Towards improved treatment of parameter uncertainty in hydrologic modeling
Vrugt, J.A.
CHAPTER 8

Epilogue

"Towards Improved Treatment of Parameter Uncertainty in Hydrologic Modeling"

To-date, uncertainty in hydrologic modeling has mostly been neglected, thereby elusively focusing on a single best parameter set that meets the needs and objectives of the modeling procedure. In this thesis, resources have been directed towards the development of stochastic and recursive single- and multi-objective parameter estimation algorithms that, combined with state-of-the-art data assimilation methods, can be used to more completely treat the various sources of uncertainty in hydrologic modeling, with a particular emphasis on parameter uncertainty. The need for parameter estimation and data assimilation tools continues to grow as advancements in remote sensing, communications, computing, and information technologies allow engineers and scientists to work with more complex computer models, spatially distributed data sets and real-time information. With the SCEM-UA, PIMLI, MOSCEM-UA, and SODA parameter and state estimation methods at our disposal we can now forcefully address a variety of research questions, which could not have been easily answered without these algorithms. In this epilogue, some of these most important research questions related to hydrologic modeling are highlighted, and a view on future research in model calibration is explicated.
CHAPTER 8

With the availability of the SCEM-UA, PIMLI, MOSCEM-UA, and SODA parameter and state estimation methods we can now forcefully address the following issues:

(1) Diagnosing systematic errors — Detailed investigations of the model prediction uncertainty ranges associated with the posterior (SCEM-UA) and Pareto (MOSCEM-UA) set of parameters, will reveal important information about systematic (auto-correlated) errors.

Theoretically, if the model structure would be correct, and the input and output data are observed without error, the model output prediction uncertainty ranges, associated with the uncertainty in the parameter estimates, will bracket the observations. Any systematic departure from this is not caused by parameter uncertainty, but due to the combined effect of input, output and model structural error (see for instance Figures 2.11 and 5.9). The insights that were developed with this analyses, led to the development of SODA in Chapter 7 of this thesis.

(2) Parameter stationarity — Detailed investigations of one-dimensional projections of the evolution of the High-Probability Density (HPD) region of the parameter space over a historical record of calibration data, provide a way for checking whether the hydrologic system has undergone changes. A calibrated model can only be reliably used for the simulation and prediction of hydrologic events outside the calibration period, if it can be reasonably assumed that the physical characteristics of the watershed and the hydrologic/climate conditions have remained similar. For instance, alterations in land-use will result in a different response of the watershed to precipitation forcing. Systematic trends in parameter variation can also be used to diagnose and quantify model structural errors, although this has yet to be satisfactorily demonstrated.

As the PIMLI and SCEM-UA algorithmic successfully infer the posterior distribution of the model parameters, these methods are suited to investigate whether model parameters are statistically stationary over a historical record of data or if they are correlated to varying characteristics of the underlying hydrologic system. As an illustrative example, consider Figure 8.1, which presents the evolution of the SCEM-UA derived marginal HPD regions for each of the parameters in the Sacramento Soil Moisture Accounting model of the National Weather Service of the USA (SAC-SMA) using a 36-year historical record of streamflow data for the Leaf River Watershed [taken from Vrugt et al., 2004]. The uncertainty bounds presented in Fig. 8.1 denote averages over a window of 6 years (1953 denotes calibration results over the WY 1953-1959 respectively and so forth).
Figure 8.1. Evolution of the 95% confidence intervals for each of the SAC-SMA parameters over the 36-year historical record of the Leaf River watershed. The gray shaded area in each parameter plot represents the HPD region, whereas the marked squares and dotted line refer to the SCE-UA solution and most likely SCEM-UA solution respectively. Results denote averages over a window of 6 years. Definition of the SAC-SMA parameters appears in Table 6.2.
To allow comparison between uncertainties of different parameters, the HPD region was scaled according to the prior uncertainty bounds of the parameters, as presented in Table 6.2. to yield normalized ranges between 0 and 1. The gray-shaded area in each parameter plot represents the HPD region, whereas the marked squares and dotted line refer to the SCE-UA solution, and most likely SCEM-UA solution within the HPD region, respectively. Definition and explanation of the SAC-SMA model parameters appears in Table 6.2. While some of the SAC-SMA parameters show little variation over the 36-year historical data record, the HPD region for other SAC-SMA parameters traverses through the feasible parameter space. Especially, there is considerable variation and uncertainty associated with the parameters lower zone parameters LZSK, and LZPK, which primarily determine the shape of the hydrograph during the recession periods, and the percolation parameters REXP and PFREE. The apparent systematic variation of the parameters ADIMP, LZSK and LZPK with time, might suggest that the watershed has undergone hydrologic changes. When some of the parameters in the SAC-SMA model are plotted against the mean areal rainfall over the calibration periods, as done in Figure 8.2, a relationship becomes apparent.

![Figure 8.2](image)

Figure 8.2. Two-dimensional plot of the most likely parameter value, indicated with a squared symbol, versus the mean areal rainfall for the SAC-SMA parameters ADIMP (a), LZTWM (b), LZSK (c), and LZPK (d). The bars around the most likely parameter value denote the size of the HPD region.
To be able to match the observed hydrograph with increasing wetness of the years, the additional impervious fraction is decreased (ADIMP), while the maximum capacity of the lower zone tension water storage (LZTWM) and depletion rate from the lower zone need is to be increased. Seemingly, parameters calibrated for relatively dry years, result in sub-optimal forecasts for the wettest years on record and vice versa. This nonstationarity with increasing wetness of the years for some of the SAC-SMA parameters, point towards aspects of the model structure that needs to be further refined.

(3) Parameter identifiability – Detailed investigations of the SCEM-UA derived posterior mean, standard deviation, coefficient of variation and Pearson correlation coefficients between the samples in the HPD region of the parameter space, (i) facilitates the selection of an adequate model structure, (ii) guides in the development of pedotransfer functions, (iii) helps to assess how much model complexity is warranted by the available calibration data, and (iv) guides the development of optimal experimental design strategies.

(3.1) With respect to issue (i), consider Figure 8.3, which presents two-dimensional scatterplots of the SCEM-UA derived posterior parameter samples for identical counterparts in the soil water retention models of Brooks and Corey [BC, 1964], van Genuchten [VG, 1980], and Kosugi [KS, 1996] [taken from Vrugt et al., 2003a].

Figure 8.3. Scatter plot of 4000 combinations of $\theta$-$b_6$ (a), $b_6$-$\lambda$ (b), $\theta$-$\lambda$ (c), $\theta$-$\alpha$ (d), $\alpha$-$\sigma$ (e), $\theta$-$r$ (f), $\theta$-$b_{6.5}$ (g), $b_{6.5}$-$\sigma$ (h) and $\theta$, $\sigma$ (i) parameters sampled for the clayey soil using the SCEM-UA algorithm for the BC-model (a-c), VG-model (d-f) and KS-model (g-i), respectively.
Each of the retention models contains four parameters, whose values need to be estimated by fitting to observed retention data. While the fit of each of the different parametric models to the observed retention data was generally excellent (not demonstrated here), Fig. 8.3 depicts that the correlation coefficients between the parameters are generally higher in the KS-model (Figs. 8.3g-i) as compared to the BC-model (Figs. 8.3a-c) and VG-model (Figs. 8.3d-f). More specifically, whereas in the BC and VG-model, only the parameters $\theta$, $\lambda$ (Fig. 8.3c), and $\alpha$, $\eta$ (Fig. 8.3f) are strongly correlated for the clayey soil, in the case of the KS-model three pairs of parameters exhibit considerable correlation, $\theta$, $h_{0.3}$ (Fig. 8.3g), $h_{0.5}$, $\sigma$ (Fig. 8.3h), and $\theta$, $\sigma$ (Fig. 8.3i).

The strong hyperbolic-shaped correlation structure being associated with the parameters $\theta$, $h_{0.3}$ and $\sigma$ in the KS-model not only decreases the identifiability of the parameters, but also enhances non-uniqueness problems of the final optimized parameters. With recourse to the quality of the fit only the competing parametric models of VG and KS appear to fit the experimental data equally well. However, when examining other aspects, such as the multivariate correlation structure as inferred from the joint distributions of the SCEM-UA generated samples (Fig. 8.3), it becomes clear that there are drawbacks associated with the correlation structure of the parameters in the KS-model for fine textured soils. It is clear that when possessing comparable predictive capabilities, the non-linear model that closely approaches linear behavior is in favor. Besides identifiability problems of the parameters, for linear models less iterations are necessary to achieve convergence in parameter estimation, and traditional computational undemanding first-order statistical inferences will be more valid.

(3.2) With respect to issue (ii), an identifiability analysis of the parameters constitutes important information for studies that aim to find pedotransfer functions, relating the calibrated parameters to catchment characteristics. Uniqueness of the parameters is a prerequisite to be able to successfully find pedotransfer functions. To successfully find pedotransfer functions, it seems most productive to concentrate on those model parameters that are well defined after direct estimation (calibration).

(3.3) To demonstrate the usefulness of studying parameter identifiability to assess the information content of calibration data (issue iii), consider Table 8.1, which presents the SCEM-UA derived posterior mean, standard deviation, coefficient of variation (CV), and correlation structure induced between the parameters of a four-parameter single-layer canopy rainfall interception model.
Table 8.1. SCEM-UA derived posterior mean, coefficient of variation (CV [%]), and Pearson correlation coefficients between the parameters in the single-layer interception model obtained when using half-hourly measurements of throughfall or canopy storage at DOY 211. The symbols $a$, $b$, $c$, and $d$ refer to the interception efficiency, drainage, storage capacity, and evaporation efficiency parameter, respectively.

<table>
<thead>
<tr>
<th>Par.</th>
<th>Unit</th>
<th>Mean.</th>
<th>CV [%]</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>[-]</td>
<td>0.92</td>
<td>4.63</td>
<td>1.00</td>
<td>-0.12</td>
<td>-0.00</td>
<td>-0.06</td>
</tr>
<tr>
<td>$b$</td>
<td>[d$^{-1}$]</td>
<td>496.08</td>
<td>38.55</td>
<td>-----</td>
<td>1.00</td>
<td>0.13</td>
<td>-0.01</td>
</tr>
<tr>
<td>$c$</td>
<td>[mm]</td>
<td>1.92</td>
<td>32.53</td>
<td>-----</td>
<td>-----</td>
<td>1.00</td>
<td>-0.86</td>
</tr>
<tr>
<td>$d$</td>
<td>[-]</td>
<td>1.26</td>
<td>18.87</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>1.00</td>
</tr>
</tbody>
</table>

These results were derived when assimilating and processing half-hourly throughfall and water storage measurements at the Speulderbos forest in the Netherlands [taken from Vrugt et al., 2003b]. The high coefficients of variations (CV) and standard deviations of the parameters in Table 8.1 demonstrate that measured throughfall dynamics contain only very limited information for the calibration of a canopy interception model. Although there does not exist an exact quantitative threshold to help judge whether a parameter is identifiable or not, the rationales on the identifiability of the interception model parameters that we adopt in this section are based on the extent of the HPD region in the physical plausible or prior parameter space. For instance, it is clear that the storage capacity ($c$), the evaporation efficiency ($d$), and the drainage parameter ($b$) are poorly defined by calibration to measured throughfall dynamics, as for each of these parameters there does not exist a well-defined region, in the sense of a compact region, interior to the physical plausible or prior parameter space. On the contrary, measured water storage dynamics contain sufficient information to be able to identify most of the model parameters with a high degree of confidence. Hence the CV values and standard deviation of the parameters, as reported in Table 8.1, illustrate that for most of the parameters the HPD region occupies only a very small portion of the prior parameter space. Especially the storage capacity and the evaporation efficiency parameter are very well determined by calibration to measured water storage dynamics. Unfortunately, like measured throughfall dynamics, canopy water storage
measurements contain insufficient information to identify the drainage parameter with a satisfying accuracy.

(3.4) With respect to issue (iv), recursive parameter estimation methods, such as PIMLI facilitate the identification of the most informative measurements for model calibration, and as such help in the development of optimal experimental design strategies. Although this has been extensively discussed and demonstrated in Chapter 5, for completeness another illustrative example in hydrometeorology is discussed here [taken from Vrugt et al., 2003b]. Figure 8.4, presents the evolution of the HPD region of the posterior probability density, in the form of one-dimensional projections for each of the parameters in the single-layer canopy interception model, derived when stepwise assimilating and processing the throughfall (Figs. 8.4b–e) or canopy water storage (Figs. 8.4g–j) measurements of Day of Year (DOY) 211 with the PIMLI algorithm. The dark-shaded line marks the evolution of the most likely parameter set at each time step, whereas the asterisks denote the most likely parameter values derived using the SCE-UA global optimization algorithm.

Starting at $t = 211.4$, the size of the HPD region of the parameters, reflects the initial or prior uncertainty of the parameters before any data are collected and processed. Immediately after the first rain event at $t = 211.55$, the uncertainty associated with the interception efficiency parameter $a$ decreases, and remains rather constant thereafter, demonstrating that rain events during the wetting stage of the forest canopy contain the most information for the identification of the $a$-parameter. Although the $a$-parameter is reasonable well determined by calibration to measured throughfall dynamics (see Fig. 8.4b), the additional throughfall measurements between the first rain event and the drying cycle starting at the beginning of DOY 212 contain very limited information for the model parameters. Hence the one-dimensional projections of the evolution of the HPD region of the posterior probability density for $b$, $c$, and $d$ (see Figs. 8.4c–e) suggest that for these parameters there does not exist a well defined region in the sense of a compact region interior to the prior parameter space when calibrating on measured throughfall dynamics.
Figure 8.4. Evolution of the HPD region of the posterior probability density in the form of one-dimensional projections of the parameters (light-gray region) using measured throughfall dynamics (b-e), or canopy water storage observations (g-j). The dotted line denotes the evolution of the most likely parameter set, whereas the asterisks indicate the “best” parameter values obtained using a traditional batch calibration with the SCE-UA global optimization algorithm. The solid circles in the figures a and f denote measured values of throughfall and canopy water storage respectively, whereas the dotted line in these figures denote the model predicted values corresponding to the parameter set with the highest posterior probability.
On the contrary, the evolution of the Bayesian confidence intervals of the parameters depicted in Figs. 8.4g–j illustrate that measured canopy water storage dynamics contain sufficient information to uniquely identify at least three of the interception model parameters \((a, c, \text{ and } d)\). Moreover, these parameters are identifiable at different stages during the wetting and drying cycles, thereby facilitating the identification of a unique set of parameters. Unfortunately, no information is found for the drainage parameter. Fig. 8.4 demonstrates very well that adding more data does not simply solve the problem of parameter identifiability. Only specific data periods with high information content can reduce the uncertainty associated with the interception model parameters.

Although not explicitly demonstrated here, the use of longer observational time series of throughfall dynamics for calibration purposes, leads to similar findings. While it might seem speculative to generalize the conclusions regarding the identifiability of the interception model parameters using measured throughfall dynamics to other climates, species, or biomasses situations, additional investigations with numerically generated throughfall “measurements” for other situations than what was presented in this Epilogue yielded similar results. We subscribe ourselves, therefore, to the view that model parameters of drainage and evaporation functions, which are obtained by calibration against measured throughfall dynamics must be interpreted with care as these parameters are subject to considerable uncertainty.

(4) Quantifying uncertainty in regionalization studies — A topic, which is currently receiving considerable attention in the hydrologic community, is the simulation of the rainfall-runoff behavior of ungauged watersheds. The International Association of Hydrological Sciences (IAHS) has recently launched a 10-year initiative, called the IAHS Decade for Prediction in Ungauged Basins (PUB), to address this problem.

When we attempt to generate hydrologic predictions for ungauged watersheds, model parameters cannot longer be estimated by calibration to streamflow observations, and need to be estimated from other sources of information, such as neighboring catchments, expert judgment or tabulated literature values. One approach to derive the values of model parameters for ungauged watersheds, which is particularly popular in hydrology, is to use established statistical relationships between model parameters and watershed characteristics for gauged basins. This process of transferring parameters from gauged to ungauged catchments is generally referred to as generalization in surface hydrology, although soil hydrologists prefer the term pedotransfer functions. While significant progress has been made in the development and application of pedotransfer functions to simulate the hydrologic behavior of ungauged watersheds, most of the
reported studies in the literature typically ignore parameter uncertainty and uniqueness in the regionalization process. Usually predictions in ungauged watersheds are given as point estimates. The availability of the SODA framework now enables the quantification of parameter uncertainty for gauged watersheds, while taking into account, input, output and model structural errors. This parameter uncertainty can be explicitly used in the derivation of transfer functions, thereby propagating modeling uncertainty from gauged to ungauged watersheds. Certainly, this will result in a stochastic estimate of the hydrologic response of the ungauged watershed, which is more useful for decision-making since it considers the confidence in the model predictions.

In principle, the summarized points under 2-4 can also be extended to multi-objective parameter estimation problems. The domain of interest is then Pareto uncertainty instead of probabilistic parameter uncertainty. This also provides new and useful ways to define the information content of data, and to evaluate different model structures and their performance.

(5) Improving hydrologic models – As mentioned in the introduction section of this thesis, the objectives of the PhD research should not be understood as that of simply developing methods for parameter estimation alone. Instead, the envisaged higher goal of the PhD research is to use the SCEM-UA, MOSCEM-UA, PIMLI and SODA algorithms to contribute to an improved understanding of the various hydrologic processes operating at or near the earth’s surface.

To be able to better understand the limitations of our models and improve our understanding and theory of hydrologic processes, we need to develop strategies, which can distinguish between input, output, parameter, and model structural uncertainty. The SODA framework presented in Chapter 7 of this thesis has been designed to facilitate this task. Because the inversely identified parameters with SODA are less corrupted by modeling errors, detailed interpretation of the time series of output and state innovations will reveal useful information about model structural errors (see Fig. 7.11). These findings can be used to reformulate the model and as such to improve our understanding of hydrologic processes.

The SODA method is an attempt to more completely treat these various sources of uncertainty. However, still progress in this matter is possible. I believe that future research on model calibration would be most productive if focused on the following issues, (1) the derivation and implementation of realistic input error models to further distinguish between input and model structural error, (2) the development and implementation of higher order filters which generate improved estimates of output data errors, (3) the extension to multicriteria problems;
SODA does not exclude the use of different parameter sets to match different portions of hydrologic behavior, (4) the use of recursive parameter estimation methods to further minimize the short term model bias using the state augmentation technique, (5) the development of filtering methods which further relax the Gaussian assumptions of the error distributions in classical KF implementations. With respect to this last issue, Gaussian mixture models are able to represent arbitrarily complex probability density functions. This fact makes them an excellent choice for representing complex likelihood functions for recursive Bayesian and multi-objective parameter estimation or filtering techniques, and (6) the extension of SODA to multiple competing models running in parallel (improved Bayesian Model Averaging).

The work presented in this thesis has primarily focused on the development and application of optimization techniques to quantify parameter uncertainty in hydrologic models. However, as emphasized in Chapter 7, it does not seem reasonable to attribute all the uncertainty in the modeling procedure to uncertainty in the parameter estimates. In the same Chapter we therefore outlined the more general SODA framework to simultaneously deal with input, output, parameter, and model structural uncertainty in hydrologic modeling. It is to be expected that further developments within this framework, in combination with the ever increasing pace of computational power, and the availability of spatially distributed data, will advance our understanding of the functioning of environmental systems.