

THESIS

EVALUATION OF PARAMETER AND MODEL UNCERTAINTY IN SIMPLE  
APPLICATIONS OF A 1D SEDIMENT TRANSPORT MODEL

Submitted by

Shaina M. Sabatine

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2011

Master's Committee:

Advisor: Jeffrey D. Niemann

Blair Greimann

Jennifer Hoeting

## ABSTRACT

### EVALUATION OF PARAMETER AND MODEL UNCERTAINTY IN SIMPLE APPLICATIONS OF A 1D SEDIMENT TRANSPORT MODEL

This paper aims to quantify parameter and model uncertainty in simulations from a 1D sediment transport model using two methods from Bayesian statistics. The first method, Multi-Variable 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 other method, Bayesian Model Averaging (BMA), determines a combined prediction based on three sediment transport equations and evaluates the uncertainty associated with the selection of a transport equation. These tools are applied to simulations of three flume experiments. Results show that MSU’s ability to consider correlation between parameters improves its estimate of the uncertainty in the model forecasts. Also, BMA results suggest that a combination of transport equations usually provides a better forecast than an individual equation, and the selection of a single transport equation substantially increases the overall uncertainty in the model forecasts.

## ACKNOWLEDGMENTS

The financial contribution from the U.S. Bureau of reclamation for this research is gratefully acknowledged along with the contributions from committee and non-committee members. Throughout the project, Dr. Jeff Niemann provided guidance and helped with understanding and integrating challenging concepts. Dr. Blair Greimann assisted with sediment transport modeling using SRH-1D. Dr. Jennifer Hoeting offered assistance with questions about complicated statistical procedures. Dr. Mazdak Arabi helped in deciding which methodologies to use.

## TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS .....	iii
1 INTRODUCTION .....	1
2 METHODOLOGY .....	6
2.1 <i>MSU</i> .....	6
2.2 <i>BMA</i> .....	11
3 SRH-1D.....	15
4 EXPERIMENTS .....	18
5 RESULTS .....	23
5.1 <i>MSU Results</i> .....	23
5.2 <i>BMA Results</i> .....	28
6 CONCLUSIONS.....	31
TABLES AND FIGURES .....	34
REFERENCES .....	49
APPENDIX A.....	52

## 1 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 [1]. Uncertainty in dynamic models can arise from several sources including approximations or simplifications in the mathematical structure of the model (model uncertainty) and ambiguity in values selected for the model parameters (parameter uncertainty) [2]. For sediment transport models, no single formula can accurately describe sediment flows for all possible fluvial conditions [3]. Therefore, a transport capacity equation among the many available is typically selected for use based on stream type and conditions, ease of application, and calibration data. The selection of a single equation introduces some uncertainty as to whether the correct mathematical description is being used to model the physical system. In addition, each equation contains multiple parameters that cannot be measured in the field and thus must be calibrated in some manner, usually by adjusting their values until the model reproduces the system behavior for some calibration period. All of these uncertainties can lead to errors in forecasts that might be used when designing channels.

Analyses of uncertainty in the field of river erosion and sedimentation have focused on parameter uncertainty [4,5,6], while less attention has been paid to the uncertainty due to the selection of a sediment transport formula. Ruark *et al.* [6]

developed a methodology to assess parameter uncertainty in sediment modeling. Their methodology uses a multi-objective version of generalized likelihood uncertainty estimation (GLUE) [7] to estimate this uncertainty. In their method, the parameters are initially assumed to conform to uniform distributions within specified ranges. Parameter sets are then generated based on the uniform distributions and used in the model to simulate a calibration period. The similarity between the observed and simulated behavior is used to judge the likelihood that each generated parameter set is correct. The calculated likelihoods are then 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 [6]. The method is relatively simple to implement, but it has some limitations. First, it uses an inefficient sampling method when developing the posterior parameter distributions. In particular, the method generates a large number of parameter sets from the specified uniform distributions and uses all of those parameter sets in model simulations before considering the results. No information gained from one of these simulations is used to decide how to configure the next simulation. As a result, many simulations are conducted in regions of the parameter space with low likelihood. For other modeling applications, methods that evolve the joint posterior parameter distribution using information (likelihood values) acquired from each simulation have been shown to decrease computation time [8]. Secondly, the method by Ruark *et al.* [6] does not consider possible correlations between parameters. Specifically, it develops the marginal posterior parameter distributions rather than the joint posterior parameter

distribution. Results in Ruark *et al.* [6] indicate that these correlations might be non-trivial, and in other modeling applications, uncertainty estimates have been shown to be substantially different when correlations are considered [8]. Finally, the method does not consider sources of uncertainty other than parameter uncertainty, which might lead to underestimations of the overall uncertainty in the model predictions. In other model applications, structural errors are known to produce more uncertainty in predictions than errors in model parameters or inputs [9,10].

Markov Chain Monte Carlo (MCMC) algorithms, tools used in Bayesian inference, have the ability to efficiently develop a sample from a sought posterior parameter distribution while including correlation between parameters [8]. Vrugt *et al.* [8] developed a MCMC algorithm called the Shuffled Complex Evolution Metropolis – Uncertainty Analysis (SCEM-UA) that evolves a sample of parameter sets generated from the assumed prior joint uniform distribution toward the joint posterior probability distribution. The calculation of likelihood in this algorithm limits it to cases where a single model output variable is used to calibrate the parameters. In sediment transport modeling applications, however, more than one model output is typically of interest (e.g., bed elevations and grain size distributions). Van Griensven and Meixner [11] proposed a likelihood function derived from Bayesian statistics for cases with more than one output variable. The improvements 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.

Bayesian Model Averaging (BMA) has been proposed as a method to account for model uncertainty [12,13]. In BMA, simulations from competing models are evaluated

based on their abilities to reproduce the calibration data for an output variable of interest. The uncertainty associated with each model (assuming it is the correct model) is modeled by a normal distribution that is centered on the model's prediction. BMA estimates the variance of each normal distribution as well as the probability that each model is correct to maximize the likelihood of these values given the calibration observations [14]. The combined distribution that is produced by BMA provides an estimate of the overall uncertainty including the uncertainty introduced by the different models under consideration. 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 a proposed uncertainty methodology that addresses the limitations of the Ruark *et al.* [6] method. To examine parameter uncertainty, we implement a multi-objective adaptation of SCEM-UA [8] following the work done by van Griensven and Meixner [11] entitled Multi-Variate Shuffled Complex Evolution Metropolis – Uncertainty Analysis (MSU). To examine model uncertainty and whether it is beneficial to combine results from multiple transport equations, we apply BMA in a similar manner to [13] and [14]. MSU is first implemented on the calibration periods of three bed load driven flume experiments that are modeled using the Sedimentation and River Hydraulics – One Dimension (SRH-1D) model [3]. These experiments include a depositional case, a data-poor erosional case, and a data-rich erosional case. Two types of observations are available for the calibration of all cases: bed profile elevation data and sediment grain size data. The impact of model uncertainty on the uncertainty of two outputs, bed profile elevations and sediment grain sizes, is then assessed for each case study using BMA. Three equations are available in SRH-1D to

treat bed load and are used in BMA: the Parker equation [15], the Wilcock and Crowe equation [16], and the modified Meyer-Peter and Müller equation [17]. The analyses included in this paper specifically focus on: (1) whether MSU's sampling method is more efficient than GLUE, (2) the relative ability of MSU to constrain parameter values for different flume experiments and transport equations, (3) whether the multi-objective likelihood function identifies a parameter set that completely optimizes both bed profile elevation and sediment grain size outputs at the same time (4) the importance of including the correlation between parameters in estimating the forecast uncertainty, and (5) whether the choice of a transport equation adds substantially to the uncertainty in the model forecast.

In the next section, the uncertainty methodology is described in detail. Following that, the model used for testing the methodology (SRH-1D) is described. Next, the application of the proposed methodology to the three case studies is summarized. In the fifth section, results are given and explained. Finally, conclusions are summarized in the last section.

## 2 METHODOLOGY

### 2.1 MSU

MSU aims to produce a sample of parameter sets from an initially unknown joint posterior parameter distribution. While iterating towards such a sample, the method simultaneously finds the parameter set that is most likely to be the correct one. This is accomplished by first generating a relatively small sample of parameter sets from the specified joint prior distribution. These parameter sets are then sorted into complexes, and one parameter set in each complex is used as the first point in a Markov chain. Then, trial parameter sets are generated based on the current point in the Markov chain and a so-called proposal distribution or transition kernel, which is determined from information in the complexes. The trial parameter set is retained using criteria based on the likelihoods of the trial parameter set and the current parameter set. Once convergence is reached, the sequence of parameter sets generated from the method conforms to the posterior distribution. Aside from its use of a multi-objective likelihood function (see below), MSU is the same as SCEM-UA, which is described and tested in detail by Vrugt et al. [8]. Readers are referred to that paper for a more in-depth discussion, but some details of the algorithm are given below.

MSU begins by generating an initial sample of parameter sets from the prior parameter 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 any calibration data is considered. A uniform distribution is used because no set of parameter values within the range is considered more likely than any other before the available observations are used. Rather than generating the initial sample randomly from the uniform distribution, Latin Hypercube Sampling (LHS) is used. LHS divides the parameter space into even increments and generates parameter sets at every point on the multi-dimensional grid formed by these increments [18].

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 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 [11] is used to calculate likelihood. The likelihood of parameter set  $\Theta$  being correct given the observations  $Y_{obs}$  is  $p(\Theta | Y_{obs})$  and is related to the GOC as:

$$p(\Theta | Y_{obs}) \propto \exp(-GOC) \quad (1)$$

where:

$$GOC = \sum_{a=1}^A \frac{SSE_a \times N_a}{SSE_{a,\min}} \quad (2)$$

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 available observations of 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, Equation (1) assumes that the residuals for each variable are independent, normally distributed, and have constant variance [11]. However, the expression allows the residuals of different output variables to have different variances. Specifically, the expression adjusts the performance for each output variable by the associated variance, where the variance of the residuals of variable  $a$  is estimated as [11]:

$$\sigma_a^2 = \frac{SSE_{a,\min}}{N_a} \quad (3)$$

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

One complex at a time, called the active complex, is then evolved using the Sequence Evolution Metropolis (SEM) algorithm. To determine the next parameter set in the Markov Chain  $\Theta^{(b+1)}$  (where  $b$  is an index for the method iterations), SEM compares the ratio of the average of the likelihoods of the points in the active complex to the average of the likelihoods of the last  $m$  parameter sets in the corresponding Markov Chain to a specified threshold. The variable  $m = s/q$  and is the number of parameter sets in a complex. If this ratio is less than the threshold, then a candidate parameter set is

drawn from a normal distribution centered on the current parameter set of the Markov Chain:

$$\Theta^{b+1} \sim N(\Theta^b, c_n^2 \Sigma^k) \quad (4)$$

where  $N$  denotes the normal distribution,  $\Theta^{(b)}$  is the current parameter set,  $c_n$  is a scaling parameter equal to 2.4 divided by the number of uncertain parameters [8, 19], and  $\Sigma^k$  is the covariance of the parameter sets in complex  $k$ . If this ratio is greater than the threshold, a candidate parameter set is drawn from a normal distribution centered on the mean of the active complex:

$$\Theta^{b+1} \sim N[\mu^k(C^k), c_n^2 \Sigma^k] \quad (5)$$

where  $\mu^k(C^k)$  is the mean of complex  $k$ . In MSU as in SCEM-UA, the threshold is a large value ( $10^6$ ) [8], so new parameter sets are regularly drawn from Equation (4). Equation (5) is included to prevent the Markov Chains from getting stuck in a particular region of the parameter space [8]. The parameter set generated by Equation (4) or (5) 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 over the range 0-1 [21]. This implies that the generated parameter set is always accepted if its likelihood is larger than the current parameter set, and it is still accepted with some probability 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,  $\Theta^{b+1} = \Theta^b$  and the following additional test is conducted. 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

[8]. If the likelihood of  $\Theta^{b+1}$  is greater than that of the worst point in the complex, the worst complex member is replaced with  $\Theta^{b+1}$ . This replacement removes points from the active complex with likelihoods that are too small. After all complexes have been updated, the individual complexes are re-sorted from most likely to least likely.

The SEM procedure is repeated  $m/5$  times for each complex [8]. 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 procedure is repeated.

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 [8, 20]. This ratio is the basis of Gelman and Rubin's Scale Reduction Score (SRS) [8]. MSU has exactly converged when the SRS of each parameter is equal to 1 [8]. Because this is very difficult to achieve in practice, SRS values of less than 1.2 are usually used to indicate approximate convergence [8, 20].

Parameter sets 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. In addition, one can examine the correlation between the values of different parameters in these generated parameter sets. Furthermore, the generated parameter sets can be used as the basis of model simulations

for the forecast period which produce samples of model outputs. The histograms of these 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 can be identified after MSU converges. This parameter set is then used in the BMA algorithm, which is described next.

## 2.2 BMA

BMA is a formal method used to analyze model uncertainty [12]. As Figure 1 shows, it develops a prediction for an output variable and associated uncertainty bounds using a weighted average of the forecasts that are produced by a collection of potential models. BMA assumes that the uncertainty associated with each model is represented by a normal distribution centered on its output value. The variance for each distribution and the weight for the associated model (i.e. the probability that it is the correct model given the observed data) are then estimated to maximize the likelihood of the observed data occurring. In the present application of BMA, the models under consideration are the sediment transport equations and the observed output variable is selected to be either bed profile elevations or measures of sediment grain sizes (see section 4). The rest of this section provides more details about BMA, but complete descriptions and evaluations of the method are provided in [13] and [14].

The central variable in BMA is the probability that the observed value of the output variable  $\Delta$  occurs under the calibration conditions given the individual model estimates  $f_1, \dots, f_I$  where  $i$  is an index of the available models and  $I$  is the total number of available models. This probability is denoted  $p(\Delta | f_1, \dots, f_I)$  and is calculated as:

$$p(\Delta | f_1, \dots, f_I) = \sum_{i=1}^I w_i g_i(\Delta | f_i) \quad (6)$$

where  $w_1, \dots, w_I$  are the probabilities that each model is the correct one given the calibration data (i.e. the posterior model probabilities) [13, 14]. The posterior model probabilities are nonnegative and add up to one, so they can be viewed as weights. The expression  $g_i(\Delta | f_i)$  is the probability of observing  $\Delta$  given model forecast  $f_i$ . Each conditional probability distribution  $g_i(\Delta | f_i)$  is assumed to be normal with a mean at its respective model forecast and a variance  $\sigma_i^2$ :

$$g_i(\Delta | f_i) \sim N(f_i, \sigma_i^2) \quad (7)$$

This distribution is meant to represent the uncertainty associated with model  $i$  assuming that it is the correct model structure. BMA estimates the weights and variances for all models in order to maximize their likelihood given the observed data. Observations are available at multiple times and locations, so this is done by maximizing the likelihood  $l$ :

$$\begin{aligned} l(w_1, \dots, w_I, \sigma_1^2, \dots, \sigma_I^2 | f_1, \dots, f_I, \Delta) \\ = \sum_{s,t}^{N_{st}} \log \left[ \sum_{i=1}^I w_i g_i(\Delta_{st} | f_{ist}) \right] \end{aligned} \quad (8)$$

where  $N_{st}$  is the total number of measurements over all  $s$  locations and  $t$  times in the calibration dataset and  $g_i(\Delta_{st} | f_{ist})$  is model  $i$ 's conditional probability for the observation given that model's forecast at space  $s$  and time  $t$ .

An iterative procedure called the Expectation-Maximization (EM) algorithm is used to solve for the values of the unknown weights and variances. This method is a widely used method for obtaining maximum likelihood estimates [22, 23]. The EM

method uses a variable denoted  $z_{ist}$ , where  $z_{ist} = 1$  if model  $i$  produces the best forecast at space  $s$  and time  $t$ ; otherwise,  $z_{ist} = 0$ . Even though the true values of  $z_{ist}$  are either zero or one, the values are not necessarily integers in the EM algorithm [13]. The algorithm begins with an initial guess for weights  $w_1, \dots, w_I$  and variances  $\sigma_1^2, \dots, \sigma_I^2$  as follows:

$$w_i = \frac{1}{I} \quad (9)$$

$$\sigma_i^2 = \hat{\sigma}_{i,obs}^2 \quad (10)$$

for  $i = 1, \dots, I$  where  $\hat{\sigma}_{i,obs}^2$  is the variance of the observations for the quantity of interest.

After this initialization, estimates of  $z_{ist}$  ( $\hat{z}_{ist}$ ) are calculating with the following equation:

$$\hat{z}_{ist}^{(j)} = \frac{w_i g[\Delta_{st} | f_{ist}, \sigma_i^{(j-1)}]}{\sum_{l=1}^I w_l g[\Delta_{st} | f_{ist}, \sigma_l^{(j-1)}]} \quad (11)$$

where  $g[\Delta_{st} | f_{ist}, \sigma_i^{(j-1)}] \sim N[f_{ist}, \sigma_i^{(j-1)}]$  and  $j$  is the iteration of the EM algorithm.

Calculation of Equation (11) represents the expectation step. The estimation of  $z_{ist}$  in Equation (11) relies on the assumption that the weights and variances are correct. Thus, subsequently, the estimates of the weights and variances are updated during the maximization step using the following formulas:

$$w_i^{(j)} = \frac{1}{N_{st}} \sum_{s,t} \hat{z}_{ist}^{(j)} \quad (12)$$

$$\sigma_i^{2(j)} = \frac{1}{N_{st}} \sum_{s,t} \hat{z}_{ist}^{(j)} (\Delta_{st} - f_{ist})^2 \quad (13)$$

The algorithm iterates between the expectation and maximization steps until changes in the likelihood, weights, variances, and  $z$  values are below predefined tolerances [14].

It is assumed that the weights and variances obtained by applying BMA to the calibration period apply to the forecast period as well [13]. Thus, the weights obtained from BMA can be applied directly to model outputs for the forecast period to obtain the BMA prediction. Confidence intervals of the BMA prediction give insight into overall uncertainty present in the model predictions. This uncertainty includes uncertainty due to the model structure (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 also determined by the MSU algorithm.

### 3 SRH-1D

SRH-1D is used to examine the uncertainty methodology as in Ruark *et al.* [6]. It is a 1D hydraulic and sediment transport model and is used by the U.S. Bureau of Reclamation in channel modeling projects [3]. The program is able to simulate channels with a variety of characteristics including non-mobile boundaries, steady flow, non-cohesive sediment [3], which are considered in this paper. SRH-1D contains several user-defined parameters in the equations to route water and sediment flows and to compute bed material mixing, which are described in the following paragraphs. For a more detailed description of SRH-1D's mathematical structure than is given below, please refer to [6] and [3].

To compute flow routing, SRH-1D solves the energy equation for steady, gradually varied flow using the standard step method [3]. This requires use of Manning's equation and, thus, 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 components: sediment routing and bed material mixing. Because all of the cases are bed load driven, suspended sediment transport is ignored and the Exner equation is used to calculate changes in the volume of sediment on the bed [3]. This simplifies the mass conservation equation that is the basis of sediment routing. Bed load transport capacity is calculated using one of the three following equations: the Parker equation

[15], the Wilcock and Crowe (W&C) equation [16], or the modified Meyer-Peter and Müller (MPM) equation [17]. Both the Parker and the W&C equations require specification of the critical shear stress  $\theta_r$ , which represents the amount of shear stress needed to initiate motion of the sediment, and a hiding factor  $\lambda$ , which accounts for hiding and exposure of particles with different sizes in mixtures. These two parameters are treated as uncertain in this paper. The MPM equation does not use a hiding factor and fixes the critical shear stress at  $\theta_r = 0.047$  [17]. SRH-1D uses a total adaptation length  $L_{tot}$  to calculate the length over which transport capacity is reached (instead of assuming that sediment discharge equals the transport capacity) [3]. The transport capacity length  $L_{tot}$  is a function of  $b_L$ , a user-defined parameter to compute the bed load adaptation length,  $f_s$ , the fraction of suspended load relative to the total load, which is estimated using an empirical function derived by Greimann *et al.* [24],  $\zeta$ , the suspended sediment recovery factor, and other characteristics related to the flow and the channel geometry [3]. Different values are used for  $\zeta$  for deposition and erosion:

$$\zeta = \begin{cases} \zeta_d & \text{if deposition} \\ \zeta_s & \text{if scour} \end{cases} \quad (14)$$

The bed load adaptation length  $b_L$  and the suspended sediment recovery factors  $\zeta_d$  and  $\zeta_s$  are considered uncertain parameters.

Bed material mixing is modeled by dividing the bed into layers: one active layer above several inactive layers. During deposition, the active layer shifts up and deposited material becomes part of the new 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  $\xi$ , which is a value between zero and one that defines the proportion of the bed load grain size distribution that is used to determine the grain size distribution of the sediment that is transferred between the active layer and the topmost inactive layer. Both  $n_{alt}$  and  $\xi$  are considered uncertain parameters.

## 4 EXPERIMENTS

Three flume experiments are considered as case studies in this paper; one is a depositional case and the other two are erosional cases. Following Ruark *et al.* [6], both the Seal *et al.* [25] and Ashida and Michiue [26] experiments are used. Because observational data are very limited in the Ashida and Michiue [26] experiment, another erosional case by Pender *et al.* [27] is examined. Bed profile elevation data and sediment size data are available for all three experiments. Sediment size data comes in the form of either profiles of grain size quantiles (i.e.  $D_{16}$ , the grain size that 16% of sediment is smaller than;  $D_{50}$ , the median grain size, and  $D_{84}$ , the grain size that 84% of sediment is smaller than), or fractions of sediment in size intervals. Table 1 provides a summary of the initial conditions, experimental inputs, and observations available for the three experiments.

The Seal *et al.* [25] experiment was designed to study sediment sorting during aggradation in three runs (named Runs 1-3 respectively). Downstream fining and armoring processes were observed in this experiment [25]. An abundance of observations from the experiment are available including bed profile elevations taken at typically 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 (Table 1), model parameters should remain unchanged between runs allowing for separate runs to be used for the calibration and forecast periods. 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 [26] experiment was designed to simulate bed degradation downstream of a dam. Bed profile elevation measurements are available at only three locations and six times within the 10 hour experiment, and the bed material distribution is reported only at the beginning and end of the experiment. The lack of sediment size measurements at an intermediate point in the experiment prevents separation of the case study into calibration and forecast periods that both contain observations. However, model predictions for hours 10 through 20 were simulated as though the experiment continued and used as a forecast period for the analyses described later.

The Pender *et al.* [26] 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 in this research since the observed data were easily acquired for this run. Bed profile elevation measurements are available for every 2 or 3 hours (over a total run time of 84.6 hours) at hundreds of locations along the bed at most times. For computational purposes, the number of observed locations used in MSU and BMA was decreased to between 21 and 42 points along the length of the bed for each time available depending on the availability of spatial observations at a given time. Hours 0 – 34.1 are used as the calibration period and hours 34.1 – 84.6 are used as the forecast period.

Table 2 shows the selected minimum and maximum possible 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. Note that  $\theta_r$  and  $\lambda$  are not variable when using the MPM equation in SRH-1D. Thus, this equation has 6 uncertain parameters while the Parker and W&C equations have 8 uncertain parameters. The range for the active layer thickness multiplier  $n_{alt}$  varies for each experiment, and in the Seal *et al.* [25] 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 need to be defined in the application of MSU. For all MSU runs, an initial population size of  $s = 500$  parameter sets is organized into  $q = 2$  complexes so that each complex contains 250 parameter sets at any one time. These values make the number of SEM evolutions for each complex before shuffling, which was defined as  $m/5$  and similar to the value used in [8], equal to 50 and were found to be the parameters that favored quick convergence for all MSU runs considered. The algorithm 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 the organization of the available observations into different variables, which are allowed to have different variances of their residuals. The first type of observation that is available for the flume experiments is bed profile elevation. It is assumed that bed profile 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. The general bed profile shape stays the same throughout each experiment and the scale of the measurements at all locations at a given time does not vary greatly. Bed profile 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. Aggregating observations from several locations together in this way allows for more reliable estimates of the variances of the residuals in the method. The other type of observation that is available is the sediment size data. Similar to bed profile elevations, all locations are assumed to have similar variances in their residuals, while different times are treated as different variables. When  $D_{16}$ ,  $D_{50}$ , and  $D_{84}$  observations are available, they are treated as three separate variables. When the fraction of sediment in size intervals is available, each size class is treated as a separate variable. Recall that the likelihood function used in MSU (Equation 2) assumes that the residuals for each variable are normally distributed. This assumption was tested after MSU was run and found to be false for some variables, however, transformations of the data did not significantly alter results from MSU or BMA and so were deemed unnecessary.

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 [13, 14] and provides a practical method of combining equations to modelers.

## 5 RESULTS

### 5.1 MSU Results

The first objective of this paper is to compare the efficiency of the MSU sampling method to the GLUE-based sampling method used by Ruark *et al.* [6]. Figure 2 plots the Scale Reduction Score for the uncertain parameters in the nine MSU runs (the three flume experiments, each simulated with the three different transport equations). The horizontal lines show a score of 1.2, which indicates approximate convergence, and the arrows indicate the approximate iteration number where convergence is achieved. MSU runs that use the MPM equation converge the fastest, and MSU runs that use the W&C equation converge the slowest. Because the MPM equation has only six uncertain parameters, it is not surprising that it converges fastest. The larger number of iterations for the W&C equation to reach convergence is unexpected because it has the same number of parameters as the Parker equation. However, the W&C equation has more difficulty simulating the observed data for all three cases than the other equations, which might produce the slower convergence. The number of iterations required for convergence with the Parker equation is very close to the sample size of 5000 used in Ruark *et al.* [6] to obtain posterior parameter distributions for the Parker equation with the GLUE methodology. Ruark *et al.* [6] found that a sample size of 5000 ensured consistent quantitative results between consecutive GLUE analyses. Based on this

comparison, it appears that the MSU methodology is not substantially quicker than GLUE for these experiments.

To evaluate the ability of MSU to constrain uncertain parameters, posterior parameter distributions were developed for all parameters. Figure 3 shows histograms that represent the posterior parameter distributions for  $n_{alt}$  for all MSU runs. For reference, the prior uniform distribution (Table 2) is also shown. Note the differences in the scales in these 9 plots. Figure 3 shows that  $n_{alt}$  can have different values and different degrees of certainty depending on the equation used and whether deposition or erosion is occurring (similar behavior is observed for other uncertain parameters). In general, MSU favors larger values for  $n_{alt}$  during deposition and smaller values for  $n_{alt}$  during erosion. To evaluate MSU's ability to constrain the parameters more generally, the percent change was calculated between the Interquartile Range (IQR) of the parameters generated from the prior distributions and the IQR of parameters generated the posterior distributions. The IQR is defined as the difference between the 75% and 25% quantiles of a data set. Table 3 lists the percent decreases for all uncertain parameters for the nine MSU runs. A value of 100% indicates that the algorithm identifies a single deterministic value of the parameter, while a very small value indicates that a large amount of uncertainty still surrounds the parameter. In general, parameters are better constrained for cases with more available observations (i.e. Seal *et al.* [25] and Pender *et al.* [27]). Also, certain bed material mixing parameters, namely the depositional and scour recovery factors, the bed load adaptation length, and the weight of bed load fractions are more easily constrained for the depositional case than for the erosional cases when the MPM equation is used. The performance of each individual

equation during the calibration period using the most likely parameter set found from MSU is compared to the observations in Figures 4 and 5. Note that none of the equations are able to capture the evolution of the sediment size fractions for the Pender *et al.* [27] case shown in Figure 5(e).

To assess the impact of using a multi-objective likelihood function on the identification of the most likely parameter sets, GOC values for parameter sets after convergence of MSU were analyzed. Recall that errors of two types of model outputs (bed elevation and sediment size) are incorporated into the calculation of the GOC. Thus, the GOC summation in Equation (2) can be separated into the terms that correspond to bed profile elevations and the terms that correspond to sediment grain sizes. Figure 6 plots the separated GOC terms on each axis after they have been normalized by the values of the most likely point (determined using the combined GOC). Each point in these plots represents a parameter set included in the sample after convergence. The normalization puts the most likely parameter set at coordinates (1,1). Any point that has a coordinate less than 1 would provide a better estimate for the observation type that is associated with that axis than the most likely point. Data from the Seal *et al.* [25] case with MPM was used to create Figure 6(a), data from the Pender *et al.* [27] case with Parker was used to create Figure 6(b), and data from the Ashida and Michiue [26] case with MPM was used to create Figure 6(c). These plots represent different cases that were observed. In Figure 6(a), the most likely point sacrifices some performance in both bed profile elevation and sediment size predictions to achieve an overall good performance. In Figure 6(b), the most likely point sacrifices bed profile elevation to obtain a better sediment size prediction. Note that if a parameter set existed that could optimize

performance of both types of model outputs at the same time, these graphs would have a square corner at point (1,1) like Figure 6(c). This situation rarely occurred in the MSU runs. Sediment grain size was the variable whose performance was usually sacrificed in the nine MSU runs. These trade-offs suggest that sediment transport modelers should be careful about merging a wide variety of data into a single, multi-objective likelihood function if one variable has particular importance in their model application.

Another 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 in the values of different parameters in the posterior distribution, the probability that the correlation observed between a pair of parameters has occurred by chance given that the true correlation is zero and approximately normally distributed was calculated at a confidence level of 95%. This analysis was done for all pairs of parameters using up to 5000 parameter sets after convergence (where available) 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  $\pm 0.1$ , and 19% of parameter pairs have a correlation coefficient stronger than  $\pm 0.4$ . More parameter pairs have stronger correlations in the more complex equations (Parker and W&C), which suggests that parameter correlations may be more important to consider when these two equations are used for modeling. The parameter pairs that are correlated and the value of this correlation both vary between cases and equations.

The implications of ignoring this correlation when assessing the uncertainty of model forecasts was explored by running two types of simulations with the parameter sets from the posterior distribution. First, up to 5000 parameter sets obtained after

convergence of MSU were used to simulate the forecast periods for each of the nine cases. Associated samples of bed elevation and sediment grain size profiles from SRH-1D at three times (beginning, middle, and end of the forecast periods) and three locations (upstream, midstream, and downstream of the flumes) were obtained. When the available sediment size data were fractions in different class sizes, three class sizes (small, medium, and large) were obtained. Then the average IQRs for the bed profile elevation data and the sediment grain size data were calculated. Second, the values for each parameter in the parameter sets from MSU after convergence were randomly reordered to remove any correlation between different parameters while maintaining the correct marginal distributions. The reordered parameters were used to simulate the forecast periods, and the IQRs of the same outputs were calculated. Figure 7 shows the average IQRs for the bed profile and sediment grain size data 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. For the Parker and W&C equations, inclusion of parameter correlation is more important when estimating bed profile elevation than sediment grain sizes for the depositional case (Seal *et al.* [25]). Interestingly, for both erosional cases, the results are opposite when examining the Parker and W&C equations. Specifically, parameter correlation is very important when estimating sediment grain sizes and not as important when estimating bed profile elevations. Overall, these results suggest that parameter correlations should be included when assessing uncertainty in sediment transport model forecasts.

## 5.2 BMA Results

BMA was used to determine weights for the three transport equations based on their ability to reproduce the observations, and these weights are reported in Table 4. Note that within each 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 case (Seal *et al.* [25]) with a weight of 0.84. Though the Parker equation performs best, the W&C equation also matches observations relatively well and has a weight of 0.16. The MPM equation dominates the profile elevation in both erosional cases (Ashida and Michiue [26], Pender *et al.* [27]) with weights of 1.00 and 0.98. The high weights associated with MPM for these experiments suggest that there is a high probability that it is the correct transport equation for these cases when predicting bed profile elevation. The forecast periods verify the inferences. In particular, the forecasts produced by weighting the transport equations as suggested by BMA match the observations better than the individual models do. Figure 8 illustrates this tendency for the Seal *et al.* [25] and the Pender *et al.* [27] cases, showing BMA predictions along with individual model predictions, the observations, and the 90% confidence interval on the BMA prediction (calculated based on the combined BMA distribution). Observations for the Ashida and Michiue [26] case were not available for the forecast period, so a similar comparison cannot be made in this case.

When sediment grain size is used as the observational data, the BMA results are rather different. As Table 4 shows, BMA suggests a different combination of equations for sediment grain sizes in all cases. For the Seal *et al.* [25] depositional case, a

combination of all three equations is suggested by BMA. For the Ashida and Michiue [26] erosional case, BMA suggests a combination of the Parker and MPM equations, while for the Pender *et al.* [26] erosional case, BMA suggests a combination of the Parker and W&C equations. Figure 9 shows the BMA predictions along with individual model predictions, observations, and the 90% confidence interval of the BMA prediction (calculated based on the combined BMA distribution) for the forecast period for sediment grain size outputs of the Seal *et al.* [25] and the Pender *et al.* [27] experiments. For the Seal *et al.* [25] case, the BMA combination provides a prediction that is a compromise of all of the  $D_{16}$ ,  $D_{50}$ , and  $D_{84}$  profiles. For the Pender *et al.* [27] case, however, the W&C equation actually outperforms BMA during the forecast period (although none of the equations performs particularly well and the uncertainty bounds are quite large). Here, it is worth noting that none of the sediment transport equations were able to capture the evolution of the sediment size distributions throughout the calibration period of the experiment.

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 output histograms generated from MSU (shown in Figure 7), which consider only parameter uncertainty, were compared to the average IQRs of the respective BMA predictive distributions. To calculate the IQRs for BMA, the same times, locations, and variables were used as in Figure 7. It should be noted that MSU and BMA are built on different statistical models for uncertainty as explained earlier, so the comparison performed here cannot be viewed as exact. Figure 10 compares the IQRs for the bed profile elevation and the sediment size outputs for the

three cases. In all cases, the IQR values from the BMA predictions, which include model uncertainty, are greater than the IQR values of the equations used in the BMA prediction, which represent only parameter uncertainty. For the Ashida and Michiue [26] case with the bed profile elevation data, the IQR for the Parker equation is larger than the BMA IQR, but the Parker equation is not used in the BMA estimate (Figure 10(c)). Recall that none of the proposed sediment transport capacity equations are able to capture the evolution of the sediment size distributions during the calibration period of the Pender *et al.* [26] case (Figure 5(e)), but parameters are constrained relatively well (Table 3). This behavior is reflected in the large difference between the IQR values of the three equations and the value that includes model uncertainty (Figure 10(f)). Here, the BMA IQR is more than 99% larger than all three equation IQRs. Compare this to Figure 10(d) which shows a smaller difference between the forecast uncertainty resulting from parameter values and the forecast uncertainty resulting from both parameter values and equations in sediment size distributions for the Ashida and Michiue [26] case. In this case, MSU is not able to constrain parameters very well, but BMA was able to match the observations (for the calibration period).

## 6 CONCLUSIONS

The research summarized in this paper explores methodologies for analyzing uncertainty due to parameter estimation (parameter uncertainty) and uncertainty due to the selection of a sediment transport capacity equation (model uncertainty) and attempts to improve on the methodology proposed by Ruark *et al.* [6]. MSU is used to evaluate parameter uncertainty, while BMA is used to explore model uncertainty. The conclusions inferred from the evaluations of these methods are summarized below.

- (1) Even though MSU has a more sophisticated means to develop parameter posterior probability distributions, it does not necessarily converge faster than the GLUE method used in Ruark *et al.* [6]. 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. The speed of convergence for MSU depends on the dimension of the optimization problem (number of uncertain parameters) and the transport equation being used. It can also depend on the prior distributions placed on the parameters. If some information is known about the correct posterior distributions of parameters beforehand, this

- information can be used in the definition of the prior distributions to possibly speed up convergence.
- (2) MSU is able to constrain the parameters in equations with fewer uncertain parameters (i.e. the MPM equation) more easily than the parameters in more complex equations (i.e. the Parker and W&C equations). When the MPM equations is used, it was also seen that certain bed material mixing parameters, specifically the depositional and scour recovery factors, the bed load adaptation length, and the weight of bed load fractions, could be more easily constrained for the depositional case than for the erosional case.
  - (3) In nearly all MSU runs, the most likely parameter set sacrificed performance with respect to one of the output variables, most often sediment size, in order to produce relatively good performance for both bed profile elevation and sediment grain size data. Rarely was one parameter set available that could produce the best performance for both types of observations at the same time. In cases where one variable is more important than the other for practical reasons, a modeler might choose to separate the variables rather than using a multi-objective likelihood function.
  - (4) Inclusion of parameter correlations substantially alters the estimation of uncertainty in the SRH-1D forecasts for some cases and thus should not be overlooked in uncertainty assessments. The importance differs between depositional and erosional cases and matters most when using equations with more parameters (Parker and W&C). In the depositional experiment by Seal *et al.* [25], it was found that parameter correlations are more important when modeling bed profile elevation than for

sediment grain sizes. This result is reversed for the two erosional experiments (Ashida and Michiue [26] and Pender *et al.* [27]). MSU accounts for parameter correlations whereas the GLUE method does not.

(5) 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 combination of equations from BMA produces a better forecast than using a single transport equation. It was also seen that 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.

The methodology examined in this paper should be tested on additional experiments as well as field cases. In particular, the case studies should be expanded to include experiments that study bed material load (which includes suspended load derived from the bed), which would substantially increase the number of transport capacity equations available for comparison. Field cases should consider reaches that involve both deposition and erosion. The multi-objective likelihood approach could also be applied to other output variables such as water profile elevation, water velocity, channel width, etc. In addition, further research could establish a multi-objective BMA method where weights represent the ability of an equation to correctly produce more than one output variable.

## TABLES AND FIGURES

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

Experiment	Seal <i>et al.</i> [25]		Ashida and Michiue [26]		Pender <i>et al.</i> [27]	
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 <sup>3</sup> /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 <sub>50</sub> ) (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 <sub>16</sub> , D <sub>50</sub> , and D <sub>84</sub> ) 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 Value	Maximum Value
Manning's roughness coefficient ( $n$ )	0.015	0.065
Critical shear stress ( $\theta_r$ )	0.01	0.06
Hiding factor ( $\lambda$ )	0	1
Active layer thickness multiplier ( $n_{alt}$ )	0.1	4 <sup>a</sup> , 6 <sup>b</sup> , 10 <sup>c</sup> , 15 <sup>d</sup>
Deposition recovery factor ( $\zeta_d$ )	0.05	1
Scour recovery factor ( $\zeta_s$ )	0.05	1
Bed load adaptation length ( $b_L$ )	0	10
Weight of bed load fractions ( $\xi$ )	0	1

<sup>a</sup>Pender *et al.* [27] case

<sup>b</sup>Ashida and Michiue [26] case

<sup>c</sup>Seal *et al.* [25] case with the Parker and W&C equations

<sup>d</sup>Seal *et al.* [25] 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 <i>et al.</i> [25]			Ashida and Michiue [26]			Pender <i>et al.</i> [27]		
	Parker	W&C	MPM	Parker	W&C	MPM	Parker	W&C	MPM
$n$	98	98	97	41	45	81	99	100	99
$\Theta_r$	94	96	-	71	63	-	92	93	-
$\lambda$	99	97	-	44	57	-	97	92	-
$n_{alt}$	68	67	95	52	78	87	92	87	84
$\zeta_d$	57	58	63	14	18	13	37	40	20
$\zeta_s$	55	63	100	19	27	44	91	100	52
$b_L$	87	82	93	15	15	12	60	94	22
$\xi$	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 <i>et al.</i> [25]	0.84	0.16	0.00
Ashida and Michiue [26]	0.00	0.00	1.00
Pender <i>et al.</i> [27]	0.02	0.00	0.98
	Sediment Grain Size BMA Weights		
Seal <i>et al.</i> [25]	0.27	0.42	0.45
Ashida and Michiue [26]	0.54	0.00	0.55
Pender <i>et al.</i> [27]	0.19	0.58	0.00

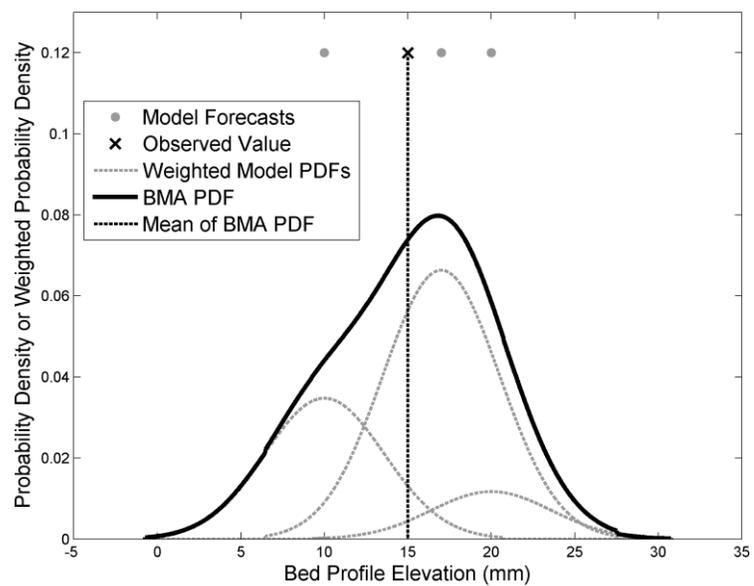


Figure 1: Example of BMA. The BMA PDF is the sum of the three weighted model PDFs.

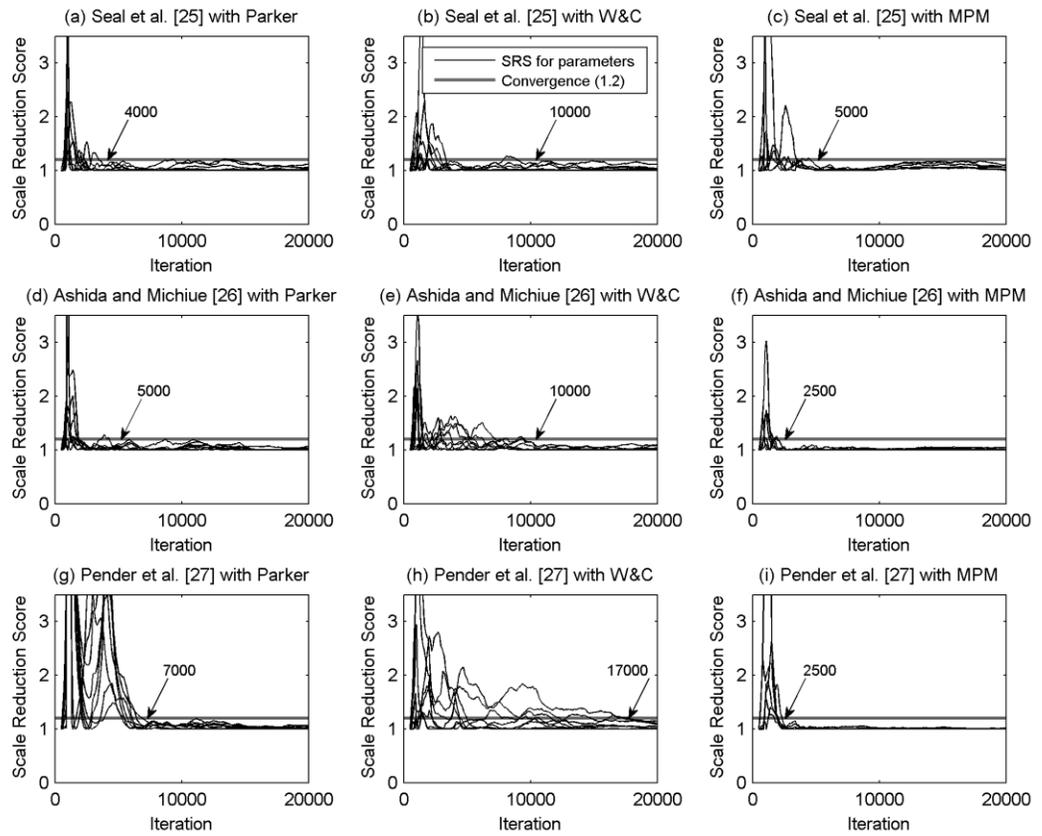


Figure 2: Gelman and Rubin's convergence diagnostic: the Scale Reduction Score (SRS) for the 20,000 iterations of MSU. The SRS is shown for all uncertain parameters, experiments, and equations. Arrows indicate the point of convergence in each plot.

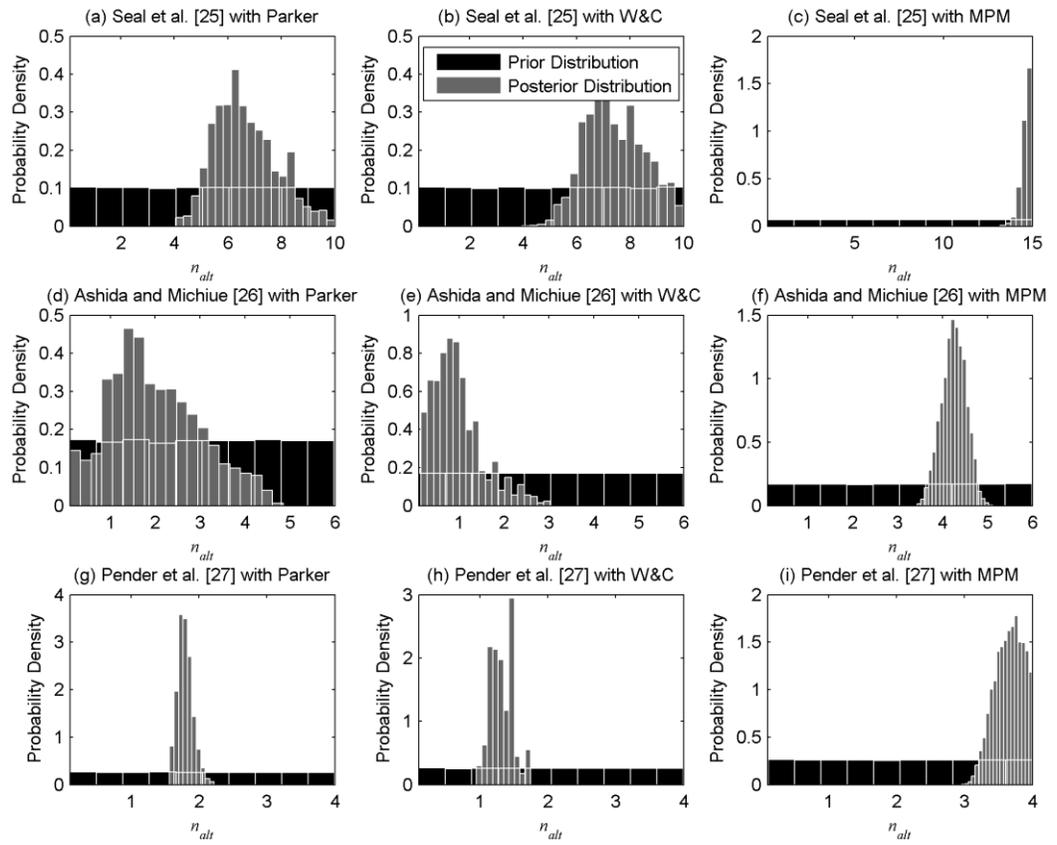


Figure 3: Histograms for the active layer thickness multiplier created from MSU samples after convergence for all experiments and equations. The original uniform distribution placed on the parameter is shown in black and its posterior distribution from MSU is shown in grey.

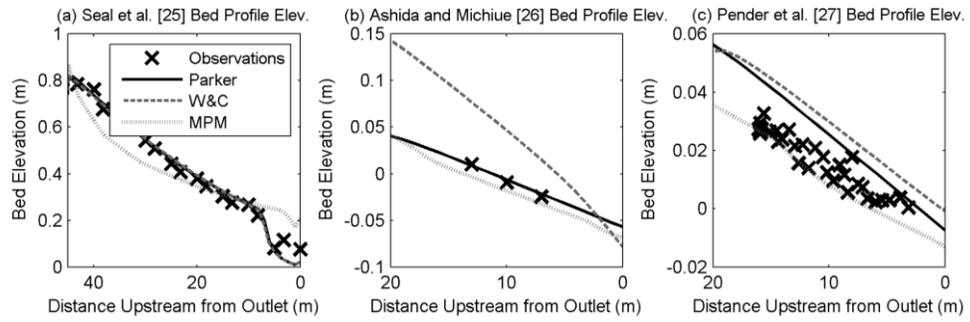


Figure 4: Individual model responses and corresponding observations of bed profile elevation for the calibration period of (a) the Seal *et al.* [25] case showing bed profile elevation at 32.4 hours, (b) the Ashida and Michiue [26] case showing bed profile elevation at 10 hours, and (c) the Pender *et al.* [27] case showing bed profile elevation at 32.1 hours

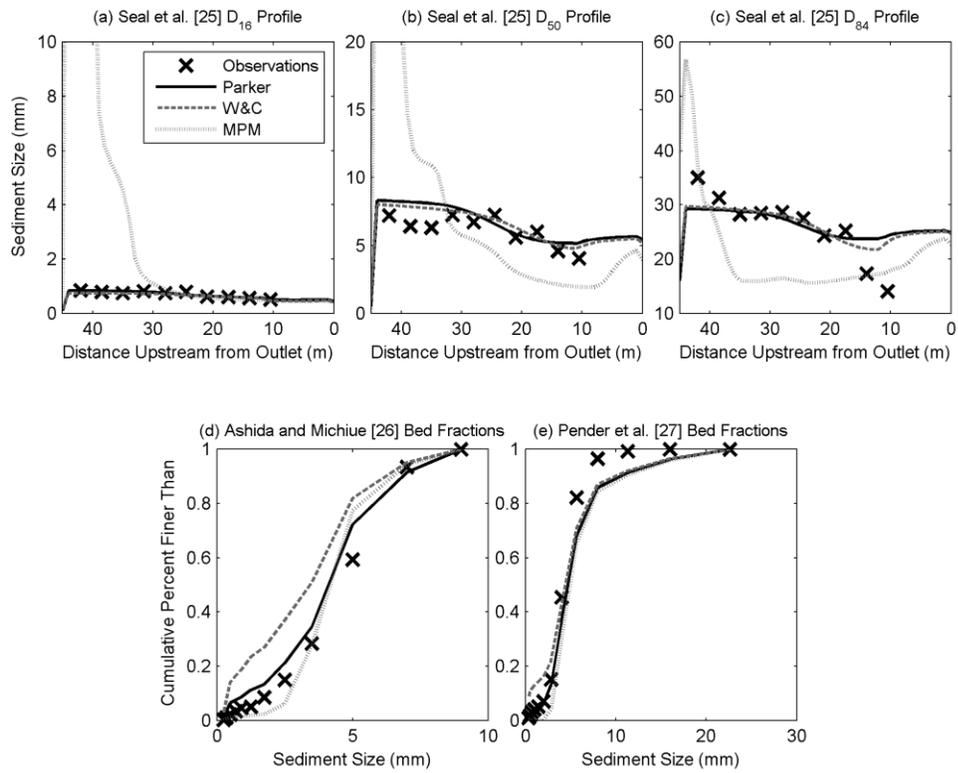


Figure 5: Individual model responses and corresponding observations of sediment size data for the calibration period of (a-c) the Seal *et al.* [25] case showing the (a)  $D_{16}$ , (b)  $D_{50}$ , and (c)  $D_{84}$  profile at 27 hours, (d) the Ashida and Michiue [26] case showing cumulative sediment size fractions at 10 hours, and (e) the Pender *et al.* [27] case showing cumulative sediment size fractions at 32.1 hours

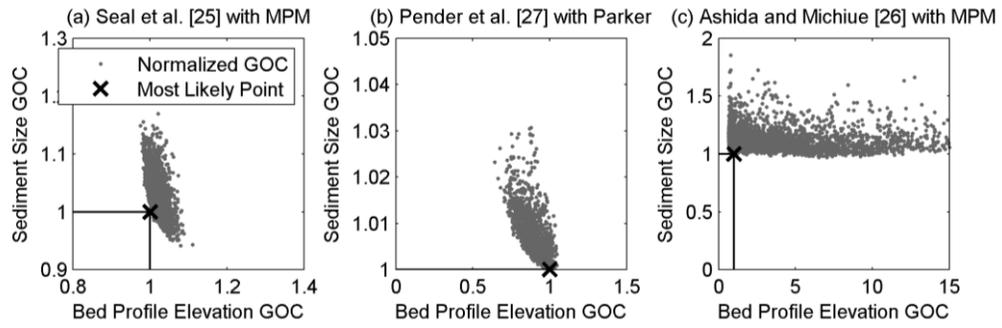


Figure 6: Plots of bed profile elevation GOC and sediment grain size GOC for the (a) Seal *et al.* [25] with MPM run, (b) Pender *et al.* [27] with Parker run, and (c) Ashida and Michiue [26] with MPM run. The most likely point is indicated with an x.

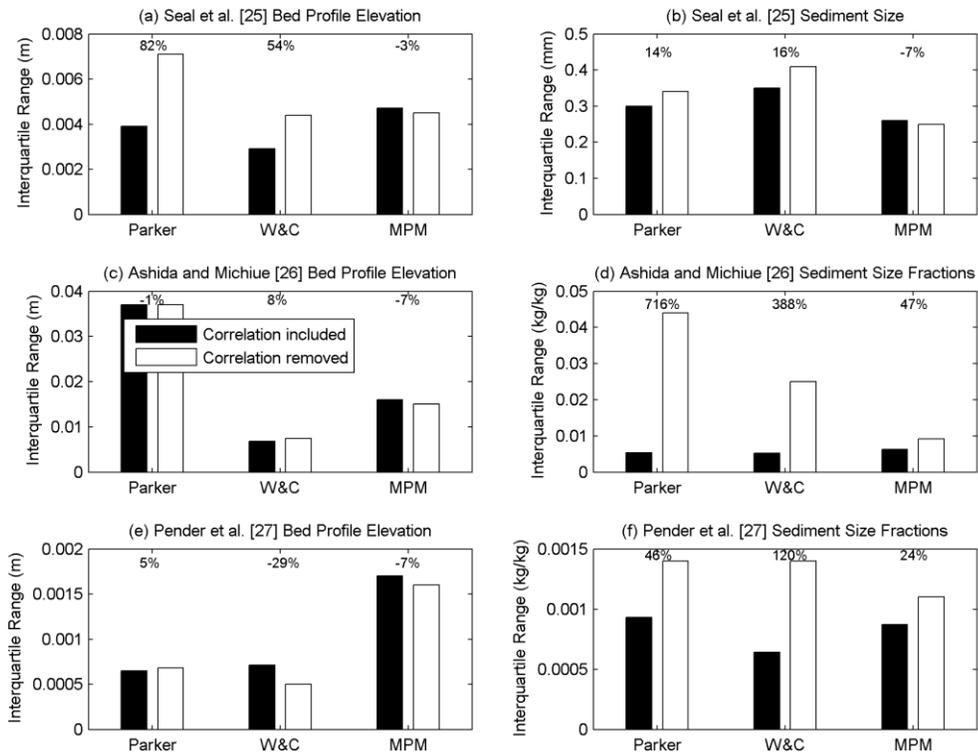


Figure 7: The average over time and space of the Interquartile Range (IQR) for the bed profile elevation or sediment size outputs (as labeled) from the forecast periods of the three experiments when simulated with the three transport equations. The black bars describe model outputs using parameter sets 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.

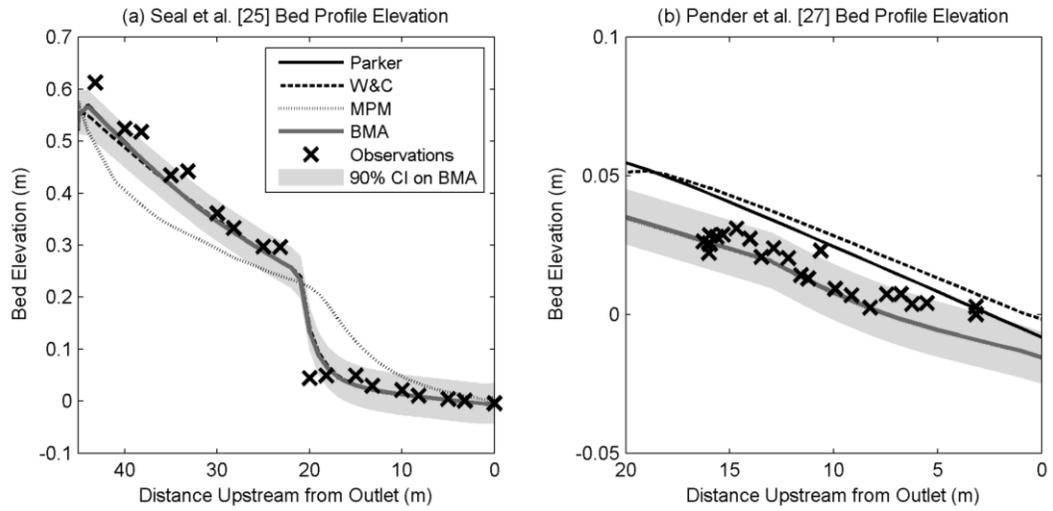


Figure 8: BMA predictions, individual model responses, corresponding observations, and the 90% Confidence Interval (CI) on the BMA prediction of bed profile elevation for the forecast period of (a) the Seal *et al.* [25] case showing bed profile elevation at 32 hours, and (b) the Pender *et al.* [27] case showing bed profile elevation at 62.4 hours.

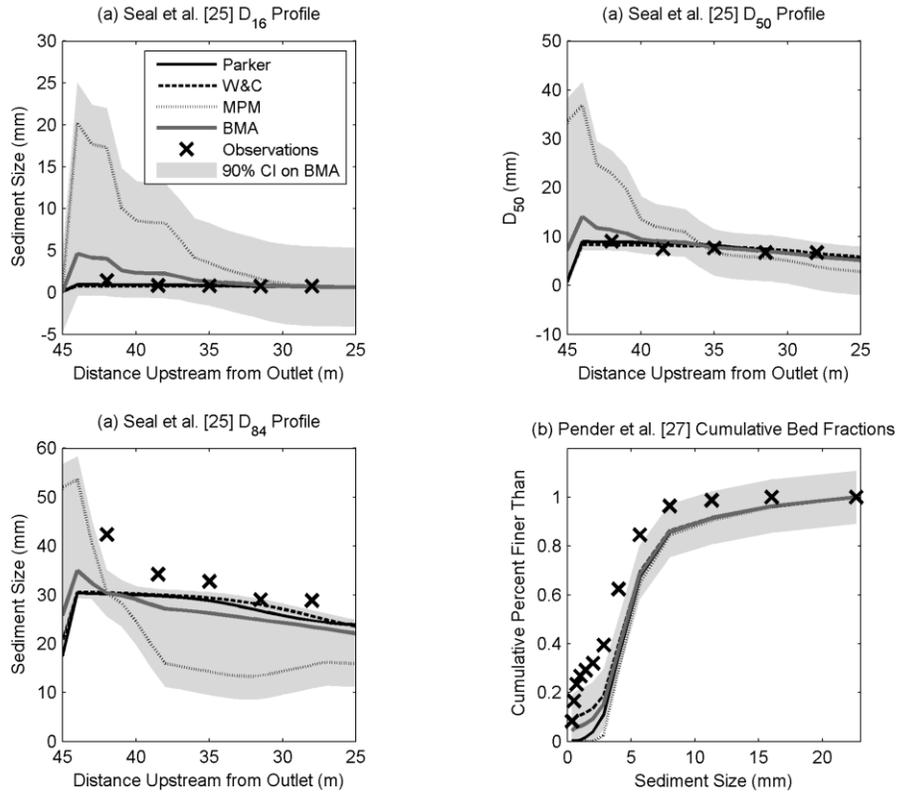


Figure 9: BMA predictions, individual model responses, corresponding observations, and the 90% Confidence Interval (CI) on the BMA prediction of sediment size data for the forecast period of (a) the Seal *et al.* [25] case showing the  $D_{16}$  profile at 34 hours, (b) the Seal *et al.* [25] case showing the  $D_{50}$  profile at 34 hours, (c) the Seal *et al.* [25] case showing the  $D_{84}$  profile at 34 hours, and (d) the Pender *et al.* [27] case showing cumulative sediment size fractions at 62.3 hours

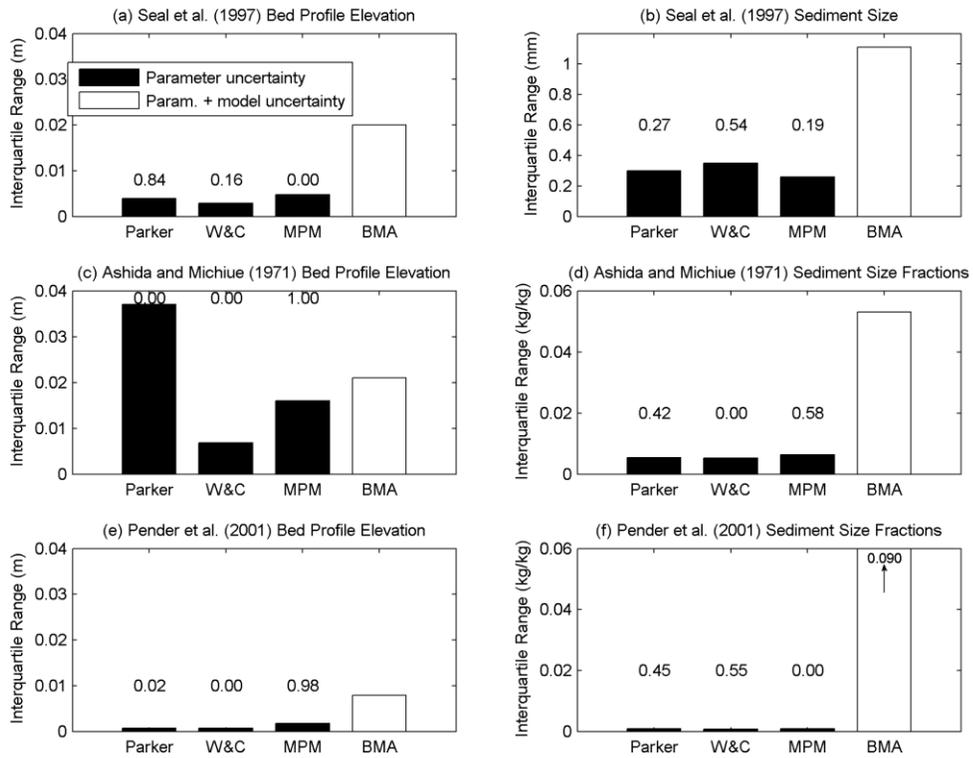


Figure 10: The average over time and space of the Interquartile Range (IQR) for the bed profile elevation or sediment size outputs (as labeled) from the forecast periods of the three experiments when simulated with the three transport equations. The black bars are associated with individual models and 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.

## REFERENCES

- [1] MacDonald LH, Smart AW, Wissmar RC. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska. Seattle: U.S. Environmental Protection Agency; 1991.
- [2] Liu Y, Gupta HV. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *J Hydraul Res* 2007;43(7).
- [3] Huang JV, Greimann B. User's manual for SRH-1D 2.6. Denver (CO): Bureau of Reclamation, Technical Service Center, Sedimentation and River Hydraulics Group; 2010 Oct. Sponsored by the Department of the Interior.
- [4] Chang CH, Yang JC, Tung YK. Sensitivity and uncertainty analysis of a sediment transport model: a global approach. *Stoch Hydrol Hydraul* 1993;7(4):299-314.
- [5] Yeh K, Lin E, Chen S. Effect of uncertainty of mobile-bed model parameters on bed evolution. In: Altinakar MS, Wang SSY, Holz KP, Kawahara M, editors. *Advances in Hydro-Science and Engineering, Volume VI*. Oxford: University of Mississippi; 2004.
- [6] Ruark MD, Niemann JD, Greimann BP, Arabi M. Method for assessing impacts of parameter uncertainty in sediment transport modeling applications. *J Hydraulic Eng* 2011;137(6):623-36.
- [7] Beven K, Binley A. The future of distributed models: Model calibration and uncertainty prediction. *Hydrol Processes* 1992;6(3):279-98.
- [8] Vrugt JA, Gupta HV, Bouten W, Sorooshian S. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour Res* 2003;39(8).
- [9] Carrera J, Neuman SP. Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data. *Water Resour Res* 1986;22(2):228-242.

- [10] Abramowitz G, Gupta H, Pitman A, Wang Y, Leuning R, Cleugh H, et al. Neural error regression diagnosis (NERD): A tool for model bias identification and prognostic data assimilation. *J Hydro-meteorol* 2006;7:160-177.
- [11] Van Griensven A, Meixner T. A global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models. *J Hydroinform* 2007;9(4):277-91.
- [12] Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian model averaging: A tutorial. *Stat Sci* 1999;14(4):382-417.
- [13] Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. Using Bayesian model averaging to calibrate forecast ensembles. *Mon Wea Rev* 2004;133(5):1155-74.
- [14] Vrugt JA, Diks CGH, Clark MP. Ensemble Bayesian model averaging using Markov Chain Monte Carlo sampling. *Environ Fluid Mech* 2008;8(5-6):579-95.
- [15] Parker G. Surface based bedload transport relationship for gravel rivers. *J Hydraul Res* 1990;28(4):417-36.
- [16] Wilcock PR, Crowe JC. Surface-based transport model for mixed-size sediment. *J Hydraulic Eng* 2003;129(2):120-28.
- [17] Wong MG, Parker G. Reanalysis and correction of bed-load relation of Meyer-Peter and Müller using their own database. *J Hydraulic Eng* 2006;132(11):1159-68.
- [18] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 1979;21(2):239-45.
- [19] Roberts GO, Gelman A, Gilks WR. Weak convergence and optimal scaling of random walk metropolis algorithms. *Ann Appl Probab* 1997;7(1):110-20.
- [20] Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci* 1992;7(4):457-72.
- [21] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. *J Chem Phys* 1953;21:1087-1091.

- [22] Givens GH, Hoeting JA. Computational Statistics. Hoboken: John Wiley & Sons, Inc.; 2005.
- [23] Dempster AP, Laird N, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B Stat Methodol* 1977;39(1):1-38.
- [24] Greimann B, Lai L, Huang J. Two-dimensional total sediment load model equations. *J Hydraul Eng* 2008;134(8):1142-46.
- [25] Seal R, Paola C, Parker G, Southard JB, Wilcock PR. Experiments on downstream fining of gravel: I. Narrow-channel runs. *J Hydraulic Eng* 1997;123(10):874-84.
- [26] Ashida K, Michiue M. An investigation of river bed degradation downstream of a dam. 14th International Association for Hydraulic Research Congress, Paris; 1971. p. 247-56.
- [27] Pender G, Hoey TB, Fuller C, McEwan IK. Selective bedload transport during the degradation of a well sorted graded sediment bed. *J Hydraul Res* 2001;39(3):269-77.

## APPENDIX A

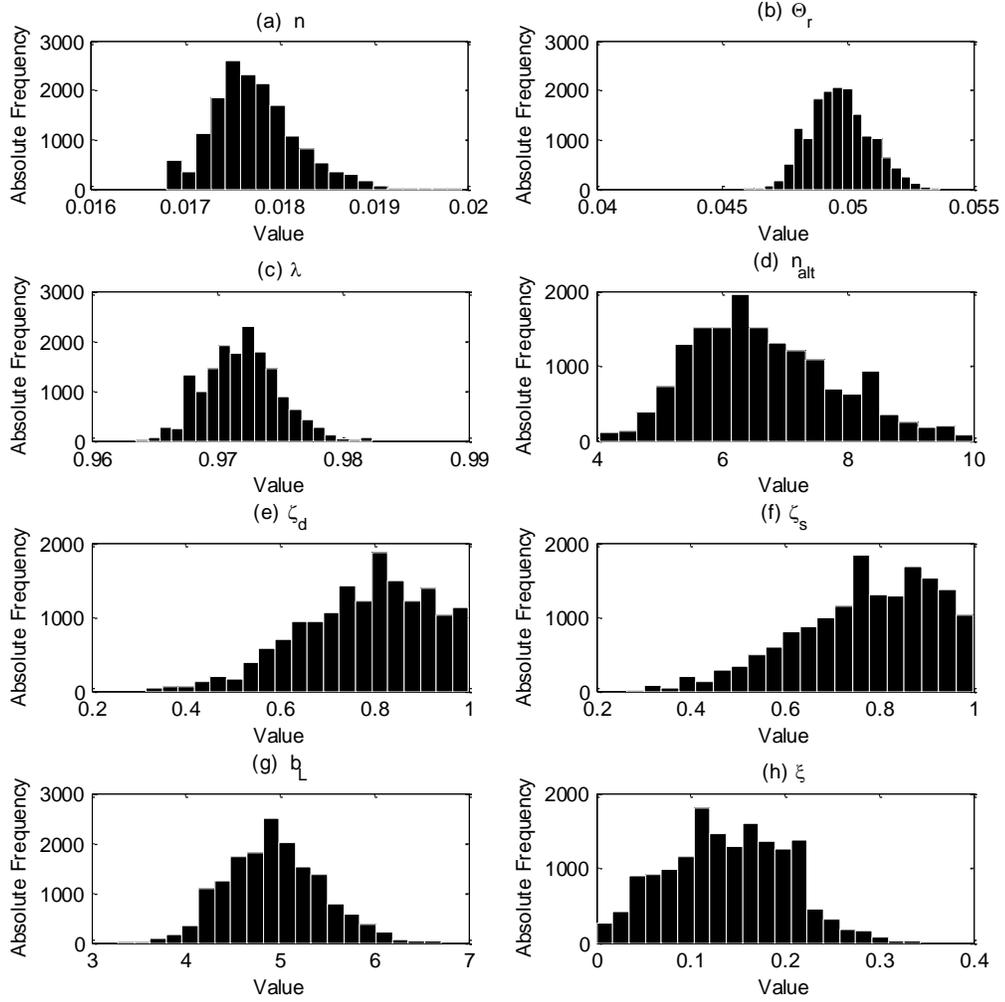


Figure 11: Posterior probability distributions for each parameter for the Seal *et al.* [25] case with the Parker equation.

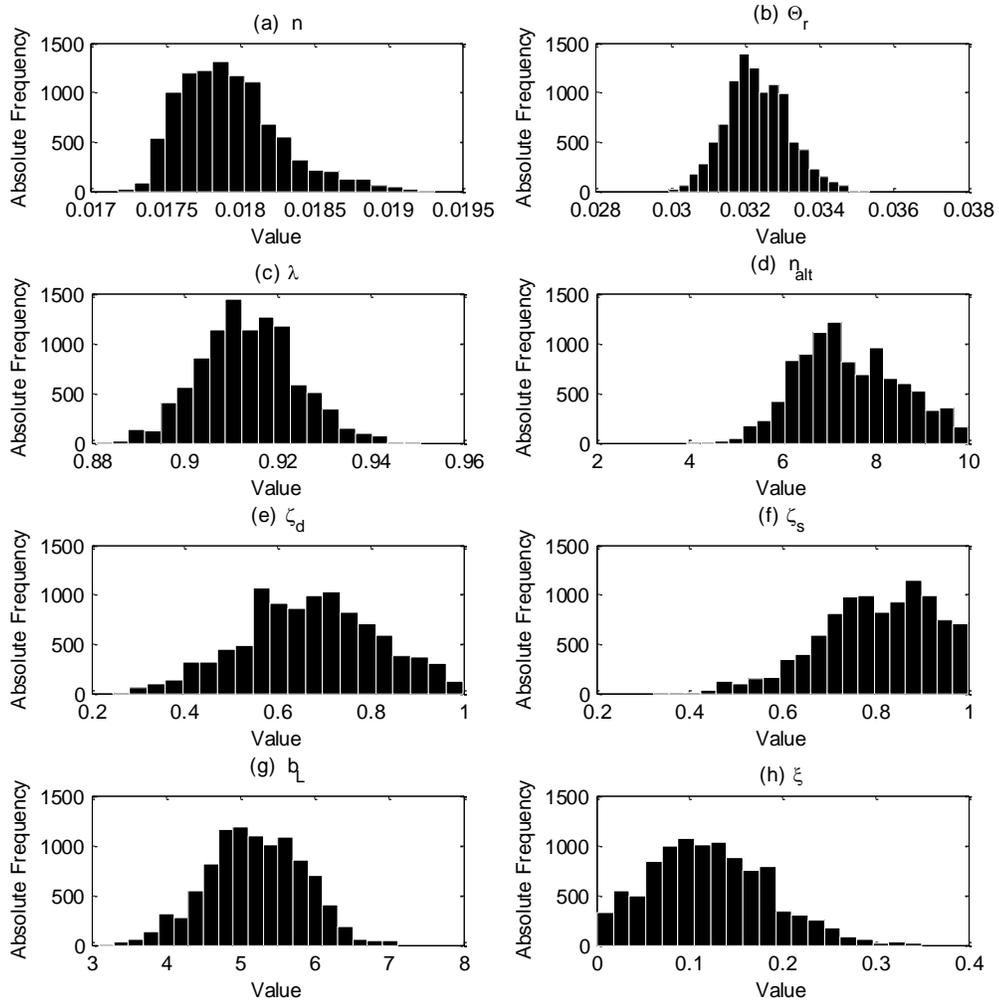


Figure 12: Posterior probability distributions for each parameter for the Seal *et al.* [25] case with the W&C equation.

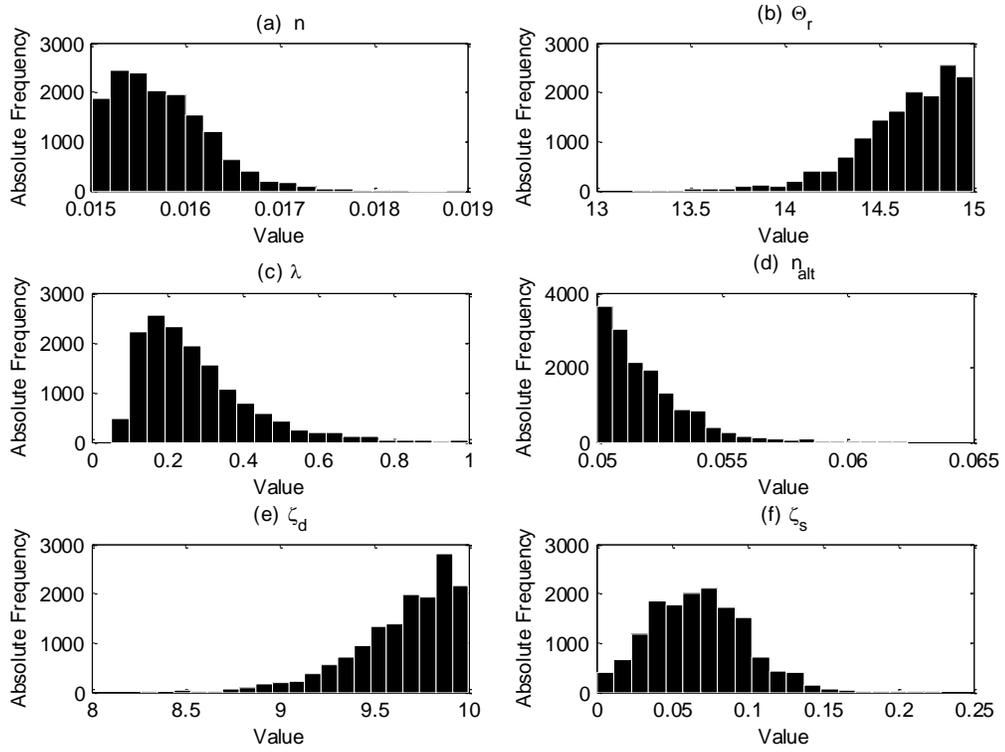


Figure 13: Posterior probability distributions for each parameter for the Seal *et al.* [25] case with the MPM equation.

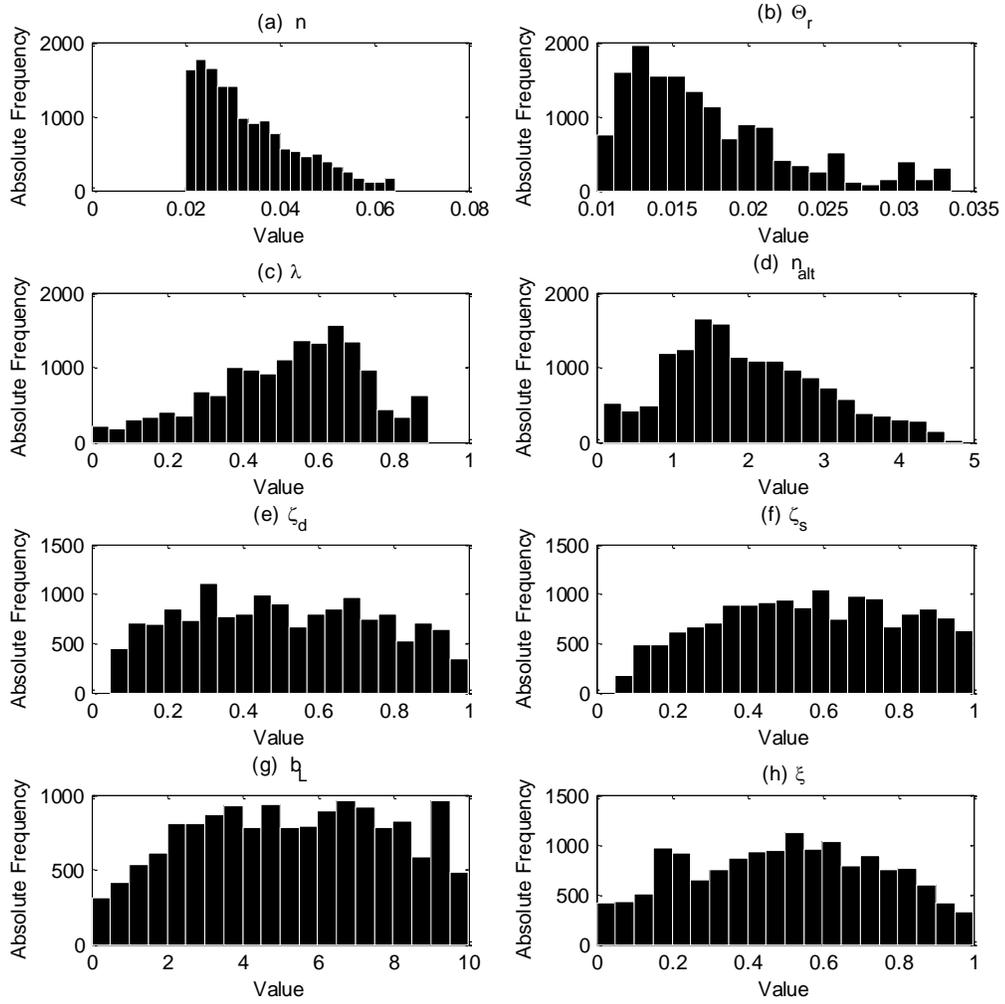


Figure 14: Posterior probability distributions for each parameter for the Ashida and Michiue [26] case with the Parker equation.

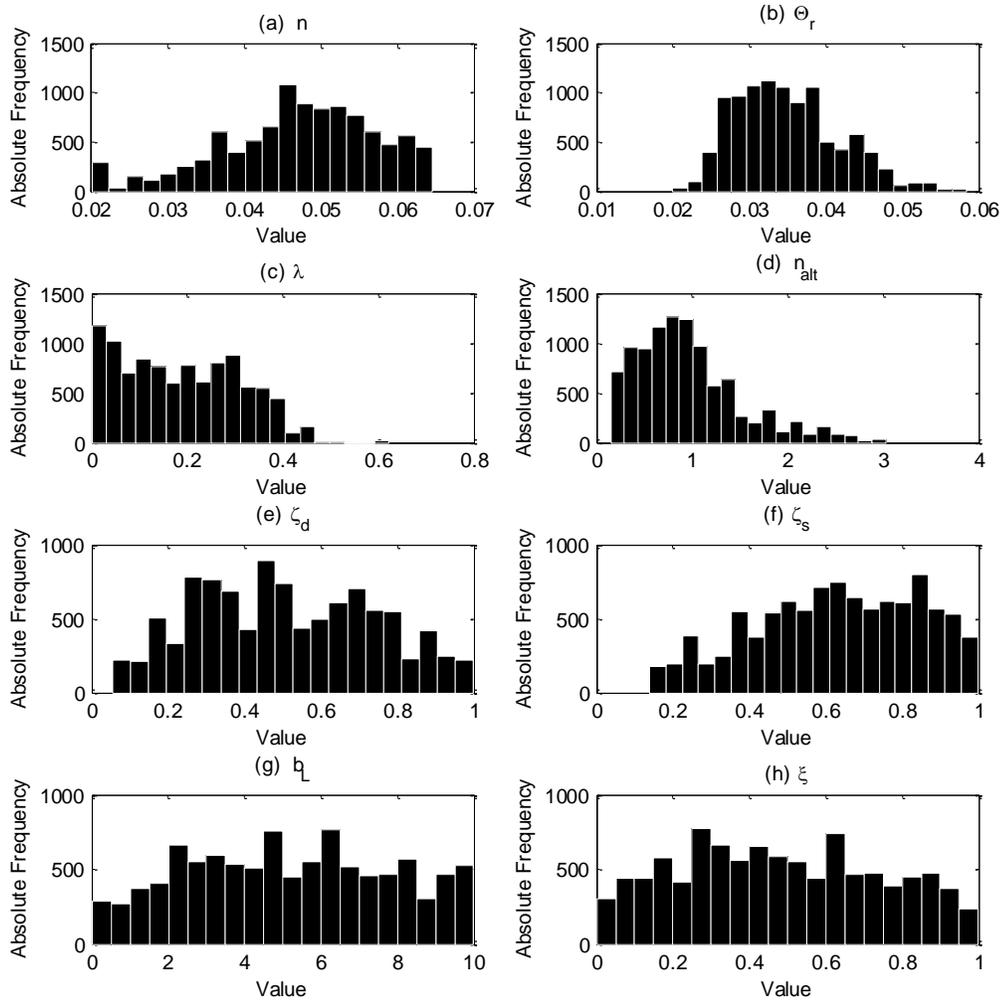


Figure 15: Posterior probability distributions for each parameter for the Ashida and Michiue [26] case with the W&C equation.

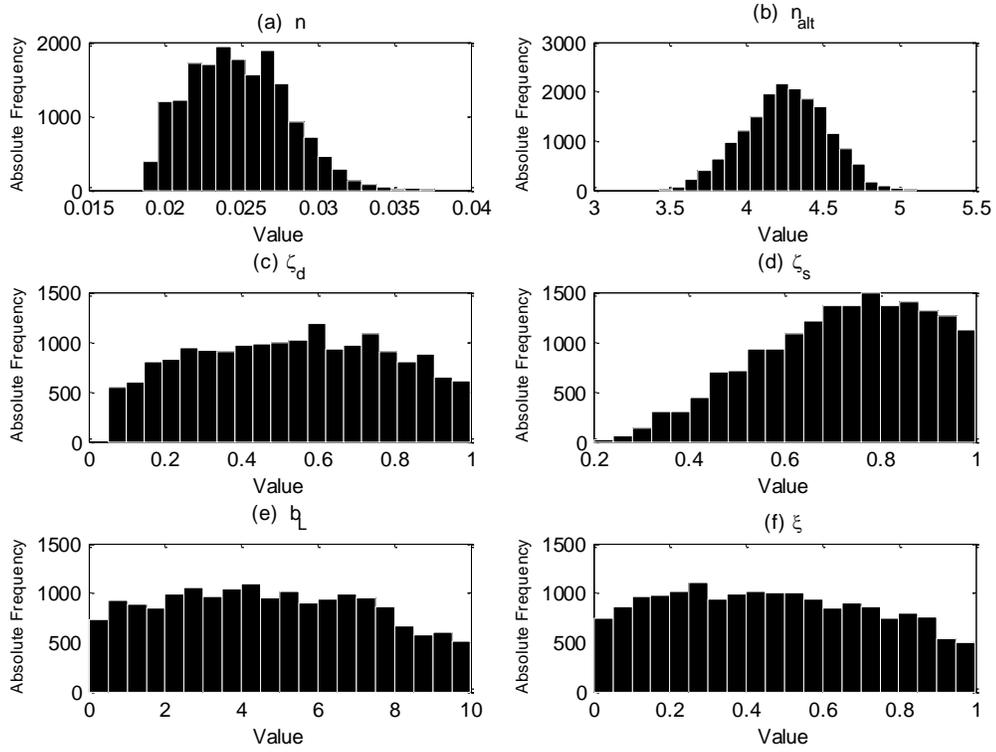


Figure 16: Posterior probability distributions for each parameter for the Ashida and Michiue [26] case with the MPM equation.

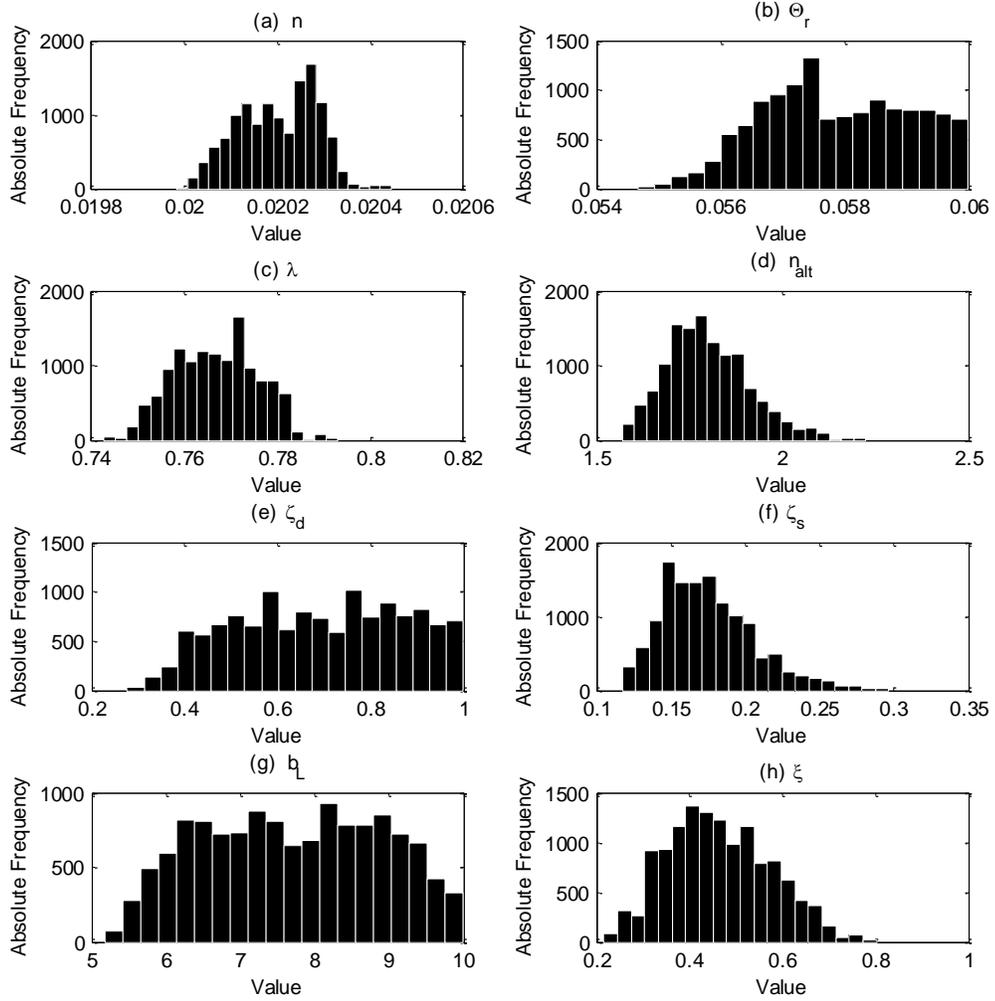


Figure 17: Posterior probability distributions for each parameter for the Pender *et al.* [27] case with the Parker equation.

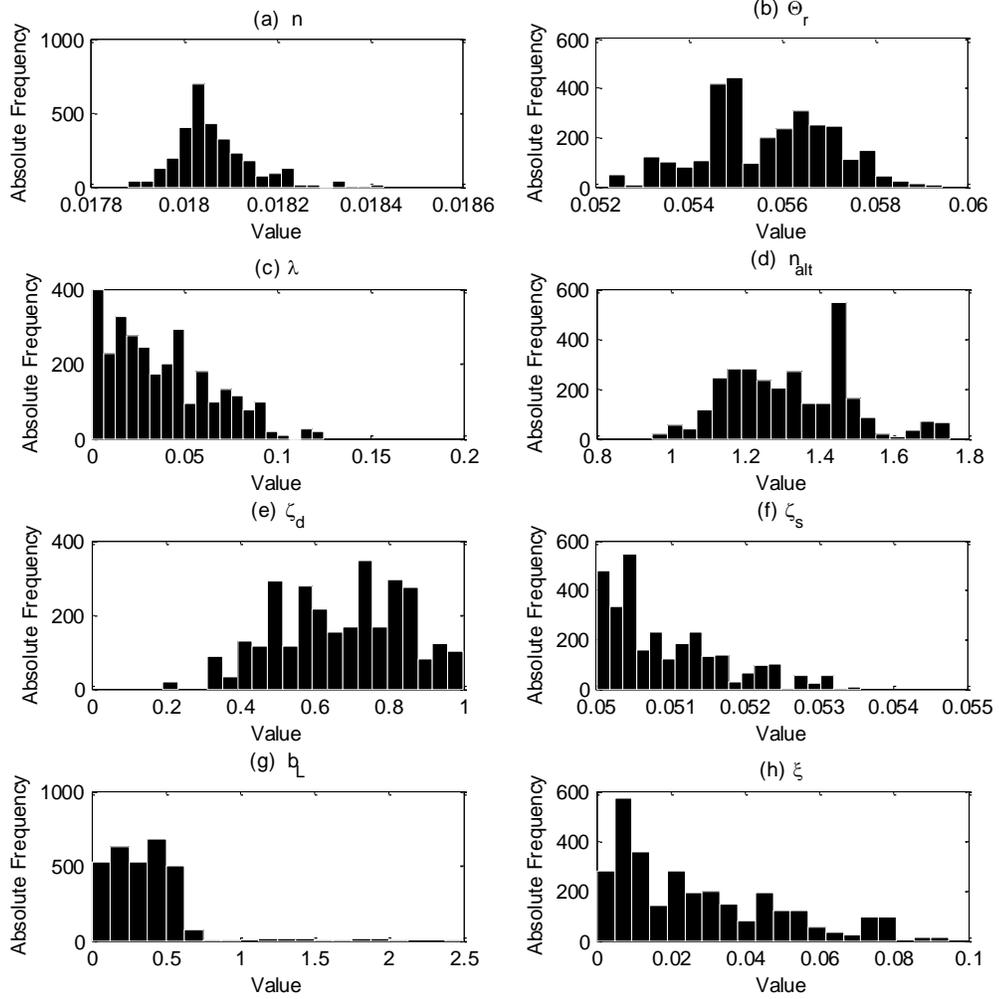


Figure 18: Posterior probability distributions for each parameter for the Pender *et al.* [27] case with the W&C equation.

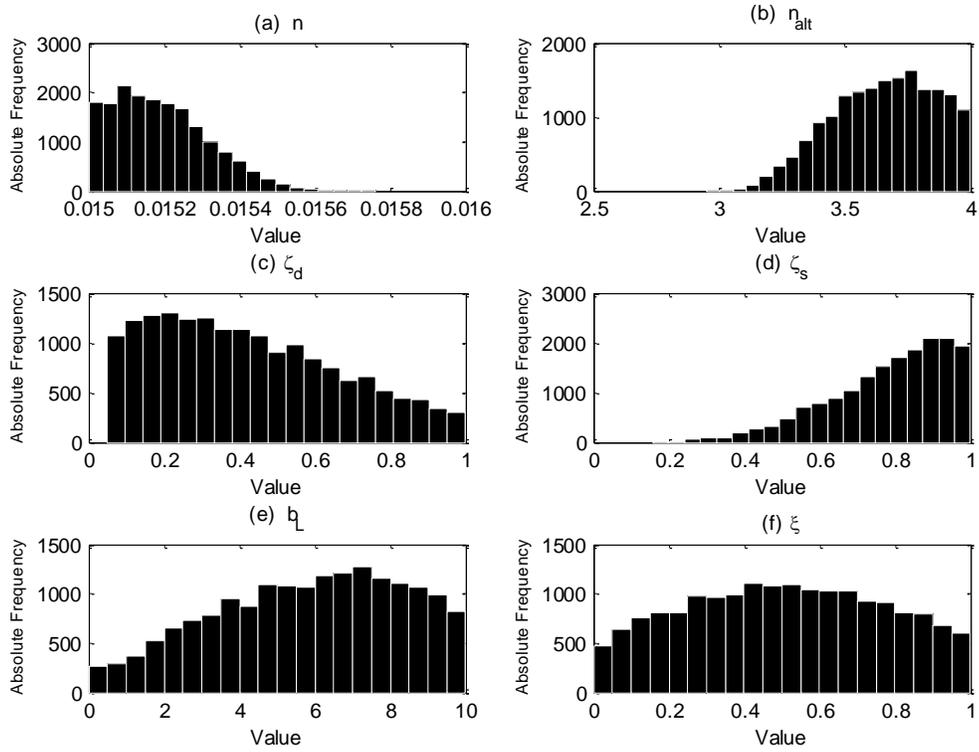


Figure 19: Posterior probability distributions for each parameter for the Pender *et al.* [27] case with the MPM equation.