Assessing and Reducing the Uncertainty of Predictions from Hydraulic and Hydrologic Models

Progress Report Dated July 2013
Project ID 9320
Jeff Niemann

1. Project Objective

The U.S. Bureau of Reclamation (USBR) uses many kinds of computational hydrologic, hydraulic, and sediment-transport models to protect and manage water resources. Such models are used to forecast future conditions and to analyze the impacts of potential changes to water systems, and the results from these models guide countless decisions made by USBR and its stakeholders. Unfortunately, the results of computational models inherently contain uncertainty due to uncertainty in the model inputs, the values of the model parameters, and the mathematical representation of the system. This uncertainty ultimately forces conservative decisions, which increases project costs and exacerbates the conflicts between the various project objectives and stakeholders.

USBR currently uses simplistic methods to assess uncertainty. These methods usually focus on simulating perceived worst-case scenarios and evaluating the behavior of the system under these scenarios. The central problem with this approach is that it is not statistically rigorous. The scenario is determined by professional expertise, so no probability of occurrence can be assigned to it, and decision makers have little confidence that the results accurately reflect the uncertainty. In addition, this approach does not directly connect the certainty in the model inputs, parameters, and mathematical structure to the determination of the worst-case scenario. As a result, this approach cannot give the project team guidance about how to reduce the uncertainty.

The overall objective of this project is to develop a method to assess and potentially reduce uncertainty in the forecasts from hydrologic, hydraulic, and sediment-transport models. We mainly focus on sediment transport models because they are typically the most complicated and computationally intensive class of models. To be feasible for widespread use, the new method must require few enough simulations to be applied to complex model applications in a reasonable amount of time yet retain enough formality to evaluate strategies to reduce uncertainty.

2. Previous Accomplishments

The first phase of the project focused on identifying a more formal method for uncertainty estimation that has been used in other fields and adapting it to apply to hydrologic, hydraulic, and sediment-transport models. The method that was identified is Generalized Likelihood Uncertainty Estimation (GLUE), which is a method to assess the uncertainty in forecasts due to the uncertainty in the parameters (Beven and Binley, 1992). This method was selected because it is among the simplest methods to assess parameter uncertainty. GLUE has been widely-used with hydrologic models, but it has a key limitation that restricts its use for hydraulic and sediment-transport models. Specifically, hydrologic models typically have calibration observations for only one output variable (discharge) whereas sediment transport models, for example, can have observations for multiple output variables (e.g., water surface elevations, stream bed elevations, and sediment grain size distributions). For this reason, GLUE was generalized to consider multiple output variables (Ruark et al., 2011). This generalization was accomplished by developing a new function in GLUE that calculates the likelihood that a given parameter set is correct given calibration observations for multiple variables. The new likelihood function weighs the observations of different output variables based on the sensitivity of each parameter to each output variable. The sensitivities are calculated using Fourier Amplitude Sensitivity Test (FAST) (Saltelli et al., 1999). The new likelihood function not only allows for multiple output variables but also uses a more statistically formal formulation than the ones that are commonly used with GLUE.

The multi-objective GLUE method was evaluated by coupling it to the Sedimentation and River Hydraulics - One Dimension (SRD-1D) model and applying it to two flume experiments (an erosional case and a depositional case). While the method is generally successful at quantifying the impacts of parameter uncertainty, three key limitations were identified. First, the method potentially requires too many model simulations to be used when simulations require long computation times. Second, the method assumes that the probability distributions that describe the uncertainty in the parameter values after calibration (i.e. the posterior distributions) are independent of each other, but the results indicate some interdependence. Third, the method does not consider the uncertainty in the forecasts due to uncertainty in the mathematical structure of the model. Structural uncertainty might be particularly important in sediment transport modeling because multiple empirical transport laws are available that can potentially give rather different forecasts.

In the second phase of the project, the importance of these limitations was evaluated by applying more sophisticated uncertainty methods and comparing their results to the multi-objective GLUE. The more sophisticated method to assess parameter uncertainty is called Shuffled Complex Evolution Metropolis – Uncertainty Estimation (SCEM-UA) (Vrugt et al., 2003). This approach is potentially quicker than GLUE because it uses information gained from the results of each model simulation to determine the configuration of the next simulation in the analysis. In GLUE, all model simulations are performed independently, which means that it can perform simulations that are potentially unimportant to the results. SCEM-UA also accounts for the dependence between the parameter distributions after calibration. SCEM-UA was generalized to allow for multiple output variables and called Multi-Variate Shuffled Complex Evolution Metropolis – Uncertainty Analysis (MSU) (Sabatine, 2012). To include the an assessment of structural uncertainty, MSU was used in combination with Bayesian Model Averaging (BMA) (Hoeting et al., 1992). BMA allows one to consider the uncertainty that is introduced by the selection of a particular sediment transport law.

MSU and BMA were coupled with SRH-1D and applied to the same flume experiments that were considered with the multi-objective GLUE as well as an additional flume experiment (an erosional case with more data). The results from MSU indicate that including the correlation between the most likely parameter values significantly alters the estimated forecast uncertainty. However, the MSU algorithm requires approximately the same number of simulations to evaluate uncertainty as GLUE. The results from BMA suggest that a combination of transport equations usually provides a better forecast than an individual equation. The BMA results also indicate that structural uncertainty is an important contribution to overall uncertainty. Overall, the results of MSU and BMA provide a rigorous evaluation of uncertainty that can be used to evaluate the performance of proposed quicker and simpler methods.

3. Most Recent Results

In the most recent period of the project, a real river system was identified for further testing of the uncertainty methodologies. All previous testing considered relatively simple flume experiments where either erosion or deposition dominated throughout the entire reach. These experiments also consider small spatial extents and short time periods, so they can be simulated relatively quickly. A real river system was selected that included regions of both erosion and deposition and reflects the computational requirements that are more commonly faced by USBR modelers. The selected case study is the Tachia River in Taiwan.

MSU was used to identify the uncertainty in the parameter values and the implications of parameter uncertainty on forecasts at the Tachia River. MSU begins by assuming that each parameter conforms to a uniform distribution between specified limits. This uniform distribution implies that there is no prior knowledge about the parameter values before they are calibrated aside from reasonable limits (the distribution bounds). In the analysis, critical shear stress, hiding factor, active layer thickness multiplier, deposition recovery factor, scour recovery factor, bed load adaptation length, and weight of bed load fractions are treated as uncertain. Manning's roughness coefficient is treated as a certain parameter using a value that was suggested by the studies of the Water Resources Agency of Taiwan.

MSU operates by first generating a relatively small sample of parameter sets from the uniform distributions and then sorting these parameter sets into complexes. One parameter set in each complex is used as the first point of a Markov chain. Trial parameter sets are then generated from a proposal distribution and the current point in the Markov chain. The trial parameter set is retained (and the chain and complex are updated) based on the likelihood that the trial parameter set is the correct choice for the case study being simulated. The likelihood is evaluated by using the parameter set in the SRH-1D model and then simulating the case study's calibration period, which has observations for both the model forcing variables and the key output variable or variables. If the simulation results closely match the observed outputs, then the parameter set is considered more likely. The chains associated with each complex are updated for several iterations, and then the parameter sets in the complexes are shuffled. From an optimization perspective, each Markov chain iterates toward a local optimum for the parameter values. The parameter sets are shuffled to help ensure that the global optimum parameter set is found. MSU also iterates so that the generated parameter sets eventually conform to the distribution that describes their uncertainty after calibration (the posterior distribution).

For the Tachia case study, an initial population size of 250 parameter sets was organized into 2 complexes so that each complex contained 125 parameter sets at any one time. The algorithm was set to iterate for a total of 10,000 simulations to be certain that all parameters converged and large samples from the posterior parameter distributions were attained. Only the volume of sediment deposition was used as an output variable for the calibration period due to data limitations for this case study. Thus, MSU simplifies to SCEM-UA in this application because only one output variable is observed. The period from 2000 to 2005 was used as the calibration period, so the conditions in 2000 were provided to SRH-1D as a starting point, and the parameters were evaluated based on the model's ability to predict the 2005 observations. Observations are available at 48 different locations for each time. In the application of the method, it is assumed that the volumes of sediment deposited at all locations at a given time have the same variance of their residuals and can be treated as a single output variable. The scale of the measurements at all locations at a given time does not vary greatly, but the scale of these measurements can change with time, which would likely imply a change in the variance of the residuals as well. The parameter uncertainty was evaluated when SRH-1D uses three different sediment transport equations: the Parker (1990) equation, the Wilcock and Crowe (2003) (W&C) equation, and the Wu (2000) equation.

The Tachia River case study provides some valuable insights into the assessment of parameter uncertainty. The MSU algorithm converges when it is sampling from the stable posterior distribution. Because more than one Markov chain is used in the method, convergence can be measured by the ratio of the variance of the average parameter value from each chain and the average of the variances of parameter values within each chain. This ratio is the basis of Gelman and Rubin's Scale Reduction Score (SRS). MSU has exactly converged when the SRS for all parameters is equal to 1. Because this is very difficult to achieve in practice, SRS values of less than 1.2 are usually used to indicate approximate convergence (Vrugt et al., 2003; Gelman and Rubin, 1992). MSU achieves convergence under 1500 iterations with the Parker, W&C, and Wu equations for the Tachia River case. This number is lower than what was observed for the data rich flume experiments, and it implies that a formal method such as MSU might be practical although time-consuming for models of real river systems.

To evaluate the extent to which the parameter values are constrained by the calibration data in this model application, one can compare the posterior parameter histograms to the prior histograms. Figure 1 shows the prior (black) and posterior (gray) histograms for the Parker, W&C, and Wu equations. For all three transport equations, both histograms for the roughness coefficient indicate a single value because the roughness coefficient was constrained as constant during the simulations. For the Parker equation, the critical shear stress and active layer thickness multiplier the best constrained by the calibration data and appear to have most likely values near the middle of their ranges. However, considerable uncertainty remains in those values. The other parameters are poorly constrained because their histograms are still wide and relatively similar to the uniform prior histograms. For the W&C equation, the critical shear stress, hiding factor, and active layer thickness multiplier are more constrained by the data and appear to have a symmetrical distribution around the middle of the ranges, but those parameters also contain considerable uncertainty. The other parameters are poorly constrained. For the Wu equation, most parameters are somewhat constrained by the calibration data except the active layer thickness multiplier and scour recovery factor. Overall, these results suggest that the available observations are not adequate to infer the values for most parameters and much uncertainty remains after calibration.

Figure 2 compares the simulated sediment volumes to the observations for the most likely parameter set associated with each transport equation. Volumes of erosion and deposition were estimated by computing the average bed change from the 2000 to 2005 cross sections, then multiplying the average bed change by the width of the active cross section and the distance between cross sections. The comparison of deposition volumes from the Parker, W&C, and Wu equations at each individual cross sections from the downstream end (section number 1) to the Shih-Gang Dam (number 36) are shown in Figure 2. These results suggest another reason that uncertainty in the parameter values is so great: even the best performing (most likely) parameter set does not reproduce the observed patterns of erosion and deposition well. The sum of squared errors (SSE) indicates that the Wu equation has the best performance, but the Parker equation best represents the drastic changes in deposition volume between sections 15 and 23.

4. Path Forward

At this stage of the project, a diverse set of test cases has been assembled, and the uncertainty in these test cases has been evaluated using the formal methods of MSU and BMA. Those results provide a benchmark against which quicker and more approximate methods can be compared. To complete the project, three major tasks remain:

1. Develop a simplified methodology that requires fewer simulations to evaluate uncertainty
2. Evaluate the performance of this method by applying it to the test cases considered previously

3. Implement the method in streamlined software, train USBR staff in its use, and publish the results in refereed journals

Faster convergence of the method can be achieved by better representing the knowledge that the modeler has prior to uncertainty estimation. The previous methods assume that the modeler has no knowledge aside from the bounds for the parameter distributions, but an experienced modeler can predict in advance the parameters that will have little effect on the results, which could be treated as certain with little effect on the uncertainty estimation. In addition, an experienced modeler has some knowledge of realistic parameter values for a given case, and such knowledge could be incorporated into the forms of the prior distributions. Also, some models may be calibrated manually before the uncertainty estimation is performed, which should also be reflected in the prior distributions.

We plan to exploit this knowledge to make the uncertainty methods quicker. First, a screening algorithm based on sensitivity analysis will be conducted to identify the parameters that contribute the most to output variability. In the screening analysis, parameter sets will be generated by individually varying the value of each parameter within its plausible range while fixing the values of the other parameters. Each generated parameter set will be used to simulate the calibration period, and the variability in the outputs will be evaluated. Parameters that introduce much variation need to be treated as uncertain, while parameters with little effect can be treated as certain. Using this approach, the number of parameters and thus the number of required simulations can be reduced. Second, the prior parameter distributions can be defined based on the modeler's prior knowledge of the parameter values. A beta distribution can be used for the uncertain parameters, and the minimum, maximum, mode, 75th percentile values of the parameter can be used to determine the shape of the beta distribution. The modeler can determine the minimum and maximum values from the physically possible range for that parameter (similar to the method used to define the uniform distributions). The mode can be determined from the manually calibrated value, and the 75th percentile can be determined from the modeler's assessment of the plausible range of values for the application case.

The revised algorithm will be tested by applying it to the same case studies that have been evaluated previously. The number of required simulations will be compared to the previous numbers of simulations to assess the improvement in speed for the simplified approach. Similarly, the estimates of forecast uncertainty that are generated by the simplified method will be compared to the more rigorous estimates determined previously to evaluate the degree of approximation. The method will also be applied to determine data collection strategies that would most effectively reduce the forecast uncertainty.

Finally, after the method has been developed and tested, an efficient implementation will be developed in Matlab and delivered to USBR staff for their routine use. The software will be freely available to USBR and to the public. Moreover, the results of the project will be published in progress reports and refereed literature to ensure broad dissemination and public use of the methodology.

References

Beven K, Binley A. 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes* **6** : 279-298. DOI: 10.1002/hyp.3360060305.

Gelman, A. and Rubin, D.B. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* **7**: 457-472.

Hoeting JA, Madigan D, Raftery AE, Volinsky CT. 1999. Bayesian model averaging: A tutorial. *Statistical Science* **14** : 382-417. DOI: 10.1214/ss/1009212519.

Parker, G. 1990. Surface-based bedload transport relation for gravel rivers. *Journal of Hydraulic Research* **28:** 417-436.

Ruark, M.D., Niemann, J.D., Greimann, B.P., Arabi, M. 2011. Method for assessing impacts of parameter uncertainty in sediment transport modeling applications. *Journal of Hydraulic Engineering* **137**: 623-636.

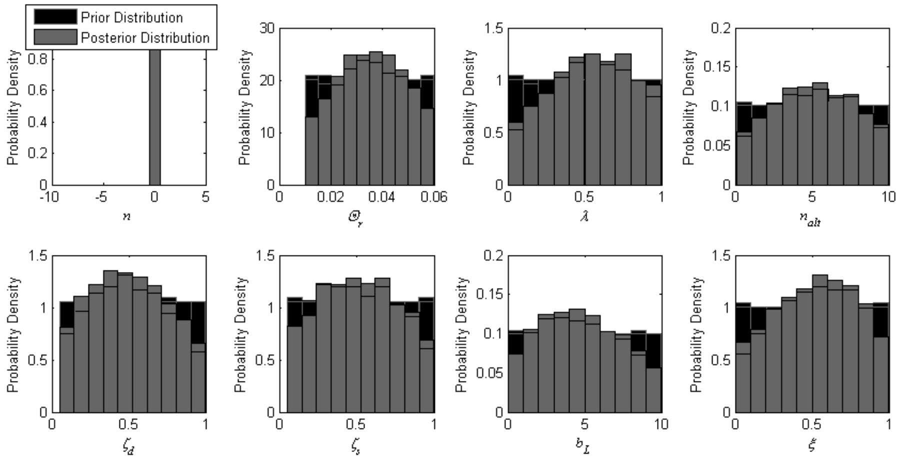
Sabatine, S.M. 2012. Evaluation of Parameter and Model Uncertainty in Simple Applications of a 1D Sediment Transport Model. Master of Science Thesis, Colorado State University, Fort Collins, Colorado.

Saltelli, A., Tarantola, S., and Chan, K.P.-S. 1999. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics* **41**(1): 39-56.

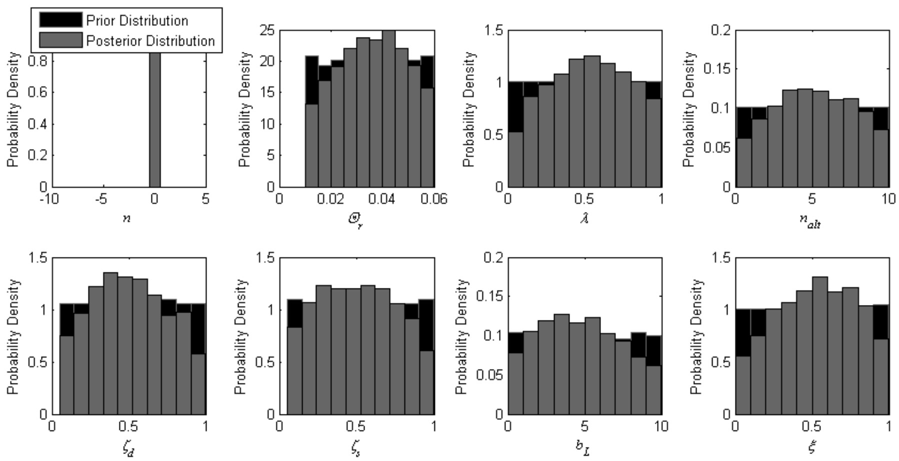
Vrugt, J.A., Gupta, H.V., Bouten, W., and Sorooshian, S. 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research* **39**: 1201-1216.

Wilcock, P.R. and Crowe, J.C. 2003. Surface-based transport model for mixed-size sediment. *Journal of Hydraulic Engineering* **129**: 120-128.

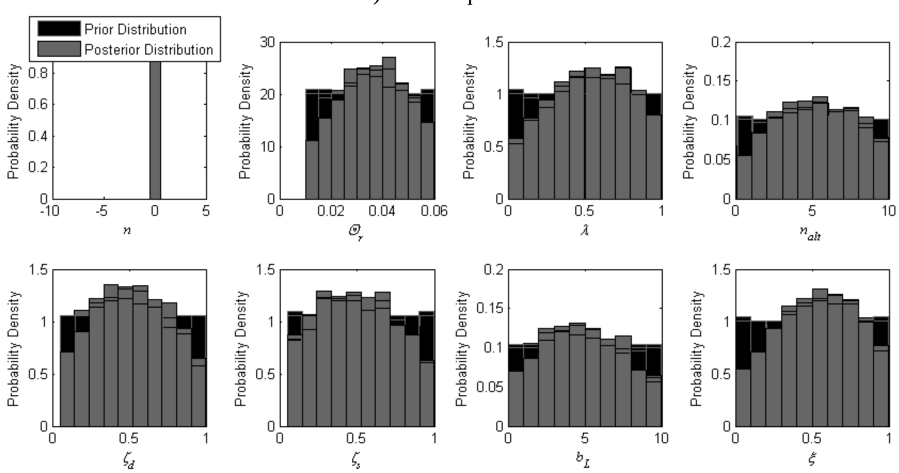
Wu, W., Wang, S.S.Y., and Jia, Y. 2000. Nonuniform sediment transport in alluvial rivers. *Journal of Hydraulic Research* **38(6)**: 427-434.



a) Parker equation



b) W&C equation



c) Wu equation

Figure 1. Prior and posterior histograms from MSU showing the likelihood that each parameter value is correct for the (a) Parker, (b) W&C, and (c) Wu equations. n is roughness coefficient, Θ_r is critical shear stress, λ is hiding factor, n_{alt} is active layer thickness multiplier, ζ_d is deposition recovery factor, ζ_s is scour recovery factor, b_L is bed load adaptation length, and ξ is the weight of bed load fractions.

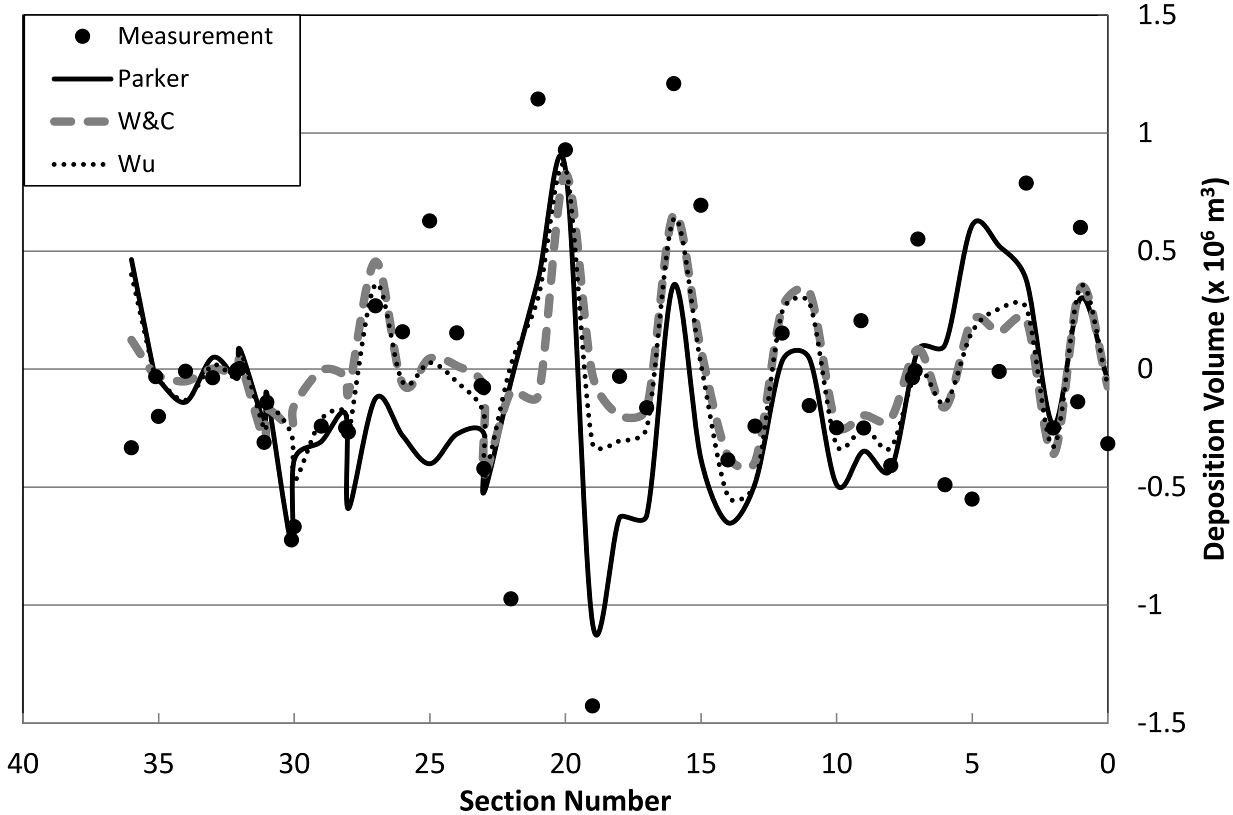


Figure 2. Comparison between measured and simulated deposition volumes downstream of Shih-Gang Dam for the Parker, W&C, and Wu equations.