Evaluation of Parameter and Model Uncertainty in Simple Applications of a 1D Sediment Transport Model

Shaina M. Sabatine[1], Jeffrey D. Niemann, M.ASCE[2 *], and Blair P. Greimann[3]

[1]Junior River Engineer & Modeling Specialist, WATERSHED Science & Engineering, 110 Prefontaine Pl. S., Suite 508, Seattle, WA 98104

[2]Associate Professor, Department of Civil & Environmental Engineering, Colorado State University, Campus Delivery 1372, Fort Collins, CO 80523, USA

[3]Hydraulic Engineer, Sedimentation and River Hydraulics Group, Technical Service Center, Bureau of Reclamation, Denver Federal Center, Building 67, Denver, CO 80225, USA

* Corresponding author. Tel.: +1 970 491 3517; fax: +1 970 491 7727.

E-mail address: jniemann@engr.colostate.edu (J.D. Niemann).

Abstract

This paper separately evaluates two methods from Bayesian Statistics to estimate parameter and model uncertainty in simulations from a 1D sediment transport model. The first method, Multi-Variate Shuffled Complex Evolution Metropolis – Uncertainty Analysis (MSU), is an algorithm that identifies the most likely parameter values and estimates parameter uncertainty for models with multiple outputs. The second method, Bayesian Model Averaging (BMA), determines a combined prediction based on three sediment transport equations that are calibrated with MSU and evaluates the uncertainty associated with the selection of the transport equation. These tools are applied to simulations of three flume experiments. For these cases, MSU does not converge substantially faster than a previously used and simpler parameter uncertainty method, but its ability to consider correlation between parameters improves its estimate of the uncertainty. Also, the BMA results suggest that a combination of transport equations usually provides a better forecast than using an individual equation, and the selection of a single transport equation substantially increases the overall uncertainty in the model forecasts.

Author Keywords: Bayesian model averaging, parameter optimization, parameter uncertainty, model uncertainty, sediment transport uncertainty

Introduction

Sediment transport models are used widely by government agencies, engineering firms, and researchers for sediment routing and sediment capacity forecasts in channels. Uncertainty in forecasts from these models can be very large. In fact, it is typical for estimates of bed load to involve 50 – 100% uncertainty (MacDonald et al. 1991). Uncertainty in sediment transport models can arise from several sources. Such models usually offer multiple equations to estimate transport capacity, and no single formula is superior for all conditions (Huang and Greimann 2010). The selection of a single equation introduces some uncertainty as to whether the correct mathematical description is being used to represent the physical system (model uncertainty). In addition, each equation contains multiple parameters that cannot be measured in the field and thus must be calibrated, usually by adjusting their values until the model reproduces some available observations during a calibration period. Thus, there is some uncertainty about whether the true values of the parameters have been identified (parameter uncertainty). Other sources of uncertainty include the representation of the flow hydraulics, channel geometry, and the model forcing data.

Analyses of uncertainty in the field of river erosion and sedimentation have focused mainly on parameter uncertainty (Chang et al. 1993; Yeh et al. 2004; Wu and Chen 2009; Ruark et al. 2011; Shen et al. 2012), while less attention has been paid to model uncertainty. Ruark et al. (2011) developed a methodology to assess parameter uncertainty in sediment modeling. Their methodology uses a multi-objective version of generalized likelihood uncertainty estimation (GLUE) (Beven and Binley 1992; Mo and Beven 2004; Werner et al. 2005;

Pappenberger et al., 2006) to estimate this uncertainty. In the Ruark et al. (2011) method, the parameters are initially assumed to conform to uniform distributions within specified ranges. Parameter sets are then generated based on these so-called prior distributions and used in the model to simulate the calibration period. The similarity between the observed and simulated behavior is used to judge the likelihood that each generated parameter set is correct without any imposed distinction between behavioral and non-behavioral parameter sets. The calculated likelihoods are used to determine the individual posterior distributions for the parameters (i.e. the parameter distributions given the available observations). Parameter sets generated from these posterior distributions are then used to simulate the forecast period and the associated distributions of the model outputs are determined to characterize the forecast uncertainty.

The Ruark et al. (2011) method is relatively simple to implement, but it has some limitations. First, it uses a potentially inefficient sampling method when developing the posterior parameter distributions. In particular, the method runs the model with a large number of parameter sets before considering the information gained from any of the simulations. As a result, many simulations are typically performed using parameter sets that have low likelihoods of being correct (van Griensven and Meixner 2007; Blasone et al. 2008). Such inefficiency is a serious consideration for sediment transport models because each simulation can be time consuming. Second, the method develops the individual posterior parameter distributions rather than the joint posterior parameter distribution. Thus, any correlation between the parameters is neglected. In other applications, uncertainty estimates have been shown to be substantially different when correlations are considered (Vrugt et al. 2003b). Finally, the method does not consider model uncertainty, which might lead to underestimations of the overall uncertainty in the model predictions. In other applications, model uncertainty has been shown to produce more uncertainty in predictions than parameter uncertainty (Carrera and Neuman 1986; Abramowitz et al. 2006).

More sophisticated algorithms have been proposed in other fields to overcome these limitations. Markov Chain Monte Carlo (MCMC) algorithms develop a sample of parameter sets from a sought posterior distribution in a way that makes use of information from each simulation as it is performed (Vrugt et al. 2003b). Thus, they are potentially more efficient. In addition, MCMC methods include correlation between parameters (Vrugt et al. 2003b). Vrugt et al. (2003b) developed a MCMC algorithm called Shuffled Complex Evolution Metropolis – Uncertainty Analysis (SCEM-UA) that evolves a sample of parameter sets from an assumed prior distribution toward the joint posterior distribution. However, the calculation of likelihood in this algorithm limits its application to cases where a single output variable is used to judge parameter likelihood. In sediment transport modeling applications, more than one model output is typically of interest (e.g., bed elevations and grain size distributions). Vrugt et al. (2003a) proposed Multiobjective Shuffled Complex Evolution Metropolis (MOSCEM), which generalizes SCEM-UA to evolve the parameter set towards the Pareto curve that reflects the trade-offs between competing objectives (and builds on previous work by Yapo et al. 1998). More recently, van Griensven and Meixner (2007) proposed a multi-objective likelihood calculation for cases with more than one output variable. The improvement in efficiency that can be achieved by applying SCEM-UA and the importance of including parameter correlation in assessing uncertainty from sediment transport models with multiple outputs remains unknown.

Another method called Bayesian Model Averaging (BMA) has been proposed to account for model uncertainty (Hoeting et al. 1999; Raftery et al. 2005; Wöhling and Vrugt. 2008). In BMA, competing models are used to simulate the calibration period. The uncertainty associated with each model (under the assumption that it is the correct model) is modeled by a normal distribution that is centered on the model's prediction. BMA then finds the most likely variance of each normal distribution and the most likely probability that each model is correct given the available observations (Vrugt et al. 2008). The combined distribution that is produced by BMA provides an estimate of the overall uncertainty including the model uncertainty. The importance of including model uncertainty in assessing the overall uncertainty in sediment transport models also remains unknown.

The goal of this paper is to evaluate proposed uncertainty methodologies that address the main limitations of the Ruark et al. (2011) method. To calibrate parameters and examine parameter uncertainty, we implement a multi-objective adaptation of SCEM-UA, which we call Multi-Variate Shuffled Complex Evolution Metropolis – Uncertainty Analysis (MSU). Unlike MOSCEM, it combines together all objectives into a single likelihood function, which is similar to van Griensven and Meixner (2007). We aim to determine whether MSU requires substantially fewer simulations to implement than the GLUE method in Ruark et al. (2011) and whether the inclusion of correlations between parameters in MSU produces important differences in the estimates of the uncertainty. To examine model uncertainty, we separately apply BMA and evaluate whether the uncertainty in the transport equation contributes substantially to the overall uncertainty in the model predictions. The only connection between MSU and BMA is that the model outputs from the calibrated parameter values from MSU are used in BMA. These methods are coupled with the Sedimentation and River Hydraulics – One Dimension (SRH-1D) model (Huang and Greimann 2010). Within this model, three equations are used to simulate bed load: the Parker (1990) equation, the Wilcock and Crowe (2003) equation, and the modified Meyer-Peter and Müller equation (Wong and Parker 2006). The model is used to simulate three bed-load driven flume experiments. The experiments include a depositional case, a data-poor erosional case, and a data-rich erosional case. In all three cases, observations are available for bed profile elevations and sediment sizes. Two of these cases are identical to those presented in Ruark et al. (2011), which allows us to compare the results from MSU to GLUE.

Methodology

MSU

MSU aims to produce a sample of parameter sets from an initially unknown joint posterior parameter distribution. While iterating towards this sample, the method simultaneously finds the parameter set that is most likely to be the correct one, which is equivalent to calibrating the model. Aside from its use of a multi-objective likelihood function, MSU is the same as SCEM-UA, which is described and tested in detail by Vrugt et al. (2003b). Although SCEM-UA lacks detailed balance and thus may not identify the exact posterior distribution, it has been shown to work well in practice (Laloy and Vrugt 2012). This section provides an overview of the method, but readers are referred to Vrugt et al. (2003b) for mathematical details.

MSU begins by generating a relatively small sample of parameter sets from the specified joint prior distribution. The sample size s is selected by the user, and the prior distribution is a joint uniform distribution with bounds that are specified by the user. The bounds represent the plausible range for each parameter before the calibration data are considered. A uniform distribution is used because initially no set of parameter values within the range is considered more likely than any other.

After the initial parameter sets have been generated, they are sorted from most likely to least likely. The likelihood of a given parameter set is judged by the model's ability to reproduce the observed values of the model outputs during the calibration period when the parameter set is used. Because sediment transport models have more than one output variable of interest, the Global Optimization Criterion (GOC) proposed by van Griensven and Meixner (2007) is used to calculate likelihood. The likelihood of parameter set Θ being correct given the observations Y_{obs} is $p(\Theta \mid Y_{obs})$ and is related to the GOC as:

$$p(\Theta \mid Y_{obs}) \propto \exp(-GOC) \tag{1}$$

where:

$$GOC = \sum_{a=1}^{A} \frac{SSE_a N_a}{SSE_{a,\min}} \tag{2}$$

In this equation, a is an index of model output variables, A is the total number of output variables, N_a is the number of observations available for variable a, SSE_a is the sum of squared errors for the model predictions of variable a, and $SSE_{a,\min}$ is the minimum sum of squared errors of variable a among all of the currently available parameter sets. Similar to the likelihood function for SCEM-UA, Eq. (1) assumes that the residuals for each variable are independent, normally distributed, and have constant variance. However, the expression allows the residuals of different output variables to have different variances.

After calculation of the likelihoods, the parameter sets are grouped into q complexes, where q is selected by the user. If two complexes are used, for example, complex one would get the 1st, 3rd, 5th, etc. most likely parameter sets, and complex two would get the 2nd, 4th, 6th, etc. most likely parameter sets. The first (and most likely) parameter set in each complex is used as the starting point for an associated Markov Chain. The complexes are used to determine how to evolve the parameter sets, while the Markov Chains track this evolution.

Trial parameter sets are then generated for each complex and considered as replacements of existing parameter sets. To generate trial parameter sets, MSU calculates the ratio of the average likelihood of the points in the selected complex and the average likelihood of the last $m \equiv s / q$ parameter sets in the corresponding Markov Chain and compares this ratio to a specified threshold. If this ratio is less than the threshold, then a candidate parameter set is drawn from a normal distribution centered on the most recent parameter set in the Markov Chain. If this ratio is greater than the threshold, a candidate parameter set is drawn from a normal distribution centered on the mean of the currently selected complex. In MSU, as in SCEM-UA

(Vrugt et al. 2003b), the threshold is large (10^6), so new parameter sets are usually generated from normal distributions centered on the current parameter set of the Markov Chain.

The candidate parameter set is accepted if the ratio of the likelihood of this parameter set to the likelihood of the current parameter set is greater than a random number generated from a uniform distribution between 0 and 1. This criterion implies that the generated parameter set is always accepted if its likelihood is larger than the current parameter set, and it is still accepted on random occasions if its likelihood is smaller. If the new parameter set is accepted, it becomes the current position of the Markov Chain and replaces the best complex member. Otherwise, the Markov Chain does not advance, but the ratio of the likelihoods of the best and worst members in the active complex is calculated. If this ratio is greater than the threshold, the covariance of the active complex might be too large (Vrugt et al. 2003b). If the likelihood of candidate parameter set is greater than that of the worst point in the complex, the worst complex member is replaced with the candidate parameter set.

This updating procedure is repeated $m/5$ times for each complex. The complexes are then shuffled to share information between them. To shuffle, the parameters sets from all complexes are re-combined into a single list and sorted from most likely to least likely as described earlier. Then, they are re-organized into complexes as previously described and the updating procedure is repeated.

The MSU algorithm has converged to the posterior distribution when it is sampling from a stable 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 (1992) Scale Reduction Score (SRS). Although the SRS is widely used, it is based on normality assumptions. In fact, it is difficult to know with certainty that convergence is reached in MCMC methods (Cowles and Carlin, 1996). The SRS indicates exact convergence for each parameter when it is equal to 1, but SRS values below 1.2 are used to indicate adequate convergence (Vrugt et al. 2003b; Gelman and Rubin 1992). Due to the difficulty in judging convergence, trace plots of the sequentially-generated parameter values are also inspected to confirm stability in the distributions.

The parameter sets that are generated after convergence are consistent with the posterior distribution. Each parameter's marginal posterior distribution can be inferred by creating histograms of the parameter values after convergence, and the correlation between the values of different parameters can be readily calculated from the generated parameter sets. Furthermore, the parameter sets can be used as the basis of model simulations for the forecast period. The histograms of the forecasted model outputs can then be used to judge the uncertainty in the model predictions that arises from the remaining parameter uncertainty. Finally, the most likely parameter set that is generated from a large sample from the posterior parameter distribution is considered to be the calibrated parameter set. Model results from this parameter set are then used in the BMA algorithm.

BMA

This section provides an overview of BMA; mathematical details and evaluations of the method are provided elsewhere (Raftery et al. 2005; Vrugt et al. 2008). BMA develops a prediction for an output variable and associated uncertainty bounds using a weighted average of the forecasts from a collection of models. In the present application, the models are different sediment transport equations within the SRH-1D program. The central variable in BMA is the probability that the observed value of the output variable Δ occurs given the individual model estimates $f_i,\ldots,f_I$ where i is an index of the available models and I is the total number of available models. This probability is denoted $p(\Delta \mid f_i,\ldots,f_I)$ and is calculated as:

$$p(\Delta \mid f_i,\ldots,f_I) = \sum_{i=1}^{I} w_i g_i(\Delta \mid f_i) \tag{3}$$

where $w_i,\ldots,w_I$ are the probabilities that each model is the correct one given the calibration data. These probabilities are nonnegative and add up to one, so they can be viewed as weights. The expression $g_i(\Delta \mid f_i)$ is the probability of observing Δ given model forecast f_i. It represents the uncertainty associated with model i and is assumed to be a normal distribution centered on the model's forecast with variance σ_i^2. The weight and variance associated with each model are estimated as the most likely values given the available observations. They are found by maximizing the likelihood l :

$$l(w_1,...,w_I,\sigma_i^2,...,\sigma_I^2 \mid f_i,...,f_I,\Delta) = \sum_{s,t}^{N_{st}} \log\left[\sum_{i=1}^{I} w_i g_i(\Delta_{st} \mid f_{ist})\right] \quad (4)$$

where N_{st} is the total number of observations over all s locations and t times in the calibration dataset and $g_i(\Delta_{st} \mid f_{ist})$ is model i's conditional probability for the observation given that model's forecast at location s and time t. An iterative procedure called the Expectation-Maximization (EM) algorithm is used to solve for the unknown weights and variances. This method is widely used for obtaining maximum likelihood estimates and is described in detail elsewhere (Dempster et al. 1977; Givens and Hoeting 2005).

The weights and variances obtained from the calibration period are assumed to apply to the forecast period as well (Raftery et al. 2005). Thus, the weights obtained from BMA can be applied directly to model outputs for the forecast period to obtain the BMA prediction (note that weights from one model application are not expected to apply to other applications of the model). Confidence (or credible) intervals of the BMA prediction give insight into overall uncertainty present in the model predictions. This uncertainty includes uncertainty due to the model selection (represented by the weights) as well as uncertainty associated with each model under the assumption that it is the appropriate model (represented by the normal distributions). The latter uncertainty estimate includes the uncertainty due to parameter values, which is separately determined by the MSU algorithm. When applying BMA, the only information that is used from MSU is the most likely parameter set associated with each transport equation.

Application

Sediment Transport Model

SRH-1D is a one dimensional hydraulic and sediment transport model that was developed and is widely used by the U.S. Bureau of Reclamation. The model is able to simulate channels with a variety of characteristics including fixed-width boundaries, steady flow, and non-cohesive sediment transport, which are considered in this paper. The model is described in detail in Huang and Greimann (2010). Here, we only highlight the parameters that are considered uncertain, which are parameters that a user typically must specify and therefore calibrate.

To compute flow hydraulics, SRH-1D solves the energy equation for steady, gradually varied flow using the standard step method. This approach uses Manning's equation and thus requires specification of Manning's roughness coefficient n, which is treated as an uncertain parameter.

Sediment transport computations in SRH-1D for the cases considered here consist of two major components: sediment routing and bed material mixing. Because all of the cases considered here are bed-load driven, the Exner equation is used to calculate changes in the volume of sediment on the bed. Bed load transport capacity is calculated using one of the following: the Parker (1990) equation, the Wilcock and Crowe (2003) (W&C) equation as modified by Gaeuman et al. (2009), or the modified Meyer-Peter and Müller (MPM) equation (Wong and Parker 2006). All these bed load equations can be written in the form:

$$\frac{q_{sj} g\left[\left(\rho_s / \rho\right) - 1\right]}{p_j \left(\tau_g / \rho\right)^{1.5}} = F\left(\phi_j\right) \tag{5}$$

where q_{sj} is volumetric sediment transport rate per unit width for grain size class j, g is gravitational acceleration, ρ_s the density of the sediment, ρ is the density of water, p_j is the bed material fraction by mass within the given size class j, and τ_g is grain shear stress. The function $F(\phi_j)$ is an empirical function fitted to field and/or laboratory data and differs between Parker, W&C, and MPM. The parameter ϕ_j is a measure of the shear stress relative to the reference shear stress:

$$\phi_j = \theta_j \Big/ \left(\xi_j \theta_r\right) \tag{6}$$

where θ_r is the reference Shield's number and θ_j is the Shield's parameter of the sediment size class j computed as:

$$\theta_j = \tau_g \Big/ \left\{\gamma \left[\left(\rho_s / \rho\right) - 1\right] d_j\right\} \tag{7}$$

ξ_j is the exposure/hiding factor, which accounts for the reduction in the critical shear stress for particles that are large relative to surrounding sediment particles and the increase in the critical shear stress for relatively small particles. It can be written:

$$\xi_j = \left(d_j / d_m\right)^{-\lambda} \tag{8}$$

where λ is a constant in the Parker (1990) equation, a function of the relative particle size in the W&C equation, and zero in the MPM equation. The representative diameter d_m is the median diameter in the Parker equation, the geometric mean diameter in the W&C equation, and not required in the MPM equation because $\lambda = 0$. The parameters θ_r and λ are constants in Parker but functions of the particle distribution in W&C. The following functions are used in SRH-1D for the W&C equation:

$$\theta_r = \theta_{r0} + 0.015\left[1 + \exp\left(10.1\sigma_{sg} - 14.14\right)\right]^{-1} \tag{9}$$

$$\lambda = 1 - \left(1 - \lambda_0\right)\left[1 + \exp\left(1.5 - d_i / d_m\right)\right]^{-1} \tag{10}$$

where σ_{sg} is the geometric standard deviation of the particle size distribution. The parameters θ_r and λ are treated as uncertain in the Parker equation, whereas the parameters θ_{r0} and λ_0 are treated as uncertain in the W&C equation. In the MPM equation, the reference shear stress is assumed to be fixed at 0.0495 and $\lambda = 0$ as specified in Wong and Parker (2006).

SRH-1D uses a total adaptation length L_{tot} to calculate the length over which transport capacity is reached:

$$L_{tot} = f_s L_b + \left(1 - f_s\right)\frac{Q}{\zeta W w_f} \tag{11}$$

where f_s is the fraction of suspended load as computed in Greimann et al. (2008), Q is the flow rate, W is the channel width, w_f is the sediment fall velocity, and $L_b = b_L h$ is the bed load adaptation length, where h is the hydraulic depth. b_L is a parameter to compute the bed load adaptation length, and ζ is the suspended sediment recovery factor. Different values are used for ζ for deposition (ζ_d) and scour (ζ_s). The bed load adaptation length parameter b_L and the suspended sediment recovery factors ζ_d and ζ_s are considered uncertain parameters.

Bed material mixing is modeled by dividing the bed into one active layer above several inactive layers. During deposition, the active layer shifts up and deposited material becomes part of the active layer while older material becomes part of the top inactive layer. During erosion, the active layer shifts down and material from the underlying inactive layers becomes part of the active layer. The thickness of the active layer is calculated by multiplying the geometric mean of the largest sediment size class by the active layer thickness multiplier n_{alt}. The user must also specify the weight of bed load fractions χ, which is the contribution of the bed load grain size distribution to the overall grain size distribution of the sediment that is transferred between the active layer and the topmost inactive layer. Both n_{alt} and χ are considered uncertain parameters.

Flume Experiments

Three flume experiments are considered as case studies. Following Ruark et al. (2011), a depositional experiment described by Seal et al. (1997) and an erosional experiment described by Ashida and Michiue (1971) are used. In both cases, we use the same observational data as Ruark et al. (2011) to allow direct comparisons to their results. Because observational data are very limited in the Ashida and Michiue (1971) experiment, another erosional case by Pender et al. (2001) is examined, which was not considered by Ruark et al. (2011). Table 1 provides a summary of the dimensions, initial conditions, experimental inputs, and observations available for the three experiments.

The Seal et al. (1997) experiment was designed to study sediment sorting during aggradation in three runs (named Runs 1-3). Downstream fining and armoring processes were observed in this experiment (Seal et al. 1997). An abundance of observations from the experiment are available including bed elevations taken typically at 18 locations every half hour, hour, and two hours for Runs 1, 2, and 3, respectively. Grain sizes (D_{16}, D_{50}, and D_{84}) were also determined at a variable number of locations along the flume profile during 4 or 5 time intervals during the experiments. These measurements were assumed to apply to the middle of the time intervals for the modeling exercises. Because the only difference between runs is the sediment feed rate, model parameters should remain the same between the runs. Thus, we used Run 2 (duration of 32.4 hours) as the calibration period and Run 3 (duration of 64 hours) as the forecast period.

The Ashida and Michiue (1971) experiment was designed to simulate bed degradation downstream of a dam. Bed elevation measurements are available at only three locations and six times within the 10 hour experiment. The bed material distribution is reported as fractions within specified size intervals and is available only at the beginning and end of the experiment. Because sediment size measurements were not collected an intermediate time, we cannot separate the case study into calibration and forecast periods that both contain observations. Thus, hours 10 through 20 were simulated as though the experiment continued and used as a forecast period.

The Pender et al. (2001) experiment was designed to study changes in bed structure and elevation during degradation in three runs (named Experiments 1-3). Experiment 1 was selected for use. This experiment has bed elevation measurements available every 2 to 3 hours at hundreds of locations for most times up to 84.6 hours. For computational purposes, we reduced the number of observed locations to between 21 and 42 points at each time, depending on the availability of observations at a given time. The bed material distribution is characterized by the fractions of sediment within specified size intervals. Hours 0 to 34.1 are used as the calibration period, and hours 34.1 to 84.6 are used as the forecast period.

Method Coupling

Table 2 shows the minimum and maximum allowed values for the eight uncertain parameters described in the previous section. These values were selected to provide broad plausible ranges for the prior joint uniform distribution provided to MSU. The range for the active layer thickness multiplier n_{alt} varies between the experiments, and in the Seal et al. (1997) experiment, the value of n_{alt} varies between sediment transport equations. The ranges for this parameter were kept as small as possible for computational purposes, but they were widened in cases where the full posterior distribution was not captured by the initial range.

Several method parameters also need to be defined to apply MSU. In all cases, an initial population size of $s = 500$ parameter sets is organized into $q = 2$ complexes. These values imply that each complex contains 250 parameter sets and each complex is updated 50 times before shuffling. The values were selected because they favor quick convergence. MSU was run for a total of 20,000 iterations to be certain that all parameters converged and large samples from the posterior parameter distributions were attained.

MSU also requires organization of the observations from the calibration periods into different variables which are allowed to have different variances for their residuals. It is assumed that bed elevations at all locations in a given flume at a given time have the same variance of their residuals and can therefore be treated as a single output variable. Aggregating observations from several locations together in this way allows for more reliable estimates of the variances of the residuals in the method. The general shape of the bed profile stays the same throughout each experiment, and the scale of the measurements at all locations at a given time does not vary greatly, so this assumption is expected to be reasonable. Bed elevations at different times are treated as different variables. If the bed aggrades or degrades substantially during the experiment, the scale of these measurements can change with time, which would likely imply a change in the variance of the residuals as well. Similar to the bed profile elevations, sediment size data from all locations are assumed to have the same variances for their residuals, while different times are treated as different variables. When D_{16}, D_{50}, and D_{84} observations are available (the Seal et al. (1997) experiment), they are treated as three separate variables. When the fraction of sediment in different size intervals is available, each size class is treated as a separate variable. Recall that the likelihood function used in MSU (Eq. (2)) assumes that the residuals for each variable are normally distributed and independent. Preliminary investigations suggested that the assumption of normality does not hold for all variables in these experiments. Transformations were used to produce normally distributed variables, but the use of the transformed variables did not substantially alter the results of MSU or BMA. Thus, the untransformed variables are used for simplicity. The assumption of independence was not evaluated in detail and remains an important assumption of this analysis.

Some differences are required in the application of MSU and BMA because BMA is not easily generalized to account for multiple variables at the same time. Thus, BMA is run twice for each case study examined: once for all bed profile elevation output and once for all sediment size output resulting in two sets of model weights for each experiment. This procedure essentially treats every bed profile elevation point as an observation from the same variable. Likewise, it treats every sediment size point as an observation from the same variable. BMA has been conducted in this manner with meteorological and hydrologic data in previous papers (Raftery et al. 2005, Vrugt et al. 2008).

Results

MSU Results

Among the key outputs of MSU are the most likely values for the uncertain parameters and associated results for the calibration period. Figs. 1 and 2 show the bed profiles and sediment size distributions that are simulated by the most likely parameter sets, respectively, along with the available observations for the calibration period. For the Seal et al. (1997) case, Parker and W&C simulate both the bed profile and the sediment sizes relatively well, while MPM is less successful. The Seal et al. (1997) case has grain sizes ranging between 0.2 mm and

65 mm. MPM is likely less successful in this case because this equation was developed assuming a single grain size class and does not represent interactions between grain size classes. The other equations were developed by considering interactions between various sediment size classes and specifically include hiding and exposure effects. For the Ashida and Michiue (1971) case, Parker and MPM reproduce the observations well, while W&C is the least successful equation for both the bed profile and sediment size distribution. W&C can simulate the bed profile well or the grain size distribution well, but no single parameter set can reproduce both types of observations simultaneously. For the Pender et al. (2001) case, the MPM equation matches the bed profile best. The median grain size and smaller sizes are predicted relatively well by the Parker equation, but all the equations fail to capture the sizes of sediment larger than the D_{50}. This disagreement might be due to the highly structured, well-sorted, and graded beds in the Pender et al. (2001) experiments. Models like SRH-1D that use an active layer to simulate sediment flows are not expected to be as accurate at estimating sediment movement in channels with such complicated bed structures.

The other key result from MSU is the parameter uncertainty that remains after calibration. Table 3 shows the percentage reduction in the Interquartile Ranges (IQRs) of the parameters generated from the prior and posterior distributions. The IQR is defined as the difference between the 75% and 25% quantiles. These percentages describe the decrease in parameter uncertainty due to calibration, so a value of 100% indicates that the algorithm converges to a single value of the parameter. As expected, the parameters are less well constrained for the data-poor Ashida and Michiue (1971) case. In fact, only the parameters with the strongest impact on the model results (n, θ_r, λ, and n_{alt}) are reasonably constrained by the available observations. In all cases, n, θ_r, and λ are among the most constrained parameters. For the Seal et al. (1997) case, the parameters are best constrained for the MPM model. As shown earlier, this model is not able to reproduce the observations well for this case, so very few combinations of parameter values are able to approach the observed system behavior. For the Ashida and Michiue (1971) case, the W&C model was shown to perform poorly, but its parameters are not well constrained because certain parameter sets are able to reproduce the bed profile or sediment sizes, but not both.

The first objective of this paper is to determine whether MSU provides a large reduction in the required number of simulations compared to the GLUE method used by Ruark et al. (2011). Fig. 3 plots the SRS for the uncertain parameters in the nine MSU runs (three flume experiments each simulated with three different transport equations). The horizontal lines show SRS = 1.2, and the arrows indicate the approximate iteration where convergence is achieved (where the SRS remains below 1.2 and trace plots indicate generation of parameter values from stable distributions). MSU converges the fastest with the MPM equation because it has two fewer parameters than the other cases. The W&C equation converges the slowest, likely because it often has more difficulty simulating the observed data. Ruark et al. (2011) found that simulation of 5000 parameter sets produces posterior parameter distributions that have consistent quantitative results between consecutive GLUE analyses. Thus, 5000 parameter sets were sufficient to produce consistent results, but additional sampling was needed to verify that such consistency was achieved. On average, MSU requires about the same number of simulations with the Parker equation (5000), more simulations with W&C (12,000), and fewer with MPM (3000) before it has converged (and the posterior distribution has been obtained). Additional simulations would be needed to actually sample from the posterior distribution. Given the differences in the structures of the two methods, a precise comparison of their efficiencies is not possible. However, the comparison does indicate that the MSU methodology does not provide a large reduction in the required number of simulations (e.g., an order of magnitude) for these experiments. When considering hydrologic models, Blasone et al. (2008) found that the computational advantage of SCEM-UA over GLUE increases when the posterior parameter distributions are narrower. For these experiments, many SRH-1D parameters remain poorly constrained after calibration (Table 3), so these results are generally consistent with Blasone et al. (2008).

The second objective of this paper is to evaluate the importance of accounting for correlation when assessing the impacts of parameter uncertainty. To assess the strength of correlation between the values of different parameters in the estimated posterior distribution, the probability that the correlation observed between a pair of parameters has occurred by chance when the true correlation is zero was calculated using the t test at a confidence level of 95%, which also assumes normality of the distributions. This analysis was done for all pairs of parameters using up to 5000 parameter sets (where available) after convergence for all nine MSU runs. 87% of the parameter pairs have a significant correlation. Of the 87%, 73% of parameter pairs have a correlation coefficient stronger than ±0.1, and 19% of parameter pairs have a correlation coefficient stronger than ±0.4. The parameter pairs that are correlated and the value of this correlation both vary between cases and equations. However, more parameter pairs have stronger correlations in the more complex equations (Parker and W&C). This result is expected because the additional parameters (θ_r and λ) both refine the description of the transport process rather than describing an additional process. The appropriate value for one parameter in the transport equation is expected to depend on the value that is used for the other parameters in that equation.

The implications of ignoring these correlations when assessing the uncertainty of model forecasts was explored by running two types of simulations with the parameter sets from the estimated posterior distributions. First, the parameter sets obtained after convergence of MSU were used to simulate the forecast periods for each of the nine cases. Second, the values for each parameter after convergence of MSU were randomly reordered to remove any correlation between different parameters while maintaining the marginal distributions estimated by MSU. The reordered parameters were also used to simulate the forecast periods. In both cases, the forecasts were characterized by defining two variables. The first variable is the average bed elevation at three selected locations (near the upstream end, midpoint, and downstream end of the flumes) and at three selected times (near the beginning, middle, and end of the forecast periods). The second variable considers the sediment sizes. For sediment size profiles, the data were averaged in the same manner as the bed elevations. For sediment size fractions, three class sizes (small, medium, and large) at three times (near the beginning, middle, and end of the forecast periods) were obtained and averaged. Then, the average IQRs for these two variables were calculated from the correlated and uncorrelated parameter sets in each case.   These IQRs are one measure of uncertainty in the forecasts. Fig. 4 shows the average IQRs for bed profile and sediment grain size for all three experiments when the parameter correlations are included and neglected. In general, removing parameter correlations has little effect on IQRs generated from the MPM equation for all three cases. This result likely occurs because the correlations in the parameters are generally smaller for MPM than the other equations as described earlier. For the Parker and W&C equations, inclusion of parameter correlation is more important when estimating the uncertainty of the bed profile elevation than sediment grain sizes for the depositional case (i.e. Seal et al. (1997)). The insensitivity of the grain sizes to the parameter correlation probably occurs because the characteristics of the deposited sediment mostly depend on the sediment that is fed to the system. This result reverses for the erosional case. In particular, parameter correlation is more important when estimating the uncertainty of the sediment grain sizes than bed profile elevations. In this case, the composition of the bed depends more directly on the erosion model, so the correlations between the parameters in this model are expected to play a larger role. Overall, these results suggest that parameter correlations should be included when assessing uncertainty in sediment transport model forecasts.

BMA Results

BMA was used to determine weights for the three transport equations based on their ability to reproduce the observations for the calibration period, and these weights are reported in Table 4. A separate set of model parameters was calibrated for each of the transport equations. For any selected experiment, BMA suggests a different set of equations for predicting bed profile elevation than it suggests for predicting sediment grain sizes. When bed profile elevation observations are used, the Parker equation dominates in the depositional (i.e. Seal et al. (1997)) case with a weight of 0.84. The probability is high that the Parker equation is the correct model because it fits the observations so well (see results in Fig. 1, for example). The W&C equation also matches the observations relatively well and has a weight of 0.16. The MPM equation dominates the bed profile elevation in both erosional cases (Ashida and Michiue (1971) and Pender et al. (2001)) with weights of 1.00 and 0.98, respectively. It is most successful in simulating the bed profile during the entire calibration period for both erosional cases (even though the Parker equation performs better for the particular time step shown in Fig. 1 for Ashida and Michiue (1971)). Fig. 5 examines whether the weightings identified in the calibration periods also apply to the forecast periods. In particular, it shows the individual model forests and the BMA forecasts for the Seal et al. (1997) and Pender et al. (2001) cases. The Ashida and Michiue (1971) case is not shown because observations are not available for the forecast period. The forecasts produced by the BMA weightings of the transport equations match the observations better than the individual models do, which suggests that the weights still have value for the forecast period.

When sediment grain size outputs are analyzed with BMA, the BMA results are rather different. As Table 4 shows, BMA suggests a different combination of equations for each flume experiment. For the Seal et al. (1997) depositional case, a combination of all three equations is suggested by BMA. For the Ashida and Michiue (1971) erosional case, BMA suggests a combination of the Parker and MPM equations, while for the Pender et al. (2001) erosional case, BMA suggests a combination of the Parker and W&C equations. Overall, more balanced weightings are observed for the sediment grain sizes than for the bed profile. Such balance suggests that the different transport equations have distinctive individual abilities in reproducing the observed grain sizes. Fig. 6 compares the individual model and BMA predictions to the observations for the forecast period for the sediment grain size outputs. For the Seal et al. (1997) case, BMA provides a prediction that is a compromise of the three models' performances in simulating the D_{16}, D_{50}, and D_{84} profiles. For the Pender et al. (2001) case, however, the W&C equation actually outperforms BMA for the time-step shown (although none of the equations performs particularly well and the uncertainty bounds are quite large). This behavior likely occurs because the SRH-1D model is not able to accurately predict the entire sediment size distribution during the calibration period, so weighting determined from these forecasts is unreliable as well. There may be multiple reasons why there is disagreement between the simulated and measured results such as the bed mixing algorithms, the unsteady nature of bed load motion, and deficiencies in the transport equations.

Another key objective of this paper is to determine how important the uncertainty in the form of the transport equation is relative to the parameter uncertainty. The uncertainty bounds produced by BMA include both parameter and model uncertainty. To estimate the amount of uncertainty attributable to the selection of a sediment transport equation, the average IQRs of the output histograms generated from MSU (shown in Fig. 4), which consider only parameter uncertainty, are compared to the average IQRs of the respective BMA distributions. To calculate the IQRs for BMA, the same times, locations, and variables were used as in Fig. 4. It should be noted that MSU and BMA are based on different statistical models for uncertainty as explained earlier, so this comparison is inexact. Fig. 7 compares the IQRs for the bed profile elevation and the sediment size outputs for the three flume experiments. In all cases, the IQR values from the BMA predictions are greater than the IQR values of the equations that are used in the BMA prediction. Examining the bed profile elevation data for the Ashida and Michiue (1971) case (Fig. 7(c)), the IQR for the Parker equation is larger than the BMA IQR, but the Parker equation is not used in the BMA estimate. Examining the sediment grain size data for the Pender et al. (2001) case (Fig. 7(f)), the IQR for BMA is much larger than the IQR for any individual model. The individual models have low uncertainty because the values of their parameters are constrained relatively well (Table 3). However, even the most likely parameter values for each equation do not produce good performance (Fig. 2(e)), which ultimately produces a large IQR for BMA. Overall, the results from Fig. 7 suggest that model uncertainty, including the selection of the transport equation, may contribute significantly to the overall uncertainty in the model forecasts.

Conclusions

(1) Even though MSU uses a more sophisticated approach to develop parameter posterior probability distributions, it does not provide a large improvement in the required number of simulations compared to the GLUE method used in Ruark et al. (2011) for the cases studied here. The GLUE method generates a large sample (5000 parameter sets) from a joint uniform distribution and generates the marginal posterior distributions based on likelihood

values calculated from model simulations. MSU begins with a smaller sample (500 parameter sets) generated from a joint uniform distribution and evolves the joint posterior parameter distribution based on frequent calculation of likelihoods and sharing of information between simulations. MSU converges more slowly when the number of uncertain parameters is greater and the ability of the model to reproduce the observations is weaker. As a result, MSU is not expected to have large computational advantages for actual river systems if they are more difficult to model than these flume experiments. Both MSU and GLUE are expected to be difficult to apply to complex sediment transport model applications unless high performance computing resources are available.

(2) Inclusion of parameter correlations substantially alters MSU's estimation of uncertainty in the SRH-1D forecasts for some cases. The importance differs between the depositional and erosional cases used and matters more when using transport equations with more parameters (Parker and W&C). For the depositional experiment by Seal et al. (1997), it was found that parameter correlations are more important for bed profile elevations than for sediment grain sizes. This result is reversed for the two erosional experiments (Ashida and Michiue (1971) and Pender et al. (2001)). MSU accounts for parameter correlations whereas the GLUE method used by Ruark et al. (2011) does not. Based on these results, correlations should not be overlooked in uncertainty assessments of natural river systems without careful evaluation of their roles in the specific circumstances that are being modeled.

(3) Results of BMA indicate that the equation(s) best suited for predicting one type of output (i.e. bed profile elevation) are not necessarily best suited for predicting a different type of output (i.e. sediment grain sizes). In most cases, using a weighted combination of equations from BMA produces a better forecast than using a single transport equation. Unlike MSU, BMA can be easily applied for a model of a natural river system because it requires little computation time. The appropriate weights can be determined from the calibration data and then used to produce a forecast. Additional testing is needed for natural river systems, but these results suggests that BMA may be a practical way of incorporating multiple transport equations and that this approach might lead to more reliable forecasts from sediment transport models. BMA also has limitations. Due to its statistical construction, it does not provide clear indications about the origins of the errors in the transport equations or a clear path for developing a more physically-based sediment transport theory.

(4) For all forecast periods, including model uncertainty along with parameter uncertainty substantially widens the bounds of uncertainty on the forecasts. This result suggests that the uncertainty associated with the selection of the transport equation should be considered when assessing overall uncertainty in sediment transport modeling applications. It should be noted that these are not the only sources of uncertainty that should be considered in sediment transport modeling. Uncertainties in the structure of the model used to simulate flow hydraulics (e.g., Apel et al. 2009), the channel geometry (Wong et al., 2014), bed mixing algorithms, and the model forcing variables can also contribute to the overall uncertainty. In

addition, the roles of these factors are expected to depend on the spatial and temporal scales over which the forecasts are generated (Wong et al., 2014).

Acknowledgements. The authors thank Mazdak Arabi and Jennifer A. Hoeting for their assistance in the selection and application of the MSU and BMA methods. Financial support from the U.S. Bureau of Reclamation Science and Technology Program for this research is also gratefully acknowledged. We also thank two anonymous reviewers, the associate editor, and the editor in chief for their help in improving this manuscript.

References

Abramowitz, G., Gupta, H., Pitman, A., Wang, Y., Leuning, R., Cleugh, H., and Hsu, K. L. (2006). "Neural error regression diagnosis (NERD): A tool for model bias identification and prognostic data assimilation." *Journal of Hydrometeorology*, 7, 160-177. DOI: 10.1175/JHM479.1.

Apel, H., Aronica, G.T., Kreibich, H., and Thieken, A.H. (2009). "Flood risk analysis—how detailed do we need to be?" *Natural Hazards*, 49, 79-98. DOI: 10.1007/s11069-008-9277-8.

Ashida, K., and Michiue, M. (1971). "An investigation of river bed degradation downstream of a dam." *14th International Association for Hydraulic Research Congress*. Paris, 247-256.

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

Blasone, R.-S., Vrugt, J.A., Madsen, H., Rosbjerg, D., Robinson, B.A., and Zyvoloski, G.A. (2008). “Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling.” *Advances in Water Resources,* 31, 630-648. DOI: 10.1016/j.advwatres.2007.12.003

Carrera, J., and Neuman, S. P. (1986). “Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data.” *Water Resources Research*, 22, 228-242. DOI: 10.1029/WR022i002p00228.

Chang, C. H., Yang, J. C., and Tung, Y. K. (1993). “Sensitivity and uncertainty analysis of a sediment transport model - a global approach.” *Stochastic Hydrology Hydraulics*, 7, 299-314. DOI: 10.1007/BF01581617.

Cowles, M.K., and Carlin, B.P. (1996). “Markov chain Monte Carlo convergence diagnostics: a comparative review.” *Journal of the American Statistical Association*, 91(434), 883-904.

Dempster, A. P., Laird, N., and Rubin, D. B. (1977). “Maximum likelihood from incomplete data via the EM algorithm.” *Journal of the Royal Statistical Society Series B – Methodological*, 39, 1-38.

Gaeuman, D., Andrews, E. D., Krause, A., and Smith, W. (2009). “Predicting fractional bed load transport rates: Application of the Wilcock-Crowe equations to a regulated gravel bed river.” *Water Resources Research*, 45, W06409. DOI: 10.1029/2008WR007320.

Gelman, A., and Rubin, D. B. (1992). “Inference from iterative simulation using multiple sequences.” *Statistical Science*, 7, 457-472. DOI: 10.1214/ss/1177011136.

Givens, G. H., and Hoeting, J. A. (2005). *Computational Statistics*. John Wiley & Sons, Inc.: Hoboken.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). "Bayesian model averaging: A tutorial." *Statistical Science*, 14, 382-417. DOI: 10.1214/ss/1009212519.

Huang, J. V., and Greimann, B. P. (2010). *User's manual for SRH-1D 2.6,* Denver Bureau of Reclamation, U.S. Department of the Interior, Denver.

Laloy, E., and Vrugt, J.A. (2012). "High-dimensional posterior exploration of hydrologic models using multiple-try DREAM$_{(ZS)}$ and high-performance computing." *Water Resources Research*, 48, W01526. DOI: 10.1029/2011WR010608.

MacDonald, L. H., Smart, A. W., and Wissmar, R. C. (1991). *Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska*. U.S. Environmental Protection Agency, Seattle.

Meyer-Peter, E., and Müeller, R. (1948). "Formulas for bed-load transport." *Proceedings of the International Association for Hydraulic Research*, 2^{nd} Meeting, Stockholm, Sweden.

Mo, X., and Beven, K. (2004). "Multi-objective parameter conditioning of a three-source wheat canopy model." *Agricultural and Forest Meteorology*, 122, 39-63. DOI:10.1016/j.agrformet.2003.09.009.

Pappenberger, F., Matgen, P., Beven, K.J., Henry, J.-B., Pfister, L., Fraipoint de, P. (2006). "Influence of uncertain boundary conditions and model structure on flood inundation predictions." *Advances in Water Resources*, 29, 1430-1449.

Parker, G. (1990). “Surface-based bedload transport relation for gravel rivers.” *Journal of Hydraulic Research*, 28, 417-436. DOI: 10.1080/00221689009499058.

Pender, G., Hoey, T. B., Fuller, C., and McEwan, I. K. (2001). “Selective bedload transport during the degradation of a well sorted graded sediment bed.” *Journal of Hydraulic Research*, 39, 269-277. DOI: 10.1080/00221680109499830.

Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). “Using Bayesian model averaging to calibrate forecast ensembles.” *Monthly Weather Review*, 133, 1155-1174. DOI: 10.1175/MWR2906.1.

Ruark, M. D., Niemann, J. D., Greimann, B. P., and Arabi, M. (2011). “Method for assessing impacts of parameter uncertainty in sediment transport modeling applications.” *Journal of Hydraulic Engineering*, 137, 623-636. DOI: 10.1061/(ASCE)HY.1943-7900.0000343.

Seal, R., Paola, C., Parker, G., Southard, J. B., and Wilcock, P. R. (1997). “Experiments on downstream fining of gravel: I. Narrow-channel runs.” *Journal of Hydraulic Engineering*, 123, 874-884. DOI: 10.1061/(ASCE)0733-9429(1997)123:10(874).

Shen, Z.Y., Chen, L., and Chen, T. (2012). “Analysis of parameter uncertainty in hydrologic and sediment transport modeling using GLUE method: a case study of SWAT model applied to Three Gorges Reservoir Region, China.” *Hydrology and Earth System Sciences*, 16, 121-132. DOI: 10.5194/hess-16-121-2012.

van Griensven, A., and Meixner, T. (2007). “A global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models.” *Journal of Hydroinformatics*, 9, 277-291. DOI: 10.2166/hydro.2007.104.

Vrugt, J. A., Diks, C. G. H., and Clark, M. P. (2008). “Ensemble Bayesian model averaging using Markov Chain Monte Carlo sampling.” *Environmental Fluid Mechanics*, 8, 579-595. DOI: 10.1007/s10652-008-9106-3.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S. (2003b). “A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters.” *Water Resources Research*, 39, 1201-1216. DOI: 10.1029/2002WR001642.

Werner, M.G.F., Hunter, N.M., and Bates, P.D. (2005). “Indentifiability of distributed floodplain roughness values in flood extent estimation.” *Journal of Hydrology*, 314, 139-157, DOI: 10.1016/j.jhydrol.2005.03.012.

Wilcock, P. R., and Crowe, J. C. (2003). “Surface-based transport model for mixed-size sediment.” *Journal of Hydraulic Engineering*, 129, 120-128. DOI: 10.1061/(ASCE)0733-9429(2003)129:2(120).

Wöhling, T., and Vrugt, J.A. (2008). “Combining multiobjective optimization and Bayesian model averaging to calibrate ensembles of soil hydraulic models.” *Water Resources Research*, 44, W12432. DOI: 10.1029/2008WR007154.

Wong, J.S., Freer, J.E., Bates, P.D., Sear, D.A., and Stephens, E.M. (2014). “Sensitivity of a hydraulic model to channel erosion uncertainty during extreme flooding.” *Hydrologic Processes*, published online, DOI: 10.1002/hyp.10148.

Wong, M. G., and Parker, G. (2006). “Reanalysis and correction of bed-load relation of Meyer-Peter and Müller using their own database.” *Journal of Hydraulic Engineering*, 132, 1159-1168. DOI: 10.1061/(ASCE)0733-9429(2006)132:11(1159).

Wu, F.-C., and Chen, C. C. (2009). “Bayesian updating of parameters for a sediment entrainment model via Markov chain Monte Carlo.” *Journal of Hydraulic Engineering*, 135(1), 22–37. DOI:10.1061/(ASCE)0733-9429(2009)135:1(22).

Yapo, P.O., Gupta, H.V., and Sorooshian, S. (1998). “Multi-objective global optimization for hydrologic models.” *Journal of Hydrology*, 204, 83– 97, 1998.

Yeh, K., Lin, E., and Chen, S. (2004). “Effect of uncertainty of mobile-bed model parameters on bed evolution.” In: Altinakar, M. S., Wang, S. S. Y., Holz, K. P., and Kawahara, M., editors, *Advances in Hydro-Science and Engineering, Volume VI* . University of Mississippi, Oxford.

Table 1. Summary of the initial conditions, experimental inputs, and observations for the three experiments

Experiment	Seal et al. (1997)		Ashida and Michiue (1971)		Pender et al. (2001)	
Period	Calibration	Forecast	Calibration	Forecast	Calibration	Forecast
Channel geometry	Shape: rectangular Length: 45 m Width: 0.3 m Slope: 0.2%		Shape: rectangular Length: 20 m Width: 0.8 m Slope: 1.0%		Shape: trapezoidal with 45° side slopes Length: 20 m Width 2.46 m Slope: 0.26%	
Volumetric flow rate (m^3/s)	0.049		0.314		0.117	
Time period of experiment (hr)	0 – 32.4	0 – 64	0 – 10	10 – 20	0 – 32.1	32.1 – 84.6
Sediment feed rate (kg/s)	0.09	0.05	0		0	
Bed material diameter range (mm)	0.125 – 65		0.2 – 10		0.25 – 22.63	
Number of observed size fraction intervals	9		12		13	
Median diameter (D_{50}) (mm)	5		1.5		4	
Number of bed profile elevation observations for calibration	518	-	18	-	597	-
Type of sediment size data	Sediment grain size (D_{16}, D_{50}, and D_{84}) profiles		Fractions of sediment in size intervals		Fractions of sediment in size intervals	
Number of sediment size observations for calibration	165	-	12	-	104	-

Table 2. Uniform distribution bounds for uncertain parameters

Parameter	Minimum	Maximum
Manning's roughness coefficient (n)	0.015	0.065
Critical shear stress (θ_r)	0.01	0.06
Hiding factor (λ)	0	1
Active layer thickness multiplier (n_{alt})	0.1	4[a], 6[b], 10[c], 15[d]
Deposition recovery factor (ζ_d)	0.05	1
Scour recovery factor (ζ_s)	0.05	1
Bed load adaptation length (b_L)	0	10
Weight of bed load fractions (χ)	0	1

[a]Pender et al. (2001) case
[b]Ashida and Michiue (1971) case
[c]Seal et al. (1997) case with the Parker and W&C equations
[d]Seal et al. (1997) case with the MPM equation

Table 3. The percent decrease in the Interquartile Range (IQR) of parameters generated from their prior uniform distributions and the IQR of parameters generated from MSU after convergence.

	Seal *et al.* (1997)			Ashida and Michiue (1971)			Pender *et al.* (2001)		
	Parker	W&C	MPM	Parker	W&C	MPM	Parker	W&C	MPM
n	98	98	97	41	45	81	99	100	99
θ_r	94	96	-	71	63	-	92	93	-
λ	99	97	-	44	57	-	97	92	-
n_{alt}	68	67	95	52	78	87	92	87	84
ζ_d	57	58	63	14	18	13	37	40	20
ζ_s	55	63	100	19	27	44	91	100	52
b_L	87	82	93	15	15	12	60	94	22
χ	81	83	91	19	15	11	68	93	16

Table 4. BMA weights for the three equations and two model outputs for calibration periods of the three experiments

Experiment	Parker	W&C	MPM
	Bed Profile Elevation BMA Weights		
Seal et al. (1997)	0.84	0.16	0.00
Ashida and Michiue (1971)	0.00	0.00	1.00
Pender et al. (2001)	0.02	0.00	0.98
	Sediment Grain Size BMA Weights		
Seal et al. (1997)	0.27	0.54	0.19
Ashida and Michiue (1971)	0.42	0.00	0.58
Pender et al. (2001)	0.45	0.55	0.00

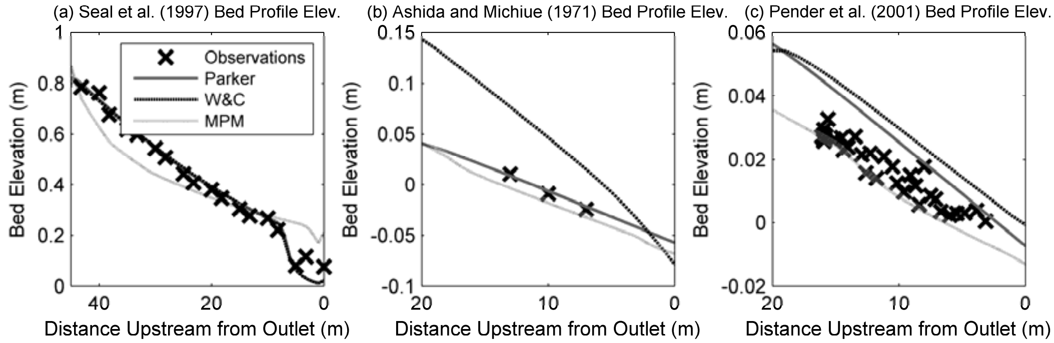

Fig. 1. Model results and corresponding observations of bed profile elevation for the calibration period of (a) the Seal et al. (1997) case at 32.4 hours, (b) the Ashida and Michiue (1971) case at 10 hours, and (c) the Pender et al. (2001) case at 32.1 hours

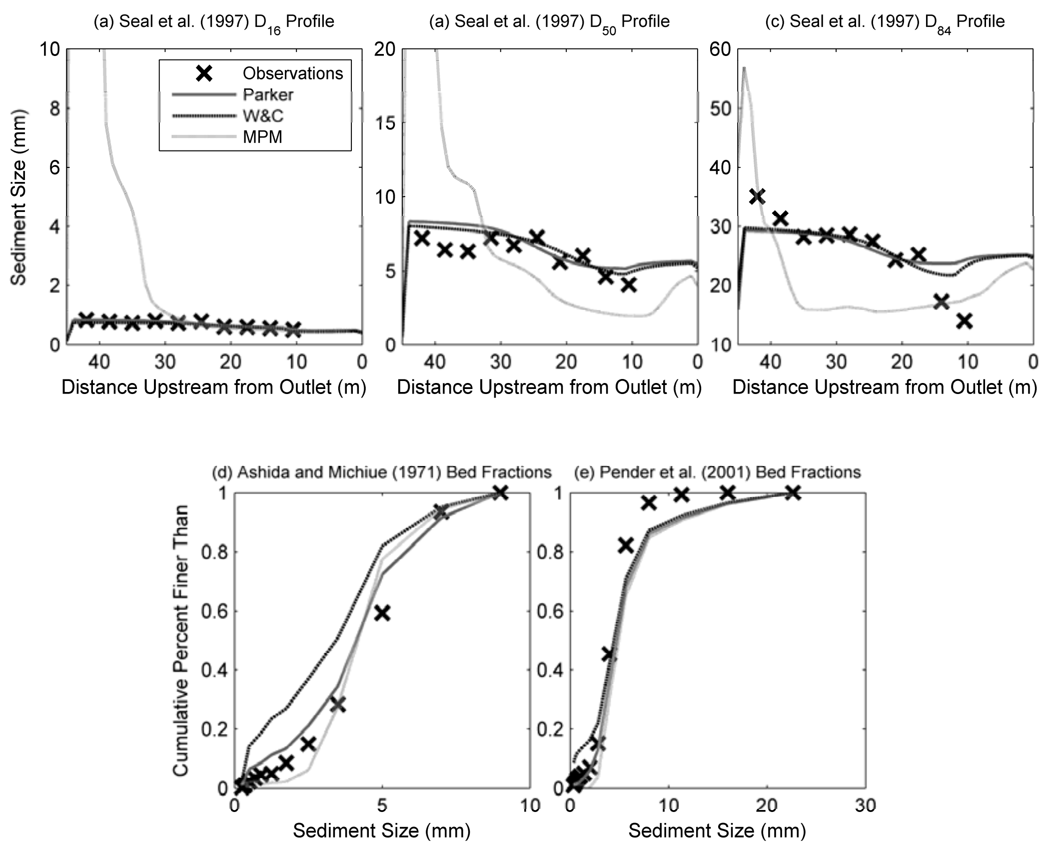

Fig. 2. Model results and corresponding observations of sediment sizes for the calibration period of (a-c) the Seal et al. (1997) case at 27 hours, (d) the Ashida and Michiue (1971) case at 10 hours, and (e) the Pender et al. (2001) case at 32.1 hours. For the seal case, profiles of D_{16}, D_{50}, and D_{84} are shown. For the other cases, the cumulative distributions at the observation locations are shown.

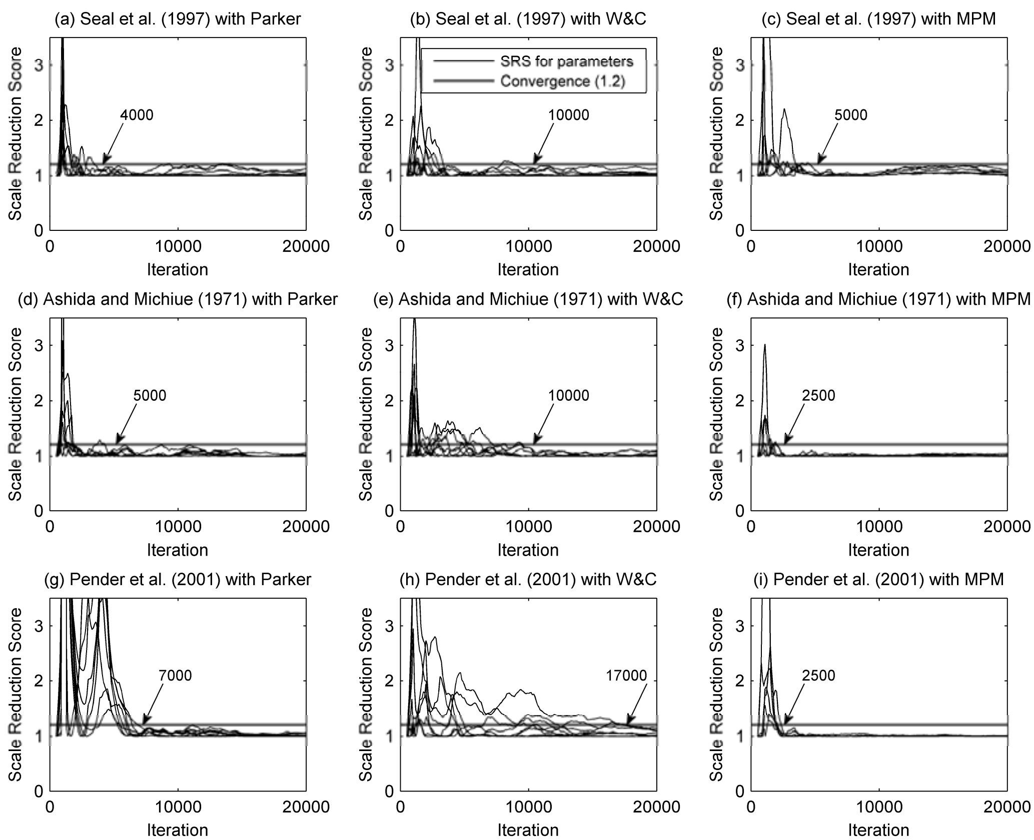

Fig. 3. Gelman and Rubin's convergence diagnostic: the Scale Reduction Score (SRS) for the 20,000 iterations of MSU. The SRS is shown for all uncertain parameters, experiments, and equations. Arrows indicate the point of convergence in each plot.

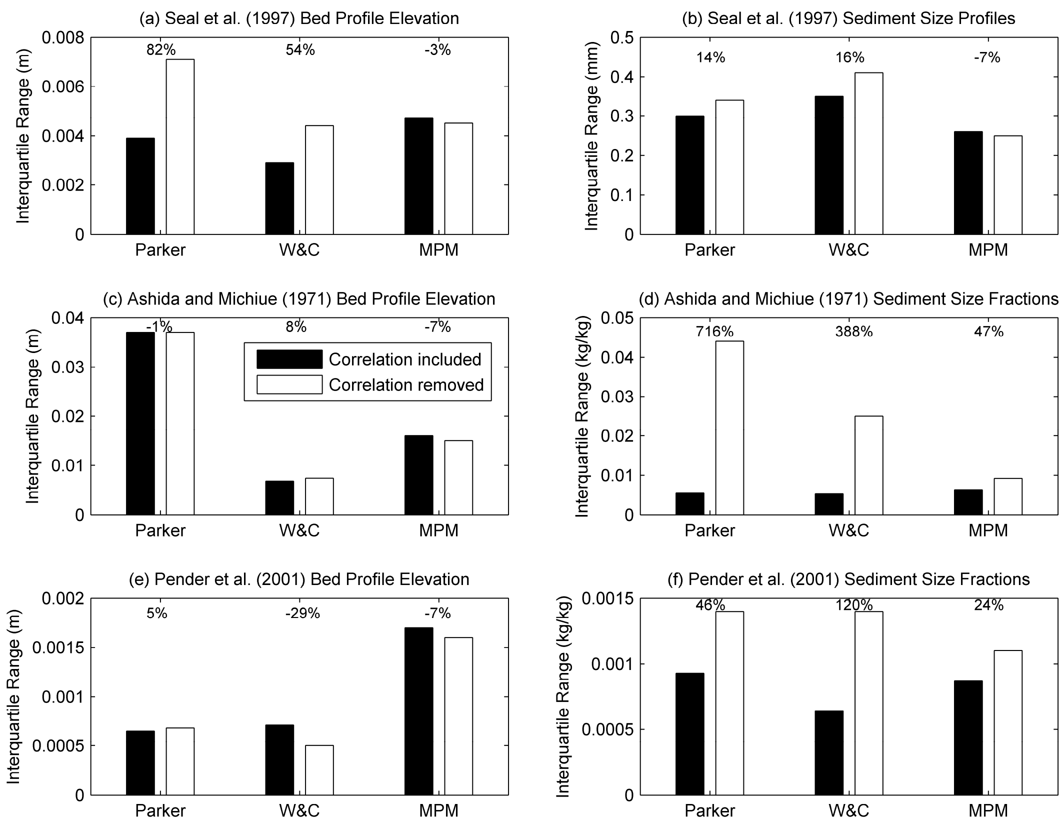

Fig. 4. The average over time and space of the Interquartile Range (IQR) for the bed profile elevation or sediment size outputs (as labeled) from the forecast periods of the three experiments when simulated with the three transport equations. The black bars describe model outputs using parameter sets generated from MSU after convergence and the white bars correspond to model outputs generated using these same parameters sets after they have been shuffled to remove correlation. The percentages indicate the change in the IQR in each case when correlation is removed.

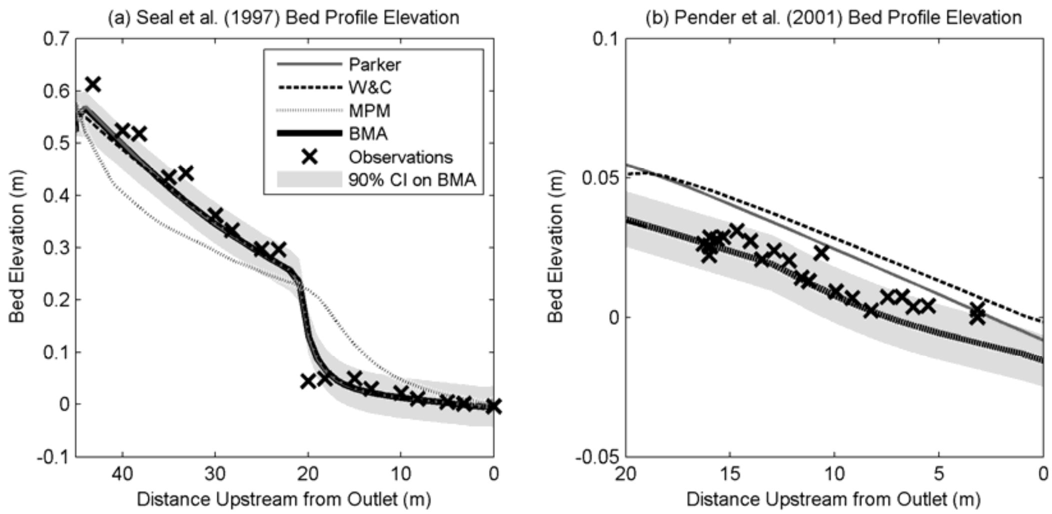

Fig. 5. BMA predictions, individual model responses, corresponding observations, and the 90% Confidence or Credible Interval (CI) on the BMA prediction of bed profile elevation for the forecast period of (a) the Seal et al. (1997) case showing bed profile elevation at 32 hours, and (b) the Pender et al. (2001) case showing bed profile elevation at 62.4 hours.

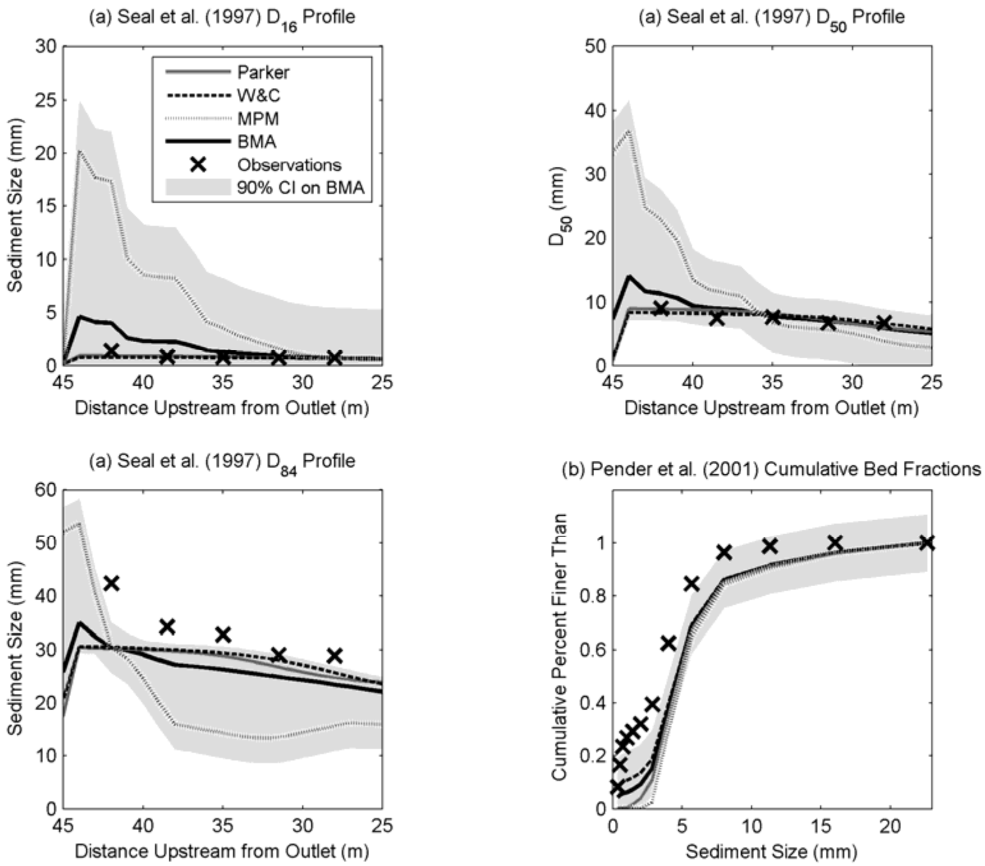

Fig. 6. BMA predictions, individual model responses, corresponding observations, and the 90% Confidence or Credible Interval (CI) on the BMA prediction of sediment size data for the forecast period of (a) the Seal et al. (1997) case showing the D_{16} profile at 34 hours, (b) the Seal et al. (1997) case showing the D_{50} profile at 34 hours, (c) the Seal et al. (1997) case showing the D_{84} profile at 34 hours, and (d) the Pender et al. (2001) case showing cumulative sediment size fractions at 62.3 hours

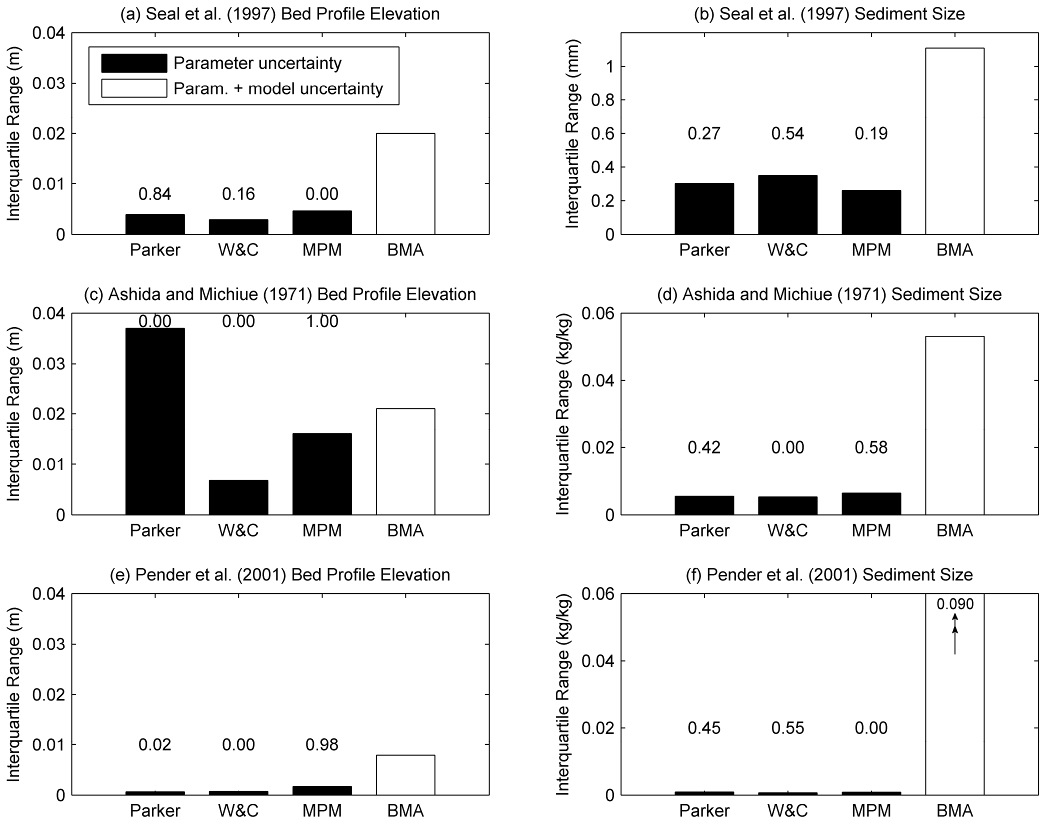

Fig. 7. The average over time and space of the Interquartile Range (IQR) for the bed profile elevation or sediment size outputs (as labeled) from the forecast periods of the three experiments when simulated with the three transport equations. The black bars are associated with individual models and represent approximate parameter uncertainty. The white bars are associated with the BMA prediction and represent both parameter and model uncertainty. For reference, the weights applied to each equation to create the BMA forecast are reported above each black bar. Note that the size of the white bar is given in (f) because it is much larger than the other bars shown.