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

Citation for published version (APA):
CHAPTER 1

Introduction

1.1. General background

Hydrology involves the study of processes describing the movement of water on the earth’s surface, from the land surface to the subsurface and from the subsurface to the surface to the atmosphere, as schematically outlined in Figure 1.1.

![Figure 1.1. Schematic overview of the various components of the hydrologic cycle [graph taken from the website http://watercycle.gsfc.nasa.gov].](image)

These processes occur on several time scales: from hours to days to weeks, for example, flash floods, water infiltration in soils, short-term droughts, and spring-time runoff events; to longer decadal to inter-decadal processes involving long-term droughts and changes in regional precipitation and water supply due to climate variability. Faced with the complexity and spatial
and temporal variability of hydrologic processes, and the difficulty of performing controlled experiments, a variety of hydrologic models have been developed to simulate the various components of the water cycle for a region of interest. These models are classified as conceptual, when water balance dynamics are represented by heuristic or empiric functions deemed to be qualitatively reasonable, or as physically-based when the functions are based on scientifically accepted principles [Kuczera, 1997]. Modeling is a framework for testing new theories and hypothesis in order to improve our understanding of the hydrological processes and how the different processes interact. Even the most elaborate physically-based model, however, cannot reflect the true complexity and heterogeneity of the processes occurring in the field, because to some degree, it must always conceptualize and aggregate complex interactions driven by a number of spatially distributed and highly interrelated water, energy, and vegetation processes by the use of only relatively simple mathematical equations. All hydrologic models are, therefore, to some degree lumped, and use effective parameters to characterize these spatial-temporal processes.

In order for a hydrologic model to reliably simulate a portion of the water cycle it is necessary to select appropriate values for the model parameters. Unfortunately, the estimation of the correct values of these effective parameters has not proved to be simple. As argued in the first paragraph, many (if not all) of the parameters in hydrologic models represent aggregated and lumped processes in space and time, which are not easily observable or measurable at the scale of interest, and therefore cannot readily be inferred from direct measurable quantities. Instead, the typical way to estimate the model parameters is to adjust them in such a way that the input-output behavior of the model approximates, as closely and consistently as possible, the underlying hydrologic system over some historical period of time. This process is known as model calibration. The parameters, which are estimated in this manner, represent effective conceptual representations of spatially and temporally heterogeneous watershed properties, insofar as they are related to inherent and invariant properties of the hydrologic system. After successful calibration the hydrologic model is ready to be used to predict the various components of the water cycle during future events under a set of naturally occurring circumstances. This modeling process is closely related to the development of understanding of hydrologic processes in the sense that a comparison of model predictions with observed system behavior (observations) is commonly used to test hypothesis about the operation of the system being modeled.
1.2. Classical paradigm of model calibration

A schematic overview of the model calibration process appears in Figure 1.2.

Consider a system $\mathcal{S}$ for which a hydrologic model $\eta$ is to be calibrated. Assume that the mathematical structure of the model is essentially predetermined and fixed and that realistic upper and lower bounds on each of the $p$ model parameters can be specified prior to any modeling being undertaken. Let $\tilde{Y} = \{\tilde{y}_1, \ldots, \tilde{y}_n\}$ denote the vector of 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 model $\eta$ with the parameter values $\theta$. The difference between the model-simulated output and measured data can be represented by the residual vector, $E$:

$$E(\theta) = G[\tilde{Y}(\theta)] - G[\hat{Y}] = \{e_1(\theta), \ldots, e_n(\theta)\}$$  \hspace{1cm} (1.1)

where the function $G(\cdot)$ allows for various user-selected linear or nonlinear transformations. The aim of model calibration now becomes finding those set of model parameters $\theta$ such that the measure $E$, commonly called the objective function, is in some sense forced to be as close to zero as possible. The development of a measure that mathematically measures the "size" of $E(\theta)$
is typically based on assumptions regarding the distributions of the measurement errors presented in the data.

The following development closely follows the work presented in Vrugt et al. [2003a] and Vrugt and Dane [2004]. The classical approach to estimating the parameters in Eq. (1.1) is to ignore input data uncertainty and to assume that the predictive model $\eta$ is a correct representation of the underlying physical data-generating system $3$. In line with classical statistical estimation theory, the residuals in Eq. (1.1) are then assumed to be mutually independent (uncorrelated), and Gaussian distributed with a constant variance. Under these circumstances, the traditional "best" parameter set in Eq. (1.1) can be found by minimizing the following additive simple least square (SLS) objective function with respect to $\theta$:

$$F_{SLS}(\theta) = \sum_{i=1}^{n} \varepsilon_i(\theta)^2 \quad (1.2)$$

This step has, in practice, proven to be quite difficult to carry out in a reliable and consistent manner, because there are typically a large number of parameters that need to be estimated, which usually either have similar or compensating effects on different parts of the simulated output. In combination with varying parameter sensitivity, this not only leads to considerable interaction among the model parameters, but perhaps more importantly, results in numerous local minima in the response surface (objective function mapped out in the parameter space).

**1.3. Manual and automatic parameter estimation**

In the process of parameter estimation or model calibration, the hydrologist adjusts the values of the model parameters such that the model is able to closely match the behavior of the real system it is intended to represent. In its most elementary form, this calibration is performed by manually adjusting the parameters while visually inspecting the agreement between observations and model predictions [Janssen and Heuberger, 1995]. In this approach, the "closeness" of the model with the measurements is typically evaluated in terms of several subjective visual measures, and a semi-intuitive trial-and-error process is used to perform the parameter adjustments [Boyle et al., 2000]. Because of the subjectivity and time-consuming nature of manual trial-and-error calibration, there has been a great deal of research into the development of automatic methods for calibration of hydrologic models [e.g., Gupta and Sorooshian, 1994].
Automatic methods for model calibration seek to take advantage of the speed and power of computers, while being objective and easier to implement than manual methods.

Many algorithms have been developed in the past to solve the nonlinear SLS optimization problem stated in Eq. (1.2). These algorithms may be classified as local search methodologies, when seeking for systematic improvement of the objective function using an iterative search starting from a single arbitrary initial point in the parameter space, or as global search methods in which multiple concurrent searches from different starting points are conducted within the parameter space. One of the simplest local search optimization methods, which is commonly used in the field of soil hydrology, is a Gauss-Newton (Levenberg-Marquardt) type of derivative based search [Marquardt, 1963]:

$$ \theta^{(k+1)} = \theta^{(k)} - H(\theta^{(k)})^{-1} (\nabla \eta(\theta^{(k)}))^T $$

(1.3)

where $\theta^{(k+1)}$ is the updated parameter set, and $\nabla \eta(\theta^{(k)})$ and $H(\theta^{(k)})$ denote the gradient and Hessian matrix, respectively, evaluated at $\theta = \theta^{(k)}$:

$$ \nabla \eta(\theta) = \left( \frac{\partial \eta}{\partial \theta_1}(\theta), \ldots, \frac{\partial \eta}{\partial \theta_p}(\theta) \right) $$

(1.4)

$$ H(\theta) = \begin{pmatrix} \frac{\partial^2 \eta}{\partial \theta_i \partial \theta_j}(\theta) \end{pmatrix}_{i,j} $$

From an initial first guess of the parameters $\theta^{(0)}$, a sequence of parameter sets, $\{\theta^{(0)}, \ldots, \theta^{(k+1)}\}$, is generated that is intended to converge to the global minimum of $F(\theta)$ in the parameter space. If the model $\eta$ depends linearly on each parameter $\theta_j (j = 1, \ldots, \phi)$ the minimization problem stated in Eq. (1.2) reduces to a linear regression problem for which analytical solutions exist.

The derivative based search defined in Eqs. (1.3) and (1.4) will evolve towards the global minimum in the parameter space in situations where the response function (the objective function mapped out in the feasible parameter space) exhibits a topographical convex shape. However, practical experience with hydrologic models suggests that the objective function seldom satisfies these restrictive conditions. To illustrate the severity of the optimization problem, we closer examine a typical example in the field of vadose zone hydrology, which considers the inverse estimation of the soil water retention and unsaturated soil hydraulic conductivity characteristics using data from a transient one-step laboratory outflow experiment.
Figure 1.3 [taken from Vrugt and Bouten, 2002], presents the results of such an analysis, in terms of marginal probability distributions (histograms) for two soil hydraulic parameters in the vicinity of the global minimum of the parameter space.

![A](image.png) ![B](image.png)

**Figure 1.3.** Marginal probability distributions of two soil hydraulic parameters in the vicinity of the global minimum, (a) residual soil water content, and (b) inverse of air-entry value.

If the objective function would exhibit a convex shape in the entire parameter domain, the histograms for each of the soil hydraulic parameters would exhibit a clear Gaussian distribution with a single mode. However, the large number of different modes (local minima) for each of the parameters depicted in Fig. 1.3 is the most probable reason for the numerous reports in the literature of the inability to find a unique set of hydraulic parameter values using observed outflow dynamics from one-step outflow experiments [Kool et al., 1985; Parker et al., 1985; Toorman et al., 1992; van Dam et al., 1992; among others]. As the local gradient-based search algorithms are not designed to handle the peculiarities of the response surface illustrated in Fig. 1.3, they will terminate their search prematurely with their final solution essentially being dependent on the starting point in the parameter space. Another emerging problem, reported by Hopmans et al. [2002], is that some of the hydraulic parameters are typically highly correlated using observed outflow data, further lowering the chance of getting a single unique solution with local search methodologies.

The existence of non-uniqueness problems with local search methodologies has led soil hydrologists to argue that there is not sufficient information in the outflow measurements to enable a unique characterization of the soil hydraulic properties. So even with laboratory experiments, where a soil can be manipulated with great flexibility, non-uniqueness problems are often experienced [van Dam et al., 1992]. Seemingly, there was a widespread conviction that the best way to solve the non-uniqueness problem would be to add more and better measurements.
INTRODUCTION

[Etting and Hopmans, 1993; van Dam et al., 1994]. However, research into data requirements has led to the understanding that the information content of the data is far more important than the amount of data used for the identification of the model parameters [Sorooshian et al., 1983; Gupta and Sorooshian, 1985; Yapo et al., 1996]. Vrugt et al. [2003a] argued, therefore, that it is not the information in the measurements that is lacking to obtain a unique set of parameters, but the fact that the classical local-search optimization procedures utilized in soil hydrology are typically not powerful enough to solve the problems illustrated in Fig. 1.3. Their arguments on uniqueness of parameters are not based on the convergence problems of the applied optimization methods, but on the shape of the posterior marginal distributions of the parameters. Hence, closer inspection of Fig. 1.3 demonstrates that for each of the parameters there is a single optimum with highest posterior probability in the global minimum. Indeed, although not shown here, the multivariate probability distribution of the parameters confirms that these regions with highest posterior probability for each parameter coincide.

Initial responses to similar failure problems of local search algorithms in the area of watershed model calibration, during the 1980s were to try and put the optimization problem onto a more rigorous statistical footing using maximum likelihood or Bayesian theory [Sorooshian and Dracup, 1980; Kuczera, 1983]. However, these approaches did not directly address the causes of the inability to find the optimum for a selected objective function.

To be able to deal with the peculiarities of the response surface so evident in Fig. 1.3, we must, therefore, develop automatic calibration algorithms which can deal with the existence of multiple optima in the parameter space (with both small and large domains of attraction), discontinuous first derivatives and curving multi-dimensional ridges. These considerations inspired Duan et al. [1992] to develop a powerful, robust, and efficient, optimization procedure, entitled, the Shuffled Complex Evolution (SCE-UA) global optimization algorithm. By merging the strengths of the Downhill Simplex procedure [Nelder and Mead, 1965] with the concepts of controlled random search, systematic evolution of points in the direction of global improvement, competitive evolution [Holland, 1975], and complex shuffling, the SCE-UA algorithm represents a synthesis of the best features of several optimization strategies. The strength and reliability of the SCE-UA global optimization algorithm have since been independently tested and proven by numerous researchers and the algorithm is now extensively used world-wide [e.g., Duan et al., 1992; Sorooshian et al., 1993; Gan and Biftu, 1996; Kuczera, 1997; Hogue et al., 2000; Boyle et al., 2000; among many others]. With the current available SCE-UA algorithm, one can now have confidence that the global minimum of a predefined objective function is found. However, the
method does not provide any information about the uncertainty associated with the final estimated parameters.

1.4. Limitations of current model calibration strategies

Despite the progress made, several contributions to the hydrologic literature in the past decade have brought into question the continued usefulness of the classical paradigm for model calibration [Beven and Binley, 1992; Gupta et al., 1998; Kavetski et al., 2003]. For the purpose of this discussion, we classify these into four categories:

(1) The classical model calibration approach typically ignores parameter uncertainty. The classical method is based on traditional statistical non-linear regression theory, which operates under the central assumption that the available model structure is correct, and therefore seeks to identify a unique "optimal" set of parameter estimates, which minimizes the SLS estimator in Eq. (1.2). Practical experience with the calibration of hydrologic models suggests, however, that it is typically difficult to find a unique "best" parameter set whose performance measure (objective function) obviates considerable of other feasible parameter sets;

(2) The approach assumes the model parameters to be time invariant and aggregates model performance errors over a large range of hydrologic behavior. It is not very difficult to see that temporal and/or spatial aggregation of model residuals into a single additive performance statistic, such as the SLS estimator in Eq. (1.2), results in considerable loss of important information that can be used to distinguish between competing parameter sets;

(3) The approach gives little guidance how to combine different types of data sets during model calibration. With the growing popularity of sophisticated physically-based hydrologic models, the complexity of the calibration problem has been multiplied many-fold. For instance, the latest hydrologic watershed or land-surface models simulate several output fluxes (e.g. water, energy, chemical constituents, etc.) for which measurement data are available, and all these data must be correctly utilized to ensure proper model calibration;
INTRODUCTION

(4) The approach typically assumes that all the uncertainty in the input-output representation of the model can be attributed to uncertainty in the parameter estimates, thereby effectively neglecting input, output and model structural uncertainty. Hence, uncertainties in the modeling procedure not only stem from uncertainties in the parameter estimates, but also from measurement errors associated with the system input and output, and from model structural errors arising from the aggregation of spatially distributed real-world processes into a mathematical model.

Summarizing, current calibration techniques are not very well suited to the task of calibrating hydrologic models, and are particularly inadequate in the face of the emerging generation of multi-input-output distributed parameter hydrologic models.

1.5. Objectives of the PhD research

The preceding sections briefly described the historical development leading to current views on model calibration, and discussed some of the techniques that have been developed for estimating parameters, thereby enabling the model to mimic the behavior of the underlying hydrologic system. Finally, in the previous section the limitations of current model calibration strategies were established. In this section, the main objectives of the PhD research are summarized. Each of the listed objectives below have to be considered in the light of trying to improve upon the classical model calibration strategy, and form a response to each of the summarized points in the previous section.

(1) Development of methods, which not only estimate the traditional “best” parameter set, but simultaneously also generate a sample set of parameter values describing the probabilistic representation of the remaining parameter uncertainty. Instead of producing a single error-free model forecast, which is the case after conventional calibration, this posterior uncertainty can be used to produce probabilistic model outputs (most likely forecast and 95% confidence intervals at each time step). In this way, at least some of the uncertainty in the input-output representation of the model is represented in the model forecasts. Moreover, by analyzing the uncertainty of the parameters and their correlation structure induced in the posterior probability distribution, the hydrologist is able to better understand which complexity of the
model is warranted by the calibration data and which parameters to select for finding
useful regionalization relationships (those which are well defined after calibration).

(2) Development of methods, which increase information retrieval from the data by
recursively assimilating and processing the calibration data, thereby explicitly taking
into account parameter uncertainty. Careful inspection of one-dimensional
projections of the evolution of the high probability density region of the parameters,
will not only reveal which measurements are most informative for the model
parameters (useful information for optimal experimental design strategies), but
perhaps more importantly, provide a way for checking for violations of the
underlying assumption that parameters are constant;

(3) Development of methods, which explicitly recognize the multi-objective nature of
the model calibration problem and use innovative ways to deal with different
observation types, data sub-sets and global search algorithms to identify the non-
inferior solution space or preferred solutions. By analyzing the tradeoffs in the fitting
of different data sub-sets the hydrologist is able to better understand the limitations
of the current model structure.

(4) Development of methods, which more completely treat input, output, parameter, and
model structural uncertainty in hydrologic model calibration. Traditional calibration
methods assign all the uncertainty in the input-output representation to uncertainty in
the parameter estimates and miss the conceptual rigor to distinguish between the
various sources of errors associated with the application of hydrologic models. We
hypothesize that the implementation of such a strategy will not only result in
meaningful prediction uncertainty bounds on the model simulations, but will also
result in parameter estimates, which better represent system properties as they are
less corrupted by modeling errors. This will further enhance the prospects of finding
useful regionalization relationships.

Note, that efficiency and effectiveness of these computerized algorithms/methods are important
considerations.

It is important to stress, however, that the objectives of the PhD research should not be
understood as that of simply developing methods for parameter estimation alone. The
development and application of these methods is a necessary step to be able to better distinguish between input, output, parameter, and model structural uncertainty in hydrologic modeling. This information will help direct resources towards model structural improvements and, as such, will help to increase our understanding and knowledge of hydrologic processes. The envisaged higher goal of the PhD-research is, therefore, to contribute to an improved understanding of the various hydrologic processes operating at or near the earth’s surface. The usefulness and applicability of these methods is, therefore, continuously demonstrated by application to a variety of different hydrologic models throughout the remainder of this thesis.

1.6. Outline of this thesis

This thesis consists of a series of eight closely related chapters, which combined address the listed objectives in the previous section. To state the case and illustrate a typical parameter estimation problem in hydrology, Chapter 2 involves the calibration of soil physical and root water uptake parameters in a physically-based three-dimensional unsaturated soil water flow model using measured spatial distributions of soil water content around a sprinkler-irrigated almond tree. This chapter serves as an illustrative example to demonstrate the capabilities of inverse modeling (parameter estimation), and highlights some of the limitations of conventional parameter estimation algorithms. They are computationally very demanding when solving for high-dimensional parameter problems. Moreover, the emphasize lies on the identification of a single best set of model parameters, thereby effectively neglecting the influence of possible sources of uncertainty on the final parameter estimates.

To overcome these limitations, Chapter 3 presents the Shuffled Complex Evolution Metropolis (SCEM-UA) global optimization algorithm, which is designed to provide an effective and efficient estimate of the traditional “best” parameter set and its underlying posterior distribution within a single optimization run. The SCEM-UA algorithm is an improvement over the original SCE-UA algorithm and uses a probabilistic offspring strategy instead of the deterministic Simplex method to search the parameter space. The power and applicability of the newly developed SCEM-UA algorithm is demonstrated by application to a hydrologic watershed model with typical conceptual components.

Chapter 4 presents another application of the SCEM-UA algorithm by inverse estimation of large-scale spatially distributed vadose zone properties using the solution of a physically-based three-dimensional distributed model combined with spatially distributed measured tile drainage data from the 9700 ha Broadview Water District in the San Joaquin Valley.
of California. To study the benefits of using a spatially distributed three-dimensional vadose zone model, the results of the three-dimensional model were compared with those obtained using a simple conceptual bucket model and a spatial-averaged one-dimensional unsaturated water flow model.

The primary focus of the previous Chapters has been on the "batch" calibration approach, which aggregates error residuals over a large range of hydrologic behavior. To reduce the information loss associated with traditional batch processing, Chapter 5 presents a recursive parameter estimation method, entitled the Parameter Identification Method based on the Localization of Information (PIMLI), which is especially designed to increase information retrieval from the calibration data and to infer the type and location of the measurements, which contain the most information for the specific model parameters.

Despite the progress made, practical experience with the calibration of hydrologic models suggests that single objective functions, no matter how carefully chosen, are often inadequate to properly measure all of the characteristics of the observed data deemed to be important. 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. Chapter 6 presents the Multi-objective Shuffled Complex Evolution Metropolis (MOSCEM-UA) global optimization algorithm, which is capable of efficiently solving the multi-objective optimization problem for hydrologic models.

Chapter 7 surveys the conceptual limitations of the PIMLI, SCEM-UA, and MOSCEM-UA algorithms (amongst others) for distinguishing between input, output, parameter and model structural error, and presents SODA, a Simultaneous Optimization and Data Assimilation method, which more completely treats the various sources of errors in hydrologic model calibration. One of the goals of the SODA framework is to produce a time series of error residuals that have the desirable properties of constant variance and independence in time and space, so that unbiased estimates of the model parameters are obtained.

Finally, Chapter 8 presents an epilogue, in which the most important research questions that can now be forcefully addressed with the availability of the SCEM-UA, PIMLI, MOSCEM-UA, and SODA parameter and state estimation methods are highlighted and a view on future research in model calibration is explicated. A summary of all the preceding chapters is given in the conclusion section of this thesis.