/s respectively. These results provide strong support for the claim by Sorooshian and Dracup [1980], that streamflow data exhibit non-homogeneous (heteroscedastic) errors. To optimally exploit the information contained in the scatter plot of Fig. 7.6, it would seem most productive to first cluster the data into various flow levels, and then to approximate the corresponding error probabilities with Gaussian mixture models. However, because the purpose of this case study is illustrative, we decided not to follow such an approach, but rather to implement the fitted nonlinear regression function in SODA, to relate the error deviation of the streamflow data to the flow level.

### 7.5.2.2. Derivation of Input Error Model and Error Deviation of the Stochastic Forcing

In the absence of a compelling basis for the assignment of a reasonable input error model to stochastically perturb the elements of $X$, we decided to merge input and model structural errors into a single forcing term. The approach that we propose here to estimating the total error deviation of this term, $\sigma^2$ in Eq. (7.24), is to use a classical post-calibration residual examination in combination with a simple error disaggregation method.
If we proceed with classical model calibration the output is a set of residuals which constitute a combination of model structural, input and output errors. The difference between this residual and the associated measurement error of the observation represents the combined effect of model structural and input error. To illustrate this, consider Figure 7.7, which presents a scatter plot of the error deviation of the residuals of the calibrated HYMOD model (see Fig. 7.1) against flow level. This plot was created using 10 years of streamflow data for the Leaf River watershed.

![Figure 7.7. Scatter plot of the computed error deviation of the residuals of the calibrated HYMOD model against flow level for a 10-year period of daily streamflow data for the Leaf River Watershed. The solid and dashed black lines represent the best possible fit of a spline function through the scattered data, and the previously estimated error model of the measurements, respectively.](image)

The dashed black line denotes the estimated measurement error model (see Fig. 7.6), whereas the solid line represents the optimal fit of a spline function through the scattered data. According to our statements, the definition of model structural and input error is the area between the two different solid and dashed black lines. Another spline function was fitted through this difference function and subsequently implemented in SODA to relate the error deviation of the stochastic forcing to flow level.

### 7.5.3. Use of SODA for Hydrologic Model Calibration: Model Precision and Accuracy

The SODA algorithm was used with the particular settings discussed in the previous two sections to simultaneously estimate the HYMOD state variables and model parameters using the
daily streamflow data of the Leaf River basin and the density criterion specified in Eq. (7.25). A search population of 250 points in combination with 5 parallel sequences and 100 ensemble members was selected. The procedure used about 1,500 parameter evaluations to converge to a reasonable estimate of the optimal model parameters, including their underlying posterior distribution. The results of this analysis are summarized in Table 7.2 and Figures 7.8, 7.9 and 7.10 and discussed below.

The usefulness of the implementation of SODA can be demonstrated in a number of different ways. One of the most straightforward is to compare the hydrograph prediction uncertainty ranges of the HYMOD model using SODA, with those obtained using a classical Bayesian SCEM-UA calibration, which ignores input uncertainty and lumps output and model error into a single white noise term. The size of the prediction uncertainty ranges is a measure of the precision of the model. Figure 7.8 presents the results of this analysis. The hydrograph prediction uncertainty ranges (light-gray region) of the HYMOD model corresponding to the optimal SODA identified parameter set and an ensemble size of 100 members is presented in Fig. 7.8a. The observed streamflows are indicated with solid circles, whereas the mean ensemble prediction is indicated with the dashed black line. Fig. 7.8b presents the prediction uncertainty ranges for the HYMOD simulated streamflows corresponding to the SCEM-UA derived posterior parameter estimates. From visual inspection of the two plots we see (as expected) that the input–output behavior of the HYMOD model when properly accounting for input, output and model structural errors, is more consistent with the observations. Hence, in the case of SODA, the prediction uncertainty ranges are reasonably small and bracket the observations. On the contrary, when applying a traditional Bayesian approach (Fig. 7.8a), without state adjustments, the model seems to be unable to match large portions of the hydrograph; this is indicated by sections where the light-gray region does not bracket the observed streamflow data. Although one might argue that this is a problem of overconditioning, and that better results can be obtained by assigning more relaxed cutoff thresholds to determine what can be considered an acceptable parameter set or not (as is done in GLUE), that in fact leads to considerably larger hydrograph uncertainty bounds than those presented for SODA, particularly during the storm events. Moreover, as argued in the first two sections of this Chapter, it is not realistic to attribute all uncertainty in the modeling exercise to uncertainty in the parameter estimates.
Figure 7.8. Results of SODA by application to hydrologic modeling using the HYMOD conceptual watershed model and historical streamflow data from the Leaf River Watershed in Mississippi; (a) Hydrograph prediction uncertainty ranges corresponding to the best SODA identified parameter set, and 100 ensemble members; The dashed black line denotes the mean ensemble prediction (b) 95% prediction uncertainty ranges of the HYMOD model forecasts associated with the SCEM-UA derived posterior parameter distribution. The dashed black line denotes the evolution of the most likely parameter set.
Another diagnostic measure, which can be used to contrast the results of SODA with those obtained using classical Bayesian model calibration, is to compare the autocorrelation functions of the SCEM-UA (see Fig. 7.1c) and SODA derived time series of forecast errors. Autocorrelation is a measure of the accuracy of the model predictions. Figure 7.9 presents the results of this analysis. To benchmark against a time series of white forecast errors, the dashed lines in these plots represent the theoretical 99% confidence intervals of a white error time series [Box and Jenkins, 1976].

![Autocorrelation functions](image)

**Figure 7.9.** Autocorrelation functions of the time series of forecast errors using (a) SCEM-UA algorithm, and (b) SODA framework. The dashed lines denote the theoretical upper and lower 99% significance intervals of a time series of white residuals.

When performing a traditional SCEM-UA calibration, there is significant autocorrelation between the residuals at the first few lags, confirming our earlier findings reported in Fig. 7.1. Notice, however, that there is considerably less autocorrelation between the forecast residuals when using the SODA framework, suggesting that recursive state adjustments remove a large part of the bias in the model predictions.

This is further demonstrated in Table 7.2, which presents summary statistics of the one-day-ahead streamflow prediction performance of the SODA and SCEM-UA methods for the 3-year calibration (WY 1952 – 1955) and 5-year evaluation (WY 1955 – 1960) period.
Table 7.2. Summary statistics (RMSE, CORR and BIAS) of the one-day-ahead streamflow forecasts using the SCEM-UA and SODA methods.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>CORR</td>
</tr>
<tr>
<td>SCEM-UA</td>
<td>17.25</td>
<td>0.903</td>
</tr>
<tr>
<td>SODA</td>
<td>13.14</td>
<td>0.96</td>
</tr>
</tbody>
</table>

The results in this Table indicate that SODA consistently provides better values of the Root Mean Square Error (RMSE), correlation (CORR), and bias (BIAS) statistics than conventional batch calibration for both time periods. Although not further demonstrated here, the performance of SODA with a simple 5-parameter conceptual model compares favorably well with other modeling approaches, including the 13-parameter Sacramento Soil Moisture Accounting model of Baraibah et al. [1973], and the multi-parameter Self-Organizing Linear Output Map (SOLO), recently developed by Hsu et al. [2003]. Thus, explicitly accounting for uncertainty in the input, output and model structure, during model calibration will lead to both an improved assessment of predictive uncertainty, and a simultaneous improvement in the model forecasts.

7.5.4. Use of SODA for Hydrologic Model Calibration: Parameter Uncertainty

Figure 7.10 presents posterior marginal probability density distributions for each of the HYMOD model parameters estimated for the Leaf River watershed using the SCEM-UA (Fig 7.10. a-e) and SODA (Fig 7.10. f-j) algorithms. The results presented in this Figure highlight two interesting observations. First, notice that for some of the model parameters the SCEM-UA and SODA algorithms have assigned the highest probability to different locations in the parameter space. If the model structure would be correct and the input and output data were observed without error, no state adjustments would be needed during model calibration and the SCEM-UA and SODA method would identify the same location for the mode of the posterior probability distribution in the full parameter space. Not properly accounting for model structural and input errors during model calibration leads to corrupted parameter estimates, which are compensating for these errors.
Figure 7.10. Marginal posterior probability distributions of the HYMOD model parameters derived for the Leaf River Watershed using the (A-E) SCEM-UA algorithm (F-J) and SODA framework.
Based on these propositions, we therefore suggest that the model parameters identified with SODA are less corrupted by model structural and input errors and, therefore, better represent the underlying properties of the physical system. Our view has also been supported by additional numerical experiments with synthetically generated streamflow data, and a corrupted input (rainfall) data during the inverse identification [see e.g., Kavetski et al., 2003 for a similar approach], which demonstrated that model parameters identified with SODA seem to be less sensitive to errors in the input and converge in close proximity of the true parameter values used to generate the synthetic streamflow data. To increase the prospects of finding useful regionalization relationships, it would therefore seem most productive to use the SODA identified mode of the parameter space.

A second significant and interesting observation is that the HYMOD model parameters become less well defined when allowing for recursive state adjustments during model calibration (compare histograms of identical counterparts in Fig. 7.10). For most of the parameters, the dispersion around the mode of the distribution has significantly increased. In the systems theoretic sense, this can be explained as part of the ‘information’ from the streamflow measurements being used in SODA to recursively estimate the state variables.

7.5.5. Use of SODA for Hydrologic Model Calibration: Towards Model Structural Improvements

An interesting byproduct of SODA, which deserves further investigation, is the computed time series of state and output innovations. Any inability of the model to represent the input – output behavior of the underlying hydrologic system, will result in state updates in the model, when new measurements of the system are assimilated and processed. We are therefore left with the intuitively reasonable hypothesis that the computed time series of state updates should contain valuable information about model structural errors. In the last section of this Chapter, we verify the validity of this hypothesis by closer examination of the time series of output innovations.

To facilitate this process, we begin by partitioning the hydrograph into a driven, nondriven quick, and a nondriven slow component, in a method similar to Boyle et al. [2000]. For each of these portions of the hydrograph, we plot the SODA-computed mean ensemble output innovation against the mean ensemble streamflow prediction (prior to the update). The results of this analysis are presented in Figure 7.11. For the driven and nondriven quick part of the hydrograph, these plots do not reveal any simple relationship between the ensemble mean predicted flow level (x-axis) and associated output innovations (y-axis). Nonlinear time series
SIMULTANEOUS PARAMETER AND STATE ESTIMATION

Analysis (not presented here) using artificial neural networks (ANN) indicates that some structure is, indeed, present in the time series of output innovations, but the process of how to translate these findings into simple mathematical equations which can be used to improve the conceptual model and as such our understanding of hydrologic processes, remains the topic of future work.

A more easily interpretable result is obtained for the nondriven slow part of the hydrograph (Fig. 7.11c). Here, a clear systematic linear relationship between flow level and associated output innovation is apparent. This relationship can easily be exploited to improve the performance of the HYMOD model during low flows. Indeed, when incorporating a linear relationship between flow level and output innovation into HYMOD, the predictive capabilities of the model during nondriven slow flow increases from an average error of 0.20 to 0.09 m$^3$/s. In contrast, the classical model calibration methodologies, which assign all the uncertainty in the input–output representation to uncertainty in the model parameters, provide little or no such guidance.

7.6. Summary and conclusions

The objective of model identification is to obtain a model where the input-state-output behavior is consistent with the measurements, and where the model predictions are accurate and precise. In practice, however, because of errors in the model structure, input (forcing), and output data, this has proven to be quite difficult, leading to considerable uncertainty in the model predictions. Classical model calibration strategies typically ignore input uncertainty, lump model structural and output errors into a single white noise term, and assign the uncertainty in the input-output representation of the model primarily to the parameters, therefore lacking the conceptual rigor required to properly account for all the important sources of uncertainty.
CHAPTER 7

In this Chapter we present a combined global optimization and data assimilation method, which improves the treatment of input, output, parameter, and model structural uncertainty in hydrologic model calibration. The method, entitled Simultaneous Optimization and Data Assimilation (SODA) merges the strengths of the parameter search efficiency and explorative capabilities of the Shuffled Complex Evolution Metropolis algorithm and the power and computational efficiency of the Ensemble Kalman Filter to simultaneously estimate parameters and state variables in hydrologic models.

The usefulness and applicability of SODA has been demonstrated for two preliminary case studies. The first case study considered the highly nonlinear three-parameter Lorenz model, and demonstrated that SODA is indeed successfully able to simultaneously estimate state variables and model parameters when confronted with highly nonlinear model dynamics. The second case study explored the usefulness of SODA by application to hydrologic modeling using the HYMOD conceptual watershed model and historical streamflow data from the Leaf River Watershed in Mississippi. The ability of SODA to deal with input, output, parameter and model structural errors, results in improved estimates of the parameter and model prediction uncertainty ranges. With this basis, a detailed investigation of the state and output innovation time series, can be used to investigate improvements to our model concepts and as such our understanding of the functioning of hydrologic systems.