Geomorphology 175–176 (2012) 1–13

Contents lists available at SciVerse ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

Accounting for uncertainty in cumulative sediment transport using Bayesian statistics

M.L. Schmelter [a,*], S.O. Erwin [b], P.R. Wilcock [c]

[a] Department of Civil and Environmental Engineering, Utah State University, Logan, UT 84322, USA
[b] Watershed Sciences Department, Utah State University, Logan, UT 84322, USA
[c] Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

ARTICLE INFO

Article history:
Received 4 November 2011
Received in revised form 8 June 2012
Accepted 13 June 2012
Available online 21 June 2012

Keywords:
Uncertainty
Probability
Sediment budgets
Sediment transport
Bayesian statistics
Gravel-bed rivers

ABSTRACT

That sediment transport estimates have large uncertainty is widely acknowledged. When these estimates are used as the basis for a subsequent analysis, such as cumulative sediment loads or budgets, treatment of uncertainty requires careful consideration. The propagation of uncertainty is a problem that has been studied in many other scientific disciplines. In recent years, Bayesian statistical methods have been successfully used to this end in hydrology, ecology, climate science, and other disciplines where uncertainty plays a major role—their applications in sediment transport, however, have been few. Previous work demonstrated how deterministic sediment transport equations can be brought into a probabilistic framework using Bayesian methods. In this paper, we extend this basic model and apply it to sediment transport observations collected on the Snake River in Wyoming, USA. These data were used previously to develop a 50-year sediment budget below Jackson Lake dam. We revisit this example to demonstrate how viewing sediment transport probabilistically can help better characterize the propagation of uncertainty in the calculation of cumulative sediment transport. We present the development of probabilistic sediment rating curves that rely on deterministic sediment transport equations and then show how these can be used to compute the distribution of sediment input and output for each year from 1958 to 2007. The Bayesian approach described provides a robust way to quantify uncertainty and then propagate it through to subsequent analyses. Results show that transport uncertainty is quantified naturally in the Bayesian approach, making it unnecessary for modelers to assume some specified error rate (e.g., ±5%) when developing estimates of cumulative transport. Further, we demonstrate that a Bayesian approach better constrains uncertainty and allows sediment deficit and surplus to be examined in terms of quantified risk.

1. Introduction

Estimates of sediment transport rate are widely known to have large uncertainty (Gomez and Church, 1989; Wilcock, 2001). Uncertainty poses a particular challenge for cases in which the cumulative transport is of interest. These include the delivery of sediment to reservoirs and other receiving waters, the supply of sediment to a river reach of concern, and the balance of input and output such that the net storage of sediment in a reach can be determined. Because these estimates involve propagating uncertainty over time, typically as a function of water discharge, defining a model that describes the uncertainty of the transport estimate is necessary.

The topic of uncertainty is widely treated in many scientific disciplines, and recent advances in statistical and computational methods have created a new set of tools at the disposal of researchers. One such method that is gaining prominence in diverse scientific fields is that of Bayesian statistical models. These tools provide a formal and theoretically solid framework in which deterministic process functions can be incorporated into a probabilistic framework thereby facilitating the quantification of parameter, structural, and predictive uncertainty. This approach has been used extensively in other disciplines, but has not been widely used in sediment transport applications. Griffiths (1982) was an early proponent of this approach and demonstrated geomorphically relevant examples, but these examples could be solved analytically. More complex models were unattainable until Geman and Geman (1984) and Gelfand and Smith (1990) demonstrated the use of Markov Chain Monte Carlo (MCMC) methods in Bayesian statistical analysis.

Schmelter et al. (2011) developed a simple Bayesian statistical model for sediment transport and demonstrated some of the benefits of this approach, including the ability to incorporate deterministic functions into a probability framework and to use prior knowledge and make predictions as probability distributions. In this paper we extend the basic Bayesian sediment transport model developed in Schmelter et al. (2011) and apply it to sediment transport observations made on a large gravel-bed river. Using the model predictions, we outline the development of a probabilistically based sediment mass balance using this modeling framework, and we evaluate the

* Corresponding author. Tel.: +1 602 618 0868.
E-mail addresses: mark.schmelter@aggiemail.usu.edu (M.L. Schmelter), s.erwin@usu.edu (S.O. Erwin), wilcock@jhu.edu (P.R. Wilcock).

0169-555X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.geomorph.2012.06.012

implications of viewing sediment transport probabilistically on the calculation of long-term sediment budgets.

1.1. Sediment budgets

One of the most challenging calculations in sediment transport is to determine the mass balance, or sediment budget, of a reach as the difference between input and output over a defined period of time. The cumulative transport at upstream and downstream sections, as well as any significant tributaries, must be determined; and an effective means is required to combine estimates of variance (due to errors or natural variability) at each section into a credible variance estimate for the net change in sediment storage. An alternative approach is to directly measure changes in sediment storage through field resurveys of channel morphology as well as the use of repeat aerial photography comparisons (e.g., Ashmore and Church, 1998; Eaton, 2001; Gaeuman, 2003). These methods, however, are not feasible for systems for which a historical survey, aerial or otherwise, is unavailable.

Thus, an alternate approach to constructing a sediment budget is to use the existing streamflow record and sediment rating curves to quantify annual sediment yield. This mass balance approach has been employed on the lower Colorado River (Schmidt, 1999; Topping et al., 2000; Hazel et al., 2006), the Sacramento River (Singer and Dunne, 2004), the Toutle River (Major, 2004), the upper Green River below Flaming Gorge dam (Andrews, 1986; Grams and Schmidt, 2005), the Fraser River (McLean et al., 1999a), and the Ebro River (Vericat and Batalla, 2006). While the reasoning behind constructing a sediment mass balance using historic streamflow records and sediment rating curves is sound, significant uncertainty is associated with any sediment rating curve. A common challenge concerns application of sediment transport relations beyond the period of available observations.

Sediment rating curves that do not account for the variability of bedload will at best provide a quantification of the mean behavior of the system, and often the extremes are of the most concern. What is not quantified in traditional approaches to sediment budgets is the variability one can expect to see in the sediment influx and outflux, as well as the magnitude of the difference between the influx and outflux required in order to be considered significant. In the absence of this information, a sediment mass balance may be indeterminate (e.g., Grams and Schmidt, 2005). While the approach to sediment budgets offered in this paper does not overcome the problems associated with non stationary sediment rating curves, it does address the issue of bedload variability for a given flow condition.

Although the uncertainties associated with estimates of sediment transport developed from sediment rating curves have been widely acknowledged, no universally accepted strategy for rigorously quantifying uncertainty in long-term estimates of sediment transport has been identified. Previous efforts to grapple with uncertainty in cumulative estimates of sediment transport include assuming that direct measurements of sediment transport are accurate to within some specified percentage of the total transport rate (Topping et al., 2000; Major, 2004; Grams and Schmidt, 2005), calculation of confidence intervals on the rating relation (Vericat and Batalla, 2006), or implementation of an error propagation analysis (Dunne et al., 1998). McLean et al. (1999a) used a Monte Carlo simulation to assess the precision of the Fraser River sediment sampling program. The Monte Carlo analysis was used to compute the coefficient of variation for replicate samples collected at a given vertical in an attempt to quantify the variability in transport rates resulting from both actual fluctuations in transport and sampling errors. In contrast, Singer and Dunne (2004) emphasized the role of streamflow variability in interannual estimates of bed material load. Toward this end, they coupled a stochastic streamflow model with calibrated sediment transport formulae to quantify variability in decadal estimates of bed material flux.

In what follows, we discuss the advantages of a Bayesian modeling framework over alternative approaches for quantifying uncertainty in sediment transport problems.

1.2. Bayesian models

Bayesian statistical models have gained increased prominence over the last 25 years owing in large part to advances in computing power and the development of sophisticated numerical methods. This modeling framework possesses several desirable properties that make it well suited for describing complex phenomena, such as sediment transport.

The first of these properties relates to the treatment of latent model parameters—those values which we cannot measure directly but must infer from observation. In deterministic approaches, model parameters (such as critical shear in sediment transport) are treated as fixed but unknown, and any variation around this fixed value is a metric of the uncertainty of the true single fixed value. It has long been acknowledged, however, that the threshold at which sediments move is not a fixed value but is, in fact, a probability distribution (e.g., Grass, 1970; Gessler, 1971; Paintal, 1971; Kirchner et al., 1990). Treating critical shear as a fixed value is therefore inappropriate and is largely for computational convenience. A Bayesian approach makes it possible to model (i.e., estimate or infer values of) latent parameters as random variables arising from probability distributions. The ability to model parameters, such as critical shear, as a random variable reconciles the long-established concepts of threshold with an appropriate modeling framework.

Because the underlying model parameters in a Bayesian approach are random variables, functions of these parameters (including predictions) are also random variables. The significance of this is that model predictions can be defined probabilistically. Deterministic approaches to sediment transport using fixed parameter values generally involve fitting a line through a distribution of observations. What is not quantified in the deterministic approach, however, is the extent to which deviations from this fitted line can be expected—it is well documented that sediment transport is variable even at steady-state conditions (Knighton, 1998; Hicks and Gomez, 2005; Turowski, 2010). One approach, then, is to make a set of Monte Carlo simulations using probability distributions for the model parameters thereby resulting in an ensemble of model predictions. This approach assumes that (i) the parameters are fixed and their variance is a reflection of uncertainty in the point estimate; and (ii) that all of the observed variability in sediment transport is directly because of parameter uncertainty. This approach will yield useful results only if the specified parameter distributions are correct. Returning to a Bayesian approach, distributions for the latent parameters (prior distributions) are assigned just as in the forward stochastic approach and these distributions are 'updated' by incorporating observations—sometimes called data assimilation (e.g., Wikle and Berliner, 2007)—of sediment transport events into the Bayesian statistical model. These updated distributions, called posterior distributions, are estimates of the underlying parameters in light of the observations and prior distributions. Posterior distributions are weighted combinations of prior knowledge and information from observation—the more observations you have, the less weight is placed on prior distributions. These inferred parameter values can then be used to calculate a posterior predictive distribution (PPD) as opposed to a single predictive line. Thus, the Bayesian model accomplishes both parameter estimation and stochastic prediction in one theoretical framework.

Bayesian statistical models, while based on probability, do not solely rely upon probability distributions or empirical relations without regard to the relevant physics of the process. In recent years, statistical approaches to bedload transport employing machine learning algorithms have been demonstrated (Bhattacharya et al., 2007; Dogan et al., 2009; Sasal et al., 2009) wherein the nonlinear functions

that govern transport are derived using sets of covariates and parameters trained to the observed bedload discharges. While machine learning algorithms seek to estimate an unknown function from covariates, the Bayesian model described in this paper incorporates deterministic equations. Machine learning techniques often result in good model fits, though the resulting parameters do not necessarily carry the same interpretability as do models that are based on equations derived from the physics of the phenomenon—this is especially true when the covariates are orthogonally transformed, such as in principal components analysis. As will be shown in what follows, incorporating deterministic or physics-based reasoning into Bayesian models is straightforward. This allows experts to select the deterministic models that are most appropriate and still retain the ability to model the phenomenon probabilistically.

Last, a Bayesian sediment transport model makes it possible to partition out variability—that is, to quantify the various sources of variability in the transport process. For instance, because model parameters are treated as random variables arising from probability distributions, we can expect some variance in predictions based on this alone. Further, sediment transport is spatially and temporally variable due to constantly changing hydraulic conditions, including changing bed topography, turbulence, varying supplies of bed material from upstream processes, and lastly irreducible noise brought about by stochasticity (Grass, 1970; Kirchner et al., 1990; Knighton, 1998; Gomez and Phillips, 1999; McLean et al., 1999b; Bunte and Abt, 2005; Hicks and Gomez, 2005; Diplas et al., 2008). An additional contributor of variance in model predictions is the conceptual model itself that we use to describe the process. Invariably, the models we use to describe physical systems are simplifications of reality, and many ways to describe the same process exist (Gomez and Church, 1989). Thus, the selection of a particular deterministic relationship will have an effect on the variability of predictions. Finally, data collection error will also contribute to observed variability in fluvial transport (McLean et al., 1999b; Diplas et al., 2008). The model presented in what follows distinguishes between variability owing to the model parameters being random variables, and variability because of stochasticity, measurement error, and model misspecification.

2. Study area

The Bayesian model described above is used here to quantify uncertainty associated with sediment rating curves developed from bedload transport data collected on the Snake River in northwestern Wyoming. The bedload transport data were collected for the purpose of developing a sediment budget for a 16-km stretch of the Snake River within Grand Teton National Park (GTNP) (Erwin et al., 2011). The study reach begins ~9 km downstream from Jackson Lake dam (JLD) at the confluence with Pacific Creek, shown in Fig. 1, and extends to Deadman's Bar.

The upper Snake River is a wandering, gravel-bed river. Through the study area, the Snake River and its tributaries flow through outwash produced during the Pinedale glaciation. Much of the Snake River is flanked by Pleistocene outwash terraces, which intermittently confine the valley (Love et al., 2003). Deadman's Bar is located in one of these confined areas, which makes it an ideal sampling location because the reach planform has changed very little over the last 50 years (Nelson, 2007). This eliminates any complications caused by transitions to multiple threads in the channel, or changes in sinuosity or braiding index.

Jackson Lake originally formed during the last glaciation when it was impounded by a recessional moraine. In 1906, JLD was constructed at the outlet of Jackson Lake to capitalize on this natural water storage location. The JLD increased the level of Jackson Lake by 11.9 m, creating 109 m^3 of storage in Jackson Lake reservoir. Importantly, the dam did not change the sediment supply to the upstream end of the study area because Jackson Lake existed prior to construction of JLD and the supply here has effectively been zero for thousands of years. Thus, the primary source of coarse sediment to the Snake River in GTNP is tributaries.

The upper Snake River and its tributaries drain the Teton and Absaroka Mountains and the Yellowstone Plateau. Although stream flow of the Snake River immediately downstream from JLD is entirely determined by dam releases, flow farther downstream results from the combined effects of dam releases and natural inflow from tributaries. The annual flood in the watershed is driven by spring snowmelt and typically occurs in May or June. However, the average annual peak flow is now significantly less than the unregulated peak flow due to dam operations (Marston et al., 2005; Erwin et al., 2011).

The field data set we used was collected from tethered rafts on three large gravel-bed rivers (see Fig. 2). The challenges in collecting the transport samples were exceptional, not least to collect sufficient samples to represent a highly variable transport field, and the resulting transport data demonstrate considerable scatter not uncommon in field data sets. Although other gravel rivers can be found with transport data sets that are more abundant and empirically define a tighter sediment rating curve, the abundance and scatter in the data is not a primary consideration in the basic purpose of this paper, which is to demonstrate application of a Bayesian approach in developing a more robust characterization of uncertainty and a more credible estimate of cumulative transport. In fact, the Snake River and its tributaries provide an appropriate illustration of how the Bayesian approach can be used to more effectively determine cumulative transport rates with field sediment transport data.

Developing a long-term sediment budget for the Snake River using the transport data of Erwin et al. (2011) faces two important challenges that merit attention. First, transport rates at the budget outlet (Snake River) were measured in 2007 whereas transport rates at the budget inlets (Buffalo Fork and Pacific Creek) were measured in 2006. Second, the cumulative sediment transport and sediment budgets are developed by applying transport observations from two years to a fifty-year flow record. This first issue requires the assumption that transport rates at the sites were similar in both years. The second issue requires the more severe assumption that the relation between flow and transport remain stationary over entire period of record. There is no specific information to inform these assumptions, although several factors suggest that the assumption is worth making. First, the broad patterns of sediment supply to the reach have remained unchanged throughout the Holocene (Jackson Lake Dam altered flows to the reach, but not sediment supply). Second, no significant changes in the reach planform have been observed since 1945 (Nelson, 2007). Third, sediment supply to the reach is abundant and tracer experiments indicate that common floods are capable of fully mobilizing the bed material. Together, these conditions suggest a transport (rather than supply) limited reach that has been in place for centuries, such that an assumption of stationary sediment rating curves is plausible.

The issue of rating curve stationarity is relevant to the accuracy of the long-term sediment budget determined from short-term data. This is a common and fundamental challenge that no method, statistical or otherwise, can resolve. Our goal in this paper is not to present and interpret a sediment budget for the reach (this has already been done by Erwin et al., 2011), but to demonstrate an approach that allows uncertainty in the budget calculation to be more effectively evaluated and applied to calculations of cumulative load. Any error arising from the assumption of stationarity does not impair this effort.

3. Methods

In this section we describe the methods used to develop the probabilistic sediment rating curves and sediment budget. Sediment input from tributaries and output from the mainstem were quantified

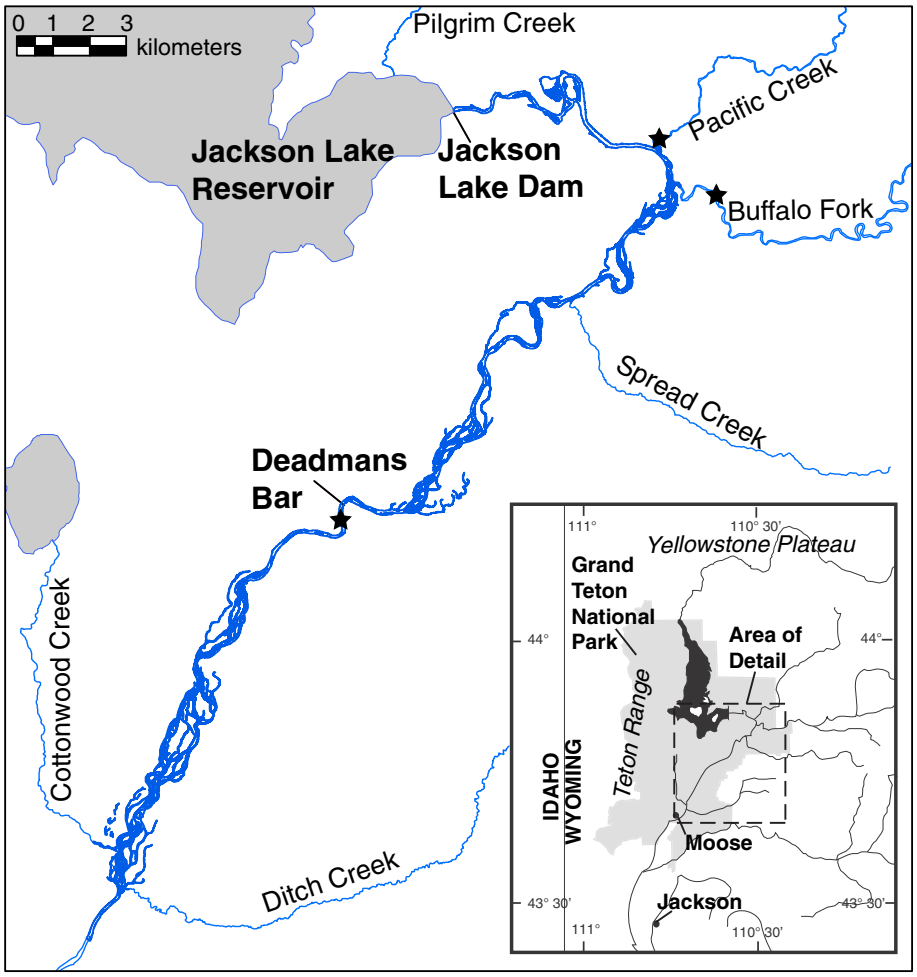

Fig. 1. The Snake River in Grand Teton National Park. The study area extends from the confluence of the Snake River with Pacific Creek downstream to Deadman's Bar. Locations of bedload sampling sites are indicated by stars.
Modified from Erwin et al. (2011).

Fig. 2. Sediment transport sampling on the Snake River at Deadman's Bar.

across a range of discharges allowing the construction of predictive sediment distributions for each location. These distributions were then used to construct long-term estimates of sediment influx and outflux using historical streamflow records.

3.1. Sediment transport observations

Three sampling sites were established for measuring bedload transport through the study reach. In 2006, bedload transport was measured on Buffalo Fork and Pacific Creek for the purpose of determining sediment inputs to the study reach. In 2007, bedload transport was measured on the Snake River at Deadman's Bar in an effort to quantify sediment outputs.

The sampling methods are described in detail in Erwin et al. (2011) but are summarized here. At each sampling site, transport rates were measured using a raft-based sampling platform and a Toutle River 2 (TR-2) bedload sampler (Childers, 1999; Wallick et al., 2010). Samples were collected using a modified version of the equal width increment (EWI) method, where each complete measurement consisted of one pass across the channel where 10–12 samples were taken at equally spaced intervals across the active bed. The sampler remained on the bed for 30–240 s at each vertical, and the time interval remained constant for each sample. All samples were sieved and weighed in 1/2-ϕ size classes. Channel conditions remained consistent during the 2-year sampling period so as to justify the assumption that there was no pronounced shift in the sediment rating curve from one year to the next.

3.2. Bayesian sediment transport model

3.2.1. Fundamentals

The basic premise of Bayesian models is discussed below. If we were to assume some generic process z as a function of some parameter θ, a Bayesian model of this process would be:

$$[\theta|z] = \frac{[z|\theta][\theta]}{\int [z|\theta][\theta]d\theta} \tag{1}$$

where the square brackets '[]' denote a probability distribution, and the vertical bar '|' denotes a 'given' such that $[\theta|z]$ is interpreted as the distribution of the model parameter θ given observations of the process z. The model in Eq. (1) is conventionally expressed as a proportionality, $[\theta|z] \propto [z|\theta][\theta]$, because the denominator in Eq. (1) is a fixed but unknown normalizing constant and MCMC does not require this normalizing constant to be known.

In words, the Bayesian model above allows us to make inference on the parameter θ by using the intrinsic information contained about it in observations of the process z. Given observations of z, it is possible to back calculate values of θ, but algebraic answers only return a single value. Because the Bayesian model treats parameters as distributions, the Bayesian answer is analogous to the back calculated parameter values for θ, except that it is a distribution instead of a single value. This back calculated distribution is the posterior distribution, $[\theta|z]$, which reads as 'the distribution of θ given observations of z'. The right-hand side of Eq. (1) consists of the likelihood and the prior. The prior distribution, $[\theta]$, is supplied to the model by the user. It is a statement of what is known about θ before the new observations were collected. This could be a summary of the literature on the parameter, an educated guess, or it can serve to constrain the parameter space to a physically plausible range. The likelihood, $[z|\theta]$, sometimes called the data model, represents the distribution of observations given the parameter θ. The likelihood describes the structure of the process being modeled and recognizes the fact that observations of z are dependent on the parameter θ.

Readers are referred to Schmelter et al. (2011) for more information on Bayesian models, specifically Bayesian sediment transport.

3.2.2. Sediment transport governing equations

As was mentioned previously, the Bayesian framework integrates deterministic functions into a probabilistic framework. It was established in Erwin et al. (2011) that the Parker (1979) and Parker (1990) models represented the observed Snake River transport data well. Further, Wilcock (2001) advocated their use because they are well-suited to predicting sediment transport over different ranges of grain shear stresses. These equations are used in the present paper to provide a representative comparison to the work done in Erwin et al. (2011) (see the cited publications for further justification). These relationships are:

$$W^* = \begin{cases} 11.2\left(1-0.846\frac{\tau_r}{\tau'}\right)^{4.5}, & \text{for } \frac{\tau'}{\tau_r} > 1 \\ 0.0025\left(\frac{\tau'}{\tau_r}\right)^{14.2}, & \text{for } \frac{\tau'}{\tau_r} < 1 \end{cases} \tag{2}$$

(Parker, 1979, 1990) where W^* is the dimensionless transport rate, τ_r is the reference shear (a surrogate for critical shear), and τ' is the skin friction. The value for W^* is defined as:

$$W^* = \frac{gq_s(s-1)}{\left(\frac{\tau'}{\rho}\right)^{1.5}} \tag{3}$$

where g is gravity, q_s is the unit bedload transport rate, s is the specific gravity (2.65), and ρ is the density of the water. The relationship used to calculate skin friction was derived in Erwin et al. (2011) and is provided here as a reference:

$$\tau' = 17k^{1.5}(SD_{65})^{0.25}Q^{1.5m} \tag{4}$$

where S is the slope, D_{65} is the 65th percentile grain size, Q is discharge, and k and m are empirical coefficients determined in Nelson (2007). Lastly, the conversion to dimensionless shear stress from shear stress follows:

$$\tau_r^* = \frac{\tau_r}{(s-1)\rho g D_{50}}. \tag{5}$$

Using the equations specified in Eq. (2) through Eq. (4), we can solve for the dimensional transport rate, Q_s, resulting in:

$$Q_s = \begin{cases} 11.2\left(1-0.846\frac{\tau_r}{\tau'}\right)^{4.5}\left(\frac{\tau'}{\rho}\right)^{1.5}\frac{W}{g(s-1)}, & \text{for } \frac{\tau'}{\tau_r} > 1 \\ 0.0025\left(\frac{\tau'}{\tau_r}\right)^{14.2}\left(\frac{\tau'}{\rho}\right)^{1.5}\frac{W}{g(s-1)}, & \text{for } \frac{\tau'}{\tau_r} < 1 \end{cases} \tag{6}$$

(Parker, 1979, 1990) where W is the channel width.

3.2.3. Bayesian model formulation

The basic model in Eq. (1) was adapted for sediment transport, resulting in the following:

$$\underbrace{\left[\tau_r,\ \sigma^2|log\left(\mathbf{Q_{s,o}}\right)\right]}_{\text{Posterior}} \propto \underbrace{\left(\prod_{i=1}^{n}\left[log\left(Q_{s,o,i}\right)|\tau_r,\ \sigma^2\right]\right)}_{\text{Likelihood}} \underbrace{[\tau_r]\left[\sigma^2\right]}_{\text{Priors}} \tag{7}$$

where $\mathbf{Q_{s,o}}$ is the vector of total cross section sediment discharge observations, $\mathbf{Q_{s,o}} = (Q_{s,o,1},\ldots,Q_{s,o,n})'$, assumed to be generally governed by Eq. (6). The model specified in Eq. (7) makes inference on both τ_r and σ^2, respectively the reference shear stress and the variance.

The likelihood is the distribution from which the observations arise, and this distribution must be specified directly. Section 3.2.2 outlined the deterministic models used to describe the dynamics of

sediment transport, and we refer to those equations as we construct the likelihood. In words, the following likelihood specification means that we believe sediment transport generally behaves according to the models developed in Parker (1979) and Parker (1990) while providing a term for the noise associated with this process, in measurements as well as in natural variability. To this end, we specified a normally distributed likelihood whose mean is defined by Eq. (6) with variance σ^2, as shown below:

$$log\left(Q_{s,0,i}\right)|\tau_r,\ \sigma^2 \sim N\left(log\left(Q_{s,i}\right),\ \sigma^2\right),\ \text{for each observation, i,} \quad (8)$$

where '~' is a statistical notation meaning 'is distributed as'. The likelihood specified above integrates a deterministic relationship that describes sediment transport into a probabilistic framework.

Because the model specified in Eq. (7) makes inference on two parameters, two prior distributions must be specified. In selecting prior distributions for parameters, care must be taken to ensure that the support of the distribution matches that of the parameter. For instance, we know that σ^2 must be greater than zero but less than infinity; and so an inverse gamma distribution ($\sigma^2 \sim I.G.(r,\ q)$, where r and q are 'hyperpriors') is an appropriate selection. For the reference shear parameter, we know that it, too, must be greater than zero but less than any transport event with $W^*<0.002$ (by construction). Thus, a lower- and upper-bound can be placed on τ_r. To this end, a truncated normal distribution was specified as the prior for reference shear ($\tau_r \sim T.N.\left(\mu_{\tau_r},\ \sigma_{\tau_r},\ \tilde{a}, \tilde{b}\right)$, respectively representing a mean, standard deviation, lower-bound, and upper-bound for τ_r). One advantage of the truncated normal distribution is that it is very flexible and can take on a variety of shapes depending on the specified hyperprior values $\left(\mu_{\tau_r},\ \sigma_{\tau_r},\ \tilde{a}, \tilde{b}\right)$, as shown in Fig. 3.

3.2.4. Parameter inference and prediction

One of the goals of formal statistical methods (both classical and Bayesian) is to make inference on latent or unobservable parameters. In order to do this, we must make observations of the process of interest after which we may then infer parameter values in context of the new observations. The posterior distribution in Eq. (7) is an updated joint distribution of the reference shear and variance parameters after considering the prior distributions (as specified by the expert) and the information available in the new observations. This distribution can be used to identify a credible interval or a range of values between which there is a $(1-\alpha)\%$ probability of containing the realized parameter value. The term 'credible interval' is a term in Bayesian statistics that is analagous to a confidence interval from classical statistics. The reason for this distinction is rooted in the interpretation of what probability truly represents. Readers are referred to any textbook on Bayesian statistics for a more detailed explanation (e.g., Gelman et al., 2004; Robert, 2007). This credible interval is computed by finding the $\alpha/2$ and $1-\alpha/2$ quantiles for the $(1-\alpha)\%$ credible interval. The distribution of these parameters is not constrained by requirements of normality (Schmelter and Stevens, in review). Most classical statistical methods assume that model parameters (or transformations thereof) are normally distributed. These assumptions are relaxed in a Bayesian approach and only prior distributions on the parameters are specified. As a result, these specifications of the prior distributions are based more on scientific merit than computational convenience (Box and Tiao, 1992). The posterior distribution for these parameters arises naturally from the prior distributions and likelihood and represents the updated knowledge on the parameters. These updated parameter distributions can then be used for prediction.

Prediction in the Bayesian framework is nicely integrated into the theoretical framework. The posterior predictive distribution represents the probability distribution of future, in this case, sediment transport events given the previously observed events. Mathematically, this is expressed as

$$\left[log\left(\tilde{Q}_{s,o}\right)|log\left(Q_{s,o}\right)\right] = \int\int\left[log\left(\tilde{Q}_{s,o}\right)|log\left(Q_{s,o}\right),\ \tau_r,\ \sigma^2\right]\left[\tau_r,\ \sigma^2|log\left(Q_{s,o}\right)\right]d\tau_r d\sigma^2. \quad (9)$$

The solution to Eq. (9) defines a distribution of sediment transport rates for a given flow condition. In essence, it is the probabilistic analog of a fitted line through observations. This distribution recognizes the natural variability in sediment transport noted in the literature (e.g., Hamamori, 1962; Knighton, 1998; Hicks and Gomez, 2005).

3.2.5. Computational methods

Because a solution to the posterior distribution in Eq. (7) is analytically intractable, the posterior must be sampled using MCMC with a Gibbs sampler and Metropolis–Hastings update to simulate observations from the posterior distribution. Readers are referred to Casella and George (1992), Chib and Greenberg (1995), Hastings (1970), and Metropolis et al. (1953) for further information on these methods. The integration required for the posterior predictive distribution in Eq. (9) is performed by composition sampling of the MCMC posterior distributions (Tanner, 1996). More detailed information is available in Schmelter et al. (2011). The *R* programming environment (R Development Core Team, 2010) was used to implement this model.

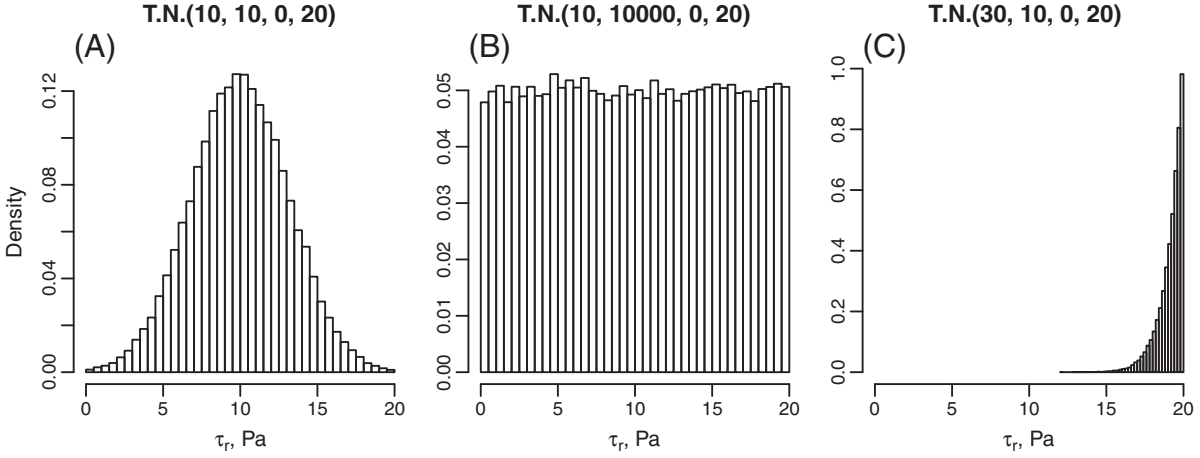

Fig. 3. Potential prior distributions for reference shear. Prior (A) specifies a mean value for reference shear centered on 10 Pa, with some variance; prior (B) assumes that every value between zero and the upper-bound is equally likely; prior (C) gives higher probability to values near the upper limit of the prior distribution. Priors are parameterized according to $T.N.\left(\mu_{\tau_r},\ \sigma_{\tau_r},\ \tilde{a}, \tilde{b}\right)$.

3.3. Sediment budget calculation

The approach to developing a probabilistic sediment budget is very similar to traditional approaches except that, in the probabilistic approach, we assume that a distribution of sediment transport rates—instead of a single prediction—may occur for any given flow condition (as defined by the posterior predictive distribution). The first step is to get the time-series of streamflow (daily mean discharge) at the location where a sediment rating curve has been developed. Assuming a period of record of I years, take day j of year i and determine the streamflow. Using this discharge, get the corresponding sample of sediment transport rates from the Bayesian rating curve. Because the posterior predictive distribution consists of a set of samples (as opposed to some analytic equation) these M samples (M should be some large number) are stored as the jth column in a matrix—each column corresponding to the sediment flux for day j in year i. This is repeated for each day in year i. Because annual sediment yield is simply the sum of all the daily sediment fluxes, the distribution of annual sediment yield is the sum across all J columns of the matrix—this results in a single column of M summations comprising the distribution of annual sediment yield (it should be noted that Nelson (2007) showed that there was no progressive channel change over the period from 1945 to 2002, thus we assumed that lateral inputs are trivial relative to the tributary inputs.) This process is repeated for all I years.

For the Snake River sediment budget presented here, annual sediment yield was calculated for Pacific Creek and Buffalo Fork to determine sediment inputs to the study area and for the Snake River at Deadman's Bar, to determine sediment output. Because JLD completely disrupts sediment supply from the mainstem to the study area, inputs are calculated as the sum of tributary inputs. The sediment budget was calculated for 1958 to the present, the period influenced by modern rules for dam releases. Because these influxes and outfluxes are distributions, their quantiles can be calculated that determine some interval for annual sediment yield at a $(1-\alpha)\%$ credibility.

4. Results

Sediment transport observations and the Bayesian transport model specified in Eq. (7) were used to estimate the parameter distributions for τ_r and σ^2 at three river sampling sites: Buffalo Fork, Pacific Creek, and the Snake River at Deadman's Bar (see Fig. 1). Table 1 shows the prior distributions used for each site and Fig. 4 shows the prior and posterior distributions for $\tau_r$and σ^2 at Deadman's Bar. The priors and posterior results for Buffalo Fork and Pacific Creek mimic the general shapes shown in Fig. 4 but with shifted locations. The prior for reference shear is diffuse and is functionally uniform over the specified interval. The prior for the variance parameter has a large variance as well to make it vague, though the inverse gamma distribution's shape is less flexible than the truncated normal. These priors were selected so that very little information is assumed, thereby allowing the observations to provide the most information about the inferred parameter values. Table 2 presents the inferred values and credible intervals for reference shear and variance from the Bayesian model; values and intervals for reference shear from Erwin et al. (2011); and channel characteristics for each sampling location. Using these posterior parameter distributions for reference shear and variance, the posterior predictive distribution was calculated for each cross-section. These predictive distributions are presented in Fig. 5(A) through (C) and are plotted as areas defining the 68%, 90%, and 95%.

Fig. 6 is a schematic diagram that compares the assumptions associated with a forward stochastic and a Bayesian model for sediment transport. The top half of Fig. 6 (dashed boxes) corresponds to the forward stochastic model and the bottom half to the Bayesian approach. The ensemble predictions for Deadman's Bar shown in Fig. 6(G) uses a uniform distribution of reference stresses on the expert-defined uncertainty envelopes from Erwin et al. (2011). The ensemble predictions were constructed using 7000 samples (to match the number of MCMC samples from the Bayesian model) of reference shear. These predictions assume τ_r to be the only random variable (no log-additive noise). Fig. 6(H) is the PPD for Deadman's Bar. Fig. 6(H) assumes τ_r and σ^2 are random variables and that the *log()* of sediment transport predictions has constant variance defined by σ^2 (see Fig. 6(B) and (F)). While the Bayesian model described in this paper only assumes two parameters, Fig. 6 illustrates that this is extensible to an arbitrary number of parameters (see Fig. 6(C) and (D)).

Figs. 7 and 8 show the credible intervals for the sediment mass balance. The sediment inputs are represented by the red regions in Figs. 7 and 8. The outflux is represented by the black regions in the same figures. Inset plot (A) of Figs. 7 and 8 represents the $(1-\alpha)\%$ credible region for annual sediment yield. For any given year there is a distribution of sediment influx and outflux, and inset plot (A) shows the width of the $(1-\alpha)\%$ credible interval for these distributions. Sediment influx and outflux, however, is not uniformly distributed over this interval. Fig. 9 shows actual predictive distributions of influx and outflux for the years 1960 and 1984 at $\alpha=0.05$ credible level along with the $\alpha/2$ and $1-\alpha/2$ quantiles shown as vertical lines. These lines are used to establish the credible interval width in inset plot (A) of Figs. 7 and 8. Fig. 9(A) shows that the 1960 influx and outflux credible intervals overlap for $\alpha=0.05$, while Fig. 9(B) shows a distinct gap between the 1984 influx and outflux values for $\alpha=0.05$. These elements help define what constitutes a 'significant' difference in influx and outflux and are also seen in Fig. 7(A) for 1960 and 1984. Inset plot (B) of Figs. 7 and 8 shows the cumulative sediment mass balance at the chosen credible level starting in 1950. These distributions are found by summing the preceding annual sediment yields up to the selected year for both influx and outflux. Inset plot (C) shows the posterior distribution of cumulative sediment yield at the end of the period of record, 2007. This distribution (with both traditional quantiles and highest posterior density (HPD) interval (Box and Tiao, 1992) denoted by the vertical lines) shows the probability density function for the difference between the cumulative influx and the cumulative outflux for the study reach.

Table 1
Prior distributions for τ_r and σ^2; the truncated normal distributions are specified as $T.N.\left(\mu_{\tau_r},\ \sigma_{\tau_r},\ \tilde{a}, \tilde{b}\right)$; the inverse gamma distribution is specified as $I.G.(\mu_{\sigma^2},\ \sigma_{\sigma^2})$.

Parameter	Snake River	Buffalo Fork	Pacific Creek
$[\tau_r]\sim$	$T.N.(7,\ 1000,\ 0,\ 15)$	$T.N.(7,\ 1000,\ 0,\ 15)$	$T.N.(10,\ 1000,\ 0,\ 20)$
$[\sigma^2]\sim$	$I.G.(0.5,\ 50)$	$I.G.(0.5,\ 50)$	$I.G.(0.5,\ 50)$

5. Discussion

5.1. Bayesian sediment transport model

The model proposed in this research is an extension of the Bayesian formulation presented in Schmelter et al. (2011). Here we demonstrate that by modifying the deterministic equation specified as the mean value of the likelihood function, Eqs. (8) and (2), the simple Bayesian formulation presented in Schmelter et al. (2011) can be used to model sediment transport in a large gravel-bed river. Using non informative priors for the model parameters, we obtained estimates for τ_r and σ^2, as shown in Table 2. From a numerical methods standpoint, the MCMC algorithm converged quickly on the parameter space and exhibited the necessary properties—e.g., mixing, acceptance rate; (Gelman et al., 2004)—characteristic of stable and valid computations. The posterior predictive distributions shown in Fig. 5 are further evidence of model validity; the predictive distributions appear to be reasonable given the observed transport rates. By

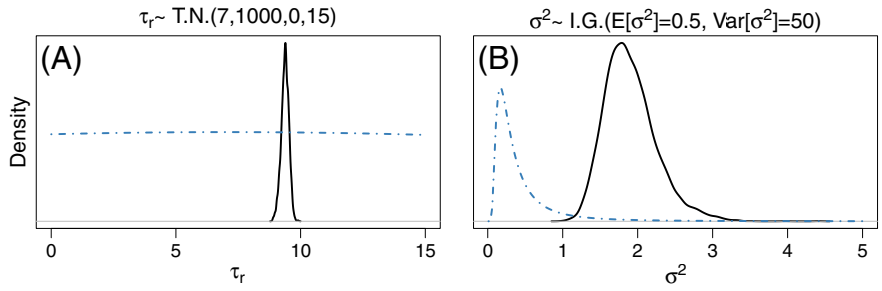

Fig. 4. Prior (dash-dot) and posterior (solid) distribution densities for reference shear and variance at Deadman's Bar on the Snake River. (A) Reference shear; (B) variance.

all accounts, the Bayesian method infers reasonable parameter values and produces predictions that are sensible given the observations.

Regarding parameter inference, we can compare the inferred values from the Bayesian model to the values fitted manually in Erwin et al. (2011). One characteristic of these parameter estimates is that while the point estimates are generally the same for reference shear between the two approaches, the credible interval of the Bayesian estimates is much tighter than the manually fitted uncertainty envelopes used in Erwin et al. (2011). This discrepancy is largely owing to the fact that in the manually fitted calibration all of the observed variation is assumed to be solely attributable to variation in τ_r (see Fig. 6). The Bayesian model, however, partitions between parameter variability (as expressed by the credible interval for τ_r) and variability because of random noise, measurement error, and model misspecification (expressed by the variance parameter σ^2).

While the uncertainty envelopes used in Erwin et al. (2011) is a more simplistic approach to quantifying uncertainty in transport predictions, they only define an upper and lower limit for reference shear and do not of themselves define a distribution. What happens between these bounds is not defined. It could be flat (uniformly distributed), skewed (e.g., lognormally or gamma distributed), or symmetric (e.g., normally distributed). To produce stochastic predictions using these uncertainty envelopes requires defining these distributions, and the resulting ensemble of predictions is directly dependent on these assumptions. Fig. 6 outlines how one might use such uncertainty envelopes in a Monte Carlo analysis by assuming a uniform distribution between the two limits. A uniform distribution for τ_r was sampled many times between these bounds, and the ensemble of predictions is shown in Fig. 6(G). Adopting a Bayesian approach to this problem, however, still allows the modeler to define, to their best knowledge, the shape and location of the parameter distributions (as in Fig. 6(C)) but the resulting predictions are based on the posterior distribution (Fig. 6(D)) and not on these prior distributions. Fig. 4 shows the difference between the prior distributions for both parameters and the updated posterior distributions. By collecting observations and incorporating them into the statistical framework, the prior knowledge (or ignorance) was updated to yield parameter estimates that produce the most realistic results. The results in Fig. 6(G) and (H) show that the more simplistic approach produces predictions with decreasing variance as discharge increases. The PPD in (H) has constant variance (in log space) at all discharges due to the specification of the likelihood. This difference is related to the fact that the forward stochastic approach in this comparison does not assume any variance structure of the observations—it assumes all variance is explained by the governing equation and variation in reference shear. The Bayesian approach incorporates a variance structure. In simple models, where only one or two parameters are involved, an argument could be made that a modeler could iteratively calibrate the parameter distributions in a forward stochastic model to produce a reasonable ensemble of predictions; but this becomes infeasible with many parameters or highly nonlinear systems such as sediment transport. To be clear, a forward stochastic approach is entirely capable of making the same predictions of a Bayesian model. Certainly, if the forward stochastic model had two parameters instead of one and the distributions of these parameters matched those of the posterior distribution of the Bayesian model then the results would be indistinguishable. In order for this to occur, however, the distributions for the parameters would need to be identical. The benefit of the Bayesian approach is that the stochastic prediction (formally, the posterior predictive distribution) and the parameter estimation (the posterior distribution) are developed in a single theoretical framework. It is conceivable that one could implement a type of non-linear regression/estimation for the model parameters for use in the forward stochastic model, but the results of non-linear regression require parameter normality, something not always guaranteed. Further, the estimate from a non-linear regression method is a fixed value and assumes normal (or t-distribution) errors. The Bayesian approach allows you to inform the model via priors, have them updated formally into the posterior distribution, and then make predictive distributions. Further, the MCMC algorithm is much more efficient in identifying the joint parameter space than a manual, iterative approach, especially because the estimates are distributions and not a point value. This difficulty only increases with dimensionality.

A further consideration that warrants the present estimation–prediction framework for sediment transport is that the threshold at which sediments begin to move is highly dependent on evolving bed configurations that are generally unknowable at the scale of large gravel-bed rivers such as the Snake (Kirchner et al., 1990; Eaton and Church, 2010). In a laboratory setting, the critical shear parameter is often well defined by its size and known bed characteristics (e.g., angle of repose; Dey, 1999). In the absence of this information, however, defining a grain shear required to produce movement is complicated by the unknown and evolving bed configurations. Thus the

Table 2
Inferred and fixed parameters for bedload transport function (top section displays results from the posterior distributions; middle section shows selected parameters from Erwin et al. (2011); bottom section shows channel properties for each site[a]).

Parameter	Snake River	Buffalo Fork	Pacific Creek
τ_r, Pa	9.38 (9.05–9.66)	9.67 (9.14–10.15)	14.30 (13.97–14.64)
τ_r^*	0.017 (0.016–0.018)	0.033 (0.031–0.035)	0.042 (0.041–0.043)
σ^2	1.90 (1.35–2.73)	1.25 (0.81–1.92)	0.62 (0.36–1.08)
τ_r, Pa[b]	9.80 (8.30–11.30)	9.50 (8.50–12.50)	14.20 (13.20–14.20)
τ_r^*[b]	0.018 (0.015–0.021)	0.032 (0.029–0.043)	0.042 (0.039–0.045)
Water surface slope, S	0.0025	0.0025	0.0035
D_{65}, mm	50	29	29
D_{50}, mm	34	18	21
Active channel width, m	70	45	43
Num. observations	62	39	24

[a] Point estimate is the expected value of the posterior distribution, and values in parentheses comprise the 95% posterior credible interval for the Bayesian estimates. Values without credible intervals are assumed fixed and known.
[b] Values as reported in Erwin et al. (2011).

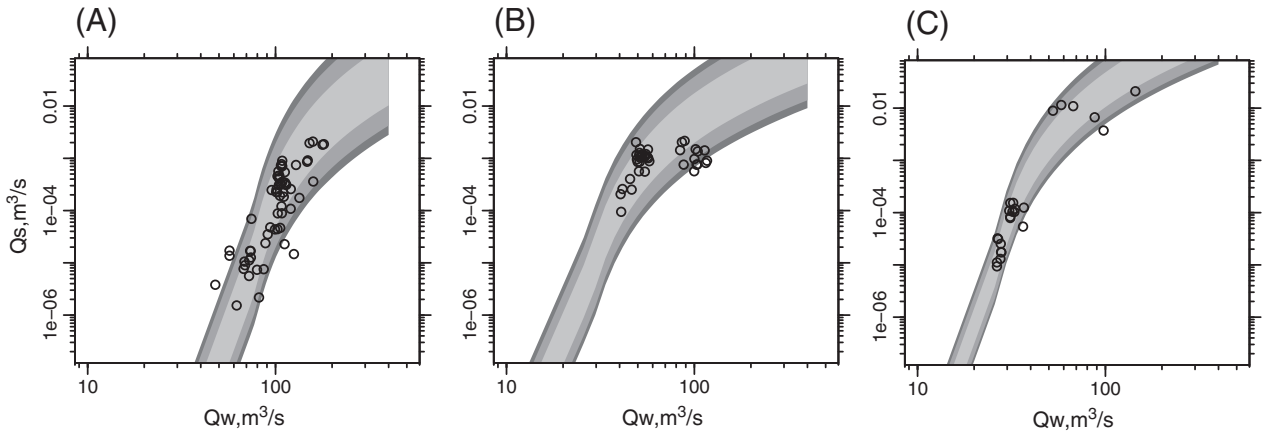

Fig. 5. Probabilistic sediment rating curves. (A) Snake River at Deadman's Bar, (B) Buffalo Fork, and (C) Pacific Creek. The shading (lightest to darkest) represents the 68%, 90%, and 95% credible intervals.

inferred reference (or critical) shear on large river systems averages over these unknowns and behaves less as a physically-meaningful parameter for the individual grains and more like a calibration parameter for the bed as an ensemble. As such, the inferred values are unlikely to be portable from one river reach to another.

5.2. Implications for sediment budgets

As described in Section 3.3, one source of uncertainty in sediment budgets is the uncertainty in sediment transport rating curves. The probabilistic rating curves described in the preceding sections form the basis on which the probabilistic sediment budget is constructed and allows us to robustly propagate uncertainty at different levels. The first level of uncertainty relates to prior knowledge of parameter values, and this is considered during the specification of the prior distributions. The next level pertains to the resulting uncertainty in the posterior distributions; this uncertainty is propagated through to the construction of the sediment rating curves. The final step in this paper is where uncertainty in the rating curves is propagated into yearly and cumulative sediment budget calculations.

Grams and Schmidt (2005) presented an example of developing a sediment budget for the Green River below Flaming Gorge dam and discussed the various uncertainties associated with this analysis. One of the conclusions of their paper is that "full consideration of the uncertainty in the sediment budget indicates that the budget is better described as indeterminate, which is not equivalent to an equilibrium condition" (Grams and Schmidt, 2005, p. 178). This conclusion was based on expert-defined error magnitudes that were applied to the measured annual loads (e.g., 5% and 10%) and underscores the reality that a simple description of system uncertainty may obscure the underlying behavior of a river system. A Bayesian approach to sediment budgets differs in that the uncertainty associated with sediment yields is estimated and not assumed. The estimation of sediment yield variability arises naturally from the specification of priors for the relevant parameters, as well as the observations that are collected to calibrate the model. At the very basic level, an expert assessment of uncertainty with respect to the parameters that govern the process can be stated to initiate the analysis (even if very approximately). After that, the uncertainties related to prediction and sediment budget calculations are not assumed because they are explicitly quantified in the posterior distributions.

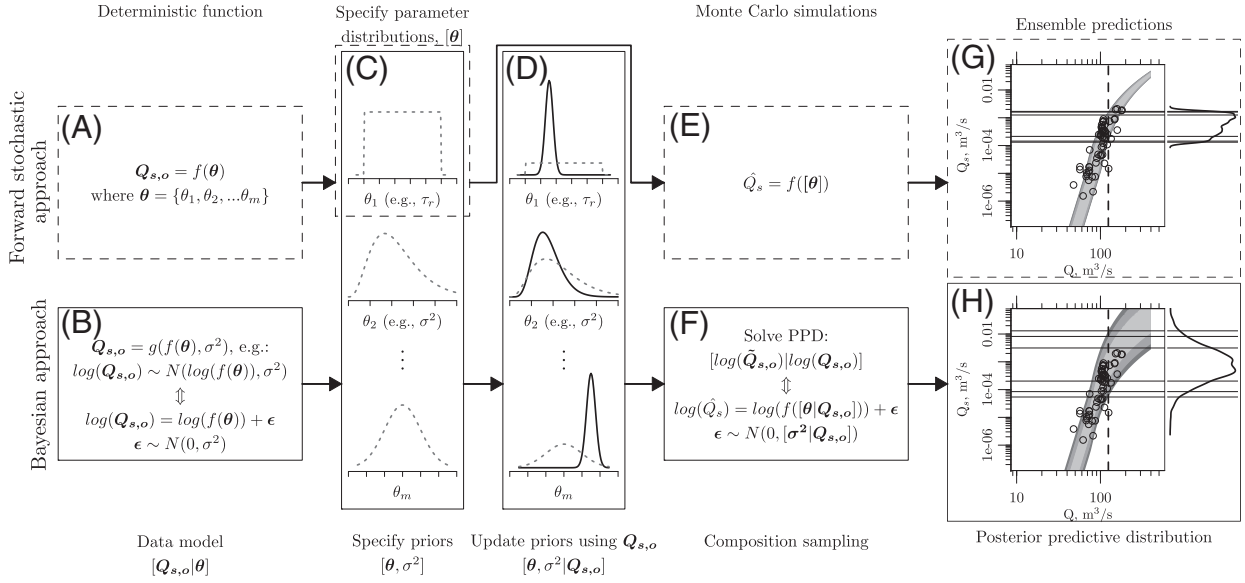

Fig. 6. Flowchart comparing workflows of a forward stochastic model and a Bayesian model for sediment transport. Forward stochastic model workflow goes from (A) to (C) to (E) to (G). The Bayesian model goes from (B) to (C) to (D) to (F) to (H). The ensemble predictions in (G) assume a uniform distribution of τ_r with upper and lower limits defined by Erwin et al. (2011) shown in Table 2 at Deadman's Bar. (H) PPD for Deadman's Bar. The density curves shown on the right of (G) and (H) are the distributions of sediment transport at the water discharges denoted by the dashed vertical line on the rating curves. The thin horizontal lines correspond to the 68%, 90%, and 95% prediction intervals.

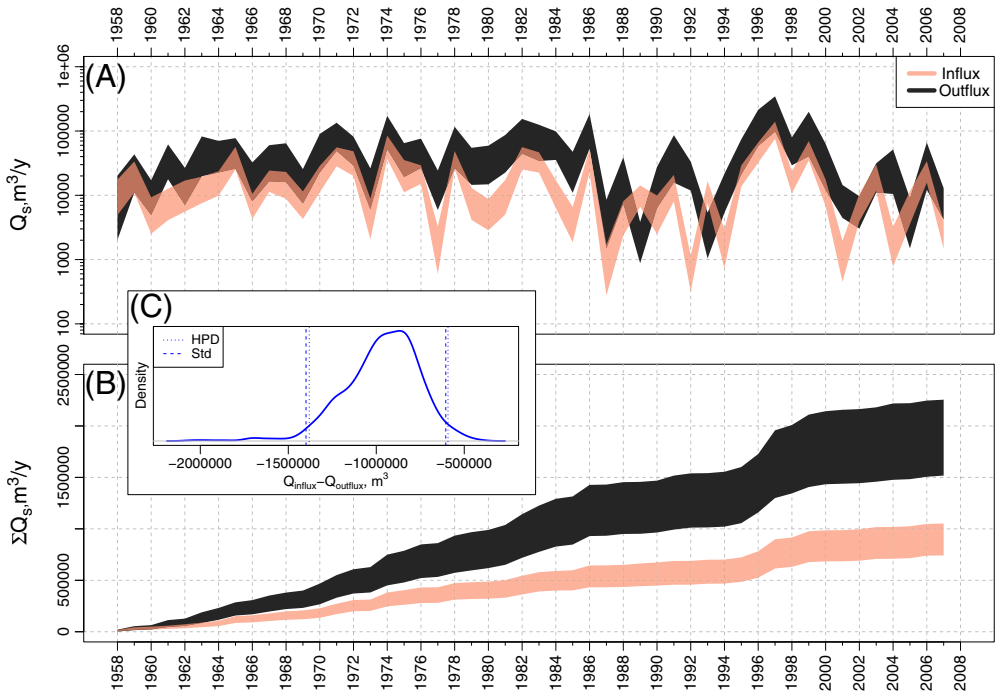

Fig. 7. Sediment budget using 95% credible interval. (A) Annual sediment yield. Shaded regions represent the 95% credible interval on the posterior predictive distribution of total annual sediment flux. A 'significant' difference between the influx and outflux occurs where the regions do not overlap. (B) Cumulative sediment yield starting in 1950. (C) Net change in sediment storage over period of record.

Another consideration when doing probabilistic sediment budgets is that the shape of the distributions will also have an effect on uncertainty. If error rates are assumed to be uniform (e.g., any value $\pm 5\%$ is equally probable) or heavy-tailed (i.e., more probability mass is located away from the center of the distribution) subsequent calculations will be less certain than if a normal distribution were assumed. When one assumes an error structure rather than estimating it, these distributions must be specified by the modeler and may influence the outcome. The approach promoted in this paper removes this burden from the modeler and allows the prior distributions and the observations to determine the form and bounds of these distributions as it naturally performs the tasks of parameter estimation and stochastic prediction.

Erwin et al. (2011) noted that the magnitude of their uncertainty in sediment flux calculations did not allow them to determine whether or not the reach is in sediment deficit or surplus. They noted, however, that the difference between influx and outflux was sufficiently large in 11 of the 50 years that a deficit could be inferred. A surplus

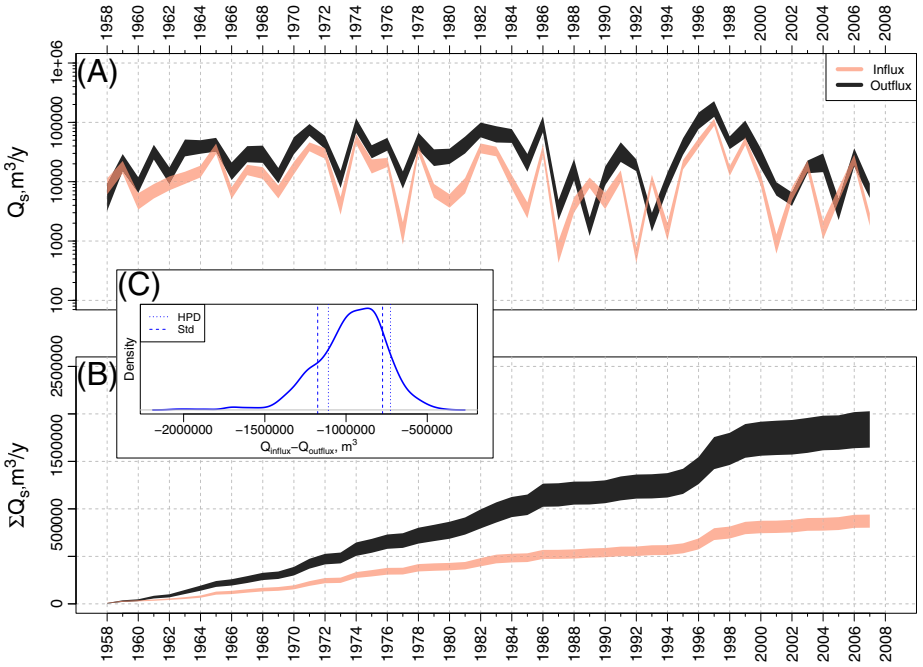

Fig. 8. Sediment budget using 68% credible interval. (A) Annual sediment yield. Shaded regions represent the 68% credible interval on the posterior predictive distribution of total annual sediment flux. A 'significant' difference between the influx and outflux occurs where the regions do not overlap. (B) Cumulative sediment yield starting in 1950. (C) Net change in sediment storage over period of record.

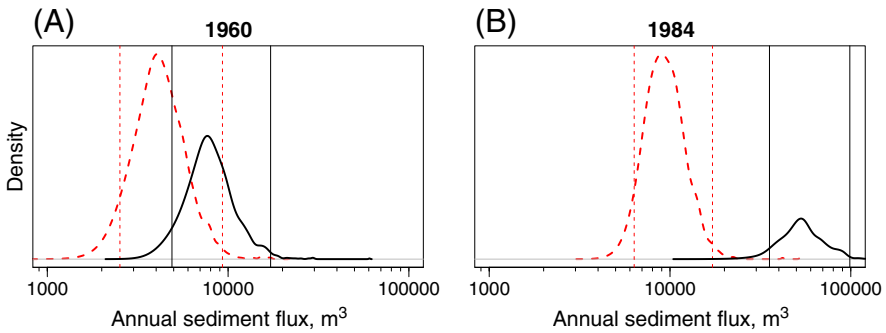

Fig. 9. Posterior predictive distributions of annual transport for 1960 and 1984. Dashed lines (red) denote sediment influx; solid lines (black) denote outflux. Vertical lines correspond to the 0.025 and 0.975 quantiles for each distribution.

could be demonstrated in 1 year. These are the results of manually defining an upper and lower reference shear that results in a region that contains 90% of the sediment transport observations. The results shown in Figs. 7 and 8 show that, at 95% credibility, the system is in deficit for 15 of the years and in surplus for two years. At 90% credibility (figure not shown), the number of years in sediment deficit increases to 20 and two years in surplus. The 68% credible interval of sediment yield indicates that 37 years are in deficit and two are in surplus. Constructing sediment budgets probabilistically in a Bayesian framework provides a robust way to propagate uncertainty and answer questions about how different influx and outflux need to be in order to be considered a real signal and not just noise. The ability to distinguish between variability attributable to noise and variability as a characteristic system behavior addresses the uncertainty challenge posed by Grams and Schmidt (2005).

Questions regarding the sediment budget can be asked and answered at a given tolerance for risk while explicitly accounting for variability and uncertainty. At the limit of taking the credible interval to 0%, the predictive distributions in Figs. 7 and 8 reduce to the purely deterministic budget calculations. At this level, every year is either in deficit or surplus. If, for a given year, the influx is 10,000 m^3 and the outflux is 10,005 m^3, we intuitively know that this is unlikely to be a significant difference. But at what difference does significance begin and at what significance level? We observe that the uncertainty in sediment rating curves is often the focus of analysis, yet this uncertainty is not always propagated through to estimates of annual yield (e.g., McLean et al., 1999a; Major, 2004; Vericat and Batalla, 2006). Without a robust method to quantify first the uncertainty of the sediment rating curve and its distribution and second how this uncertainty affects estimates of sediment yield, point estimates of sediment yield will always be distinct and left to intuition to determine what constitutes a material difference versus an immaterial difference. Fig. 9 shows the predictive distributions for sediment influx and outflux in the study reach for 1960 and 1984. This figure illustrates how calculating the quantiles of the predictive sediment yield distributions facilitates the distinction between significant and non significant differences. In 1960, the 95% credible interval on the influx and outflux overlap—thus one can conclude that at 95% credibility no significant difference is present. The values for 1984, however, illustrate that because the credible regions do not overlap the distance between the influx 0.975 quantile and the outflux 0.025 quantile is the sediment deficit that can be justifiably reported at the 95% credible level. These quantiles form the boundaries shown in Fig. 7(A) and the regions shown in Figs. 7 and 8 are not flat.

Over the study time period (1958–2007), we accumulated the sediment yields to develop a cumulative sediment yield distributions shown as (B) in Figs. 7 and 8. This cumulative distribution sums the preceding years' annual yield distributions (e.g., Fig. 9) to produce a new distribution. These annual loads are then accumulated from 1958 to 2007. The total difference in sediment influx and outflux, shown as (C) in the same figures, demonstrates that every computation on the sediment budget data is done as a distribution. Thus, we can say with 95% credibility that by 2007 the study reach was in sediment deficit by at least 600,000 m^3 and at most 1,400,000 m^3. These intervals can be reduced by relaxing the credibility requirements (compare the quantiles in Fig. 8(C) to Fig. 7(C)). The analysis provided in this paper confirms the conclusion of Erwin et al. (2011) that the study reach on the Snake River is generally in deficit (on a yearly basis) but goes further to provide a robust estimate of the accumulated sediment deficit over the period of record.

6. Conclusions

The main purposes of this paper were to demonstrate the applicability of the Bayesian sediment transport model developed in Schmelter et al. (2011) to a large gravel-bed river, provide increased visibility of the Bayesian approach because of its ability to model complex systems and to accommodate uncertainty in predictions and parameters, and evaluate the suitability of the Bayesian approach as the basis for constructing probabilistic sediment budgets.

The model results indicate that the Bayesian approach to sediment transport is readily extensible to large gravel-bed rivers and provides a robust method to incorporate expert knowledge, estimate model parameters, and define sediment transport predictions in terms of probabilities. This approach naturally incorporates deterministic models, thereby providing a physically-based approach to model systems that are generally predictable by governing relationships but have elements of randomness associated with them. The model described here is able to partition out variability attributable to parameter variability as well as stochasticity and measurement error. More simplistic approaches, such as the forward stochastic model described in Fig. 6, implicitly assume that all variability is because of parameter uncertainty and would be difficult to calibrate manually in a nonlinear system where an error structure was incorporated and or more model parameters were introduced. The ability of MCMC to search parameter spaces of complex and nonlinear systems provides an attractive option for quantifying parametric uncertainty of such systems.

The posterior predictive distributions produced from the Bayesian analysis form the foundation of the probabilistic sediment budgets described in this paper. Because the uncertainty associated with sediment transport for a given flow is robustly quantified and propagated into the PPD, the experts modeling the system need not assume error rates of annual sediment loads for a given location. This further removes the necessity for modelers to choose the type of distribution for the uncertainty (e.g., normal, uniform, etc.). This is particularly attractive because uniformly distributed error rates (e.g., $\pm 5\%$) may be unnecessarily conservative, potentially resulting in indeterminate sediment budgets. The Bayesian sediment budget shown in Figs. 7 and 8 indicates that either deficit or surplus can be reliably inferred for 17 years at the 95% credible interval. Erwin et al. (2011) estimated this number to be 13 at 90% confidence. Because the sediment yields

are probability distributions, a credibility interval of sediment flux can be determined for any given credibility level or tolerance for risk. Credible intervals for influx, efflux, sediment transport, sediment accumulation/evacuation are easily obtained at any risk profile by adjusting the quantiles appropriately. Further, the approach described in this paper makes it easy to calculate the cumulative sediment yield as a distribution in which uncertainties have been fully propagated from the initial model assumptions.

While the approach that we promote here has advantages over deterministic and some simple uncertainty estimation schemes, the Bayesian model itself is quite modest. This general approach offers significant opportunities for more complex models to be constructed for fluvial sediment transport that estimate more parameters, incorporate competing deterministic functions, use different variance structures, and evaluate the suitability of different submodels such as those used for the partitioning of grain shear from total shear. One weakness of the current model is that it assumes stationary (time-independent) sediment rating curves for each location—while certainly not ideal, this assumption was driven by necessity because no other records from which the sediment rating curves could be estimated for previous years were available. This is not a reality unique to the data of Erwin et al. (2011) but is reflective of many applied geomorphic studies at large. Provided, however, that adequate data to develop time-dependent sediment rating curves are available, this could be integrated into the sediment budget framework presented above. Future work could also focus on the uncertainty of other model unknowns, such as grain size distributions and variability of estimated water discharges, though this uncertainty may be extremely small relative to the others. The current paper discusses the use of Bayesian statistics for bedload transport, and we believe that this approach will also prove useful in other areas of the field such as suspended sediment transport, multi-fraction bed load, determination of the optimal deterministic model for a dataset, and general prediction and inference related to sediment transport phenomena.

Acknowledgments

This research was supported by the U.S. Bureau of Reclamation, the National Park Service (Cooperative Agreement number: H1200040001), the U.S. Geological Survey Northern Rocky Mountain Science Center, Bozeman, MT (Cooperative Agreement number: 05CRAG0036), and the U.S. National Science Foundation via the National Center for Earth-surface Dynamics (Agreement EAR-0120914). The authors wish to thank Mevin Hooten and David Stevens for their technical recommendations, Jürgen Symanzik for his constructive criticisms on the graphics, and Jack Schmidt for his helpful comments on an early version of this manuscript This manuscript was greatly improved through critical review and feedback from Richard Marston, Scott Wright, Andre Zimmerman, and two anonymous reviewers.

References

Andrews, E., 1986. Downstream effects of Flaming Gorge Reservoir on the Green River, Colorado and Utah. Bulletin of the Geological Society of America 97, 1012–1023.
Ashmore, P., Church, M., 1998. Sediment transport and river morphology: a paradigm for study. In: Kingman, P., Beschta, R., Komar, P., Bradley, J. (Eds.), Gravel-bed Rivers in the Environment. Water Resource Publications, LLC, Highlands Ranch, CO, pp. 115–148.
Bhattacharya, B., Price, R.K., Solomatine, D.P., 2007. Machine learning approach to modeling sediment transport. Journal of Hydraulic Engineering 133, 440.
Box, G.E.P., Tiao, G.C., 1992. Bayesian Inference in Statistical Analysis, Wiley class edition. John Wiley and Sons Ltd., New York.
Bunte, K., Abt, S., 2005. Effect of sampling time on measured gravel bed load transport rates in a coarse-bedded stream. Water Resources Research 41, W11405.
Casella, G., George, E.I., 1992. Explaining the Gibbs sampler. The American Statistician 46, 167.
Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. The American Statistician 49, 327.
Childers, D., 1999. Field comparisons of six pressure-difference bedload samplers in high-energy flow. U.S. Geological Survey Water Resources Investigations Report 92-4068: Technical Report.
Dey, S., 1999. Sediment threshold. Applied Mathematical Modeling 23, 399–417.
Diplas, P., Kuhnle, R., Gray, J., Glysson, D., Edwards, T., 2008. Sediment transport measurements. In: Garcia, M.H. (Ed.), Sedimentation Engineering: Processes, Measurements, Modeling, and Practice. American Society of Civil Engineers, Reston, pp. 307–353. chapter 5.
Dogan, E., Tripathi, S., Lyn, D.A., Govindaraju, R.S., 2009. From flumes to rivers: can sediment transport in natural alluvial channels be predicted from observations at the laboratory scale? Water Resources Research 45, W08433.
Dunne, T., Mertes, L., Meade, R., Richey, J., Forsberg, B., 1998. Exchanges of sediment between the flood plain and channel of the Amazon River in Brazil. Geological Society of America Bulletin 110, 450.
Eaton, B., 2001. Effects of large floods on sediment transport and reach morphology in the cobble-bed Sainte Marguerite River. Geomorphology 40, 291–309.
Eaton, B., Church, M., 2010. A rational sediment transport scaling relation based on dimensionless stream power. Earth Surface Processes and Landforms 36, 901–910.
Erwin, S., Schmidt, J., Nelson, N., 2011. Downstream effects of impounding a natural lake: the Snake River downstream from Jackson Lake dam, Wyoming, USA. Earth Surface Processes and Landforms 36, 1421–1434.
Gaeuman, D.A., 2003. Evaluation of in-channel gravel storage with morphology-based gravel budgets developed from planimetric data. Journal of Geophysical Research 108, 1–16.
Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85, 398.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis, second edition. Chapman & Hall, Boca Raton.
Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721–741.
Gessler, J., 1971. Beginning and ceasing of sediment motion. In: Shen, H.W. (Ed.), River Mechanics, volume 1. Hsieh Wen Shen, Fort Collins, CO. chapter 7.
Gomez, B., Church, M., 1989. An assessment of bed load sediment transport formulae for gravel bed rivers. Water Resources Research 25, 1161–1186.
Gomez, B., Phillips, J., 1999. Deterministic uncertainty in bed load transport. Journal of Hydraulic Engineering 125, 305.
Grams, P., Schmidt, J., 2005. Equilibrium or indeterminate? Where sediment budgets fail: sediment mass balance and adjustment of channel form, Green River downstream from Flaming Gorge dam, Utah and Colorado. Geomorphology 71, 156–181.
Grass, A.J.A., 1970. Initial instability of fine bed sand. Journal of the Hydraulics Division 96, 619–632.
Griffiths, G.A., 1982. Stochastic prediction in geomorphology using Bayesian inference models. Mathematical Geology 14, 65–75.
Hamamori, A., 1962. A theoretical investigation on the fluctuation of bedload transport. Technical Report. Delft Hydraulics Laboratory, Delft, The Netherlands.
Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.
Hazel, J.E., Topping, D.J., Schmidt, J.C., Kaplinski, M., 2006. Influence of a dam on fine-sediment storage in a canyon river. Journal of Geophysical Research 111, 1–16.
Hicks, D.M., Gomez, B., 2005. Sediment transport. In: Kondolf, G.M., Piegay, H. (Eds.), Tools in Fluvial Geomorphology. John Wiley & Sons Ltd., Chichester, p. 688. chapter 15.
Kirchner, J.W., Dietrich, W.E., Iseya, F., Ikeda, H., 1990. The variability of critical shear stress, friction angle, and grain protrusion in water-worked sediments. Sedimentology 37, 647–672.
Knighton, D., 1998. Fluvial Forms and Processes. Arnold, London.
Love, J., Reed, J., Pierce, K., 2003. Creation of the Teton Landscape: a Geologic Chronicle of Jackson Hole and the Teton Range. Grand Teton Natural History Association, Moose, WY.
Major, J.J., 2004. Posteruption suspended sediment transport at Mount St. Helens: decadal-scale relationships with landscape adjustments and river discharges. Journal of Geophysical Research 109, 1–22.
Marston, R.A., Mills, J.D., Wrazien, D.R., Bassett, B., Splinter, D.K., 2005. Effects of Jackson Lake Dam on the Snake River and its floodplain, Grand Teton National Park, Wyoming, USA. Geomorphology 71, 79–98.
McLean, D.G., Church, M., Tassone, B., 1999a. Sediment transport along lower Fraser River: 1. measurements and hydraulic computations. Water Resources Research 35, 2533.
McLean, S., Wolfe, S., Nelson, J., 1999b. Predicting boundary shear stress and sediment transport over bed forms. Journal of Hydraulic Engineering 125, 725.
Metropolis, N., Rosenbluth, A.W., Rosenblulth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. Journal of Chemical Physics 21, 1087–1093.
Nelson, N., 2007. Hydrology and geomorphology of the Snake River in Grand Teton National Park, Wyoming. M.S. thesis. Utah State University. Logan, UT.
Paintal, A.S., 1971. Concept of critical shear stress in loose boundary open channels. Journal of Hydraulic Research 9, 91–113.
Parker, G., 1979. Hydraulic geometry of active gravel rivers. Journal of the Hydraulics Division 105, 1185–1201.
Parker, G., 1990. Surface-based bedload transport relation for gravel rivers. Journal of Hydraulic Research 28, 417–436.
R Development Core Team, 2010. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Robert, C.P., 2007. The Bayesian Choice: from Decision-theoretic Foundations to Computational Implementation, second edition. Springer Science+Business Media, LLC, New York City.
Sasal, M., Kashyap, S., Rennie, C., Nistor, I., 2009. Artificial neural network for bedload estimation in alluvial rivers. Journal of Hydraulic Research 47, 223–232.

Schmelter, M.L., Hooten, M.B., Stevens, D.K., 2011. Bayesian sediment transport for unisize bed load. Water Resources Research 47, W11514.
Schmelter, M.L., Stevens, D.K., in review. Traditional and Bayesian statistical models in fluvial sediment transport. Submitted to Journal of Hydraulic Engineering.
Schmidt, J.C., 1999. Summary and synthesis of geomorphic studies conducted during the 1996 controlled flood in Grand Canyon. In: Webb, R., Schmidt, J., Marzolf, G., Valdez, R. (Eds.), The Controlled Flood in the Grand Canyon. : Geophysical Monograph, 110. AGU, Washington, D.C., pp. 329–342.
Singer, M., Dunne, T., 2004. Modeling decadal bed material sediment flux based on stochastic hydrology. Water Resources Research 40, W03302.
Tanner, M.A., 1996. Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions, third edition. Springer-Verlag, New York.
Topping, D.J., Rubin, D.M., Vierra, L.E., 2000. Colorado River sediment transport — 1. natural sediment supply limitation and the influence of Glen Canyon Dam. Water Resources Research 36, 515–542.
Turowski, J., 2010. Probability distributions of bed load transport rates: a new derivation and comparison with field data. Water Resources Research 46, W08501.
Vericat, D., Batalla, R.J., 2006. Sediment transport in a large impounded river: the lower Ebro, NE Iberian Peninsula. Geomorphology 79, 72–92.
Wallick, J., Anderson, S., Cannon, C., O' Connor, J.E., 2010. Channel change and bed-material transport in the lower Chetco River, Oregon: U.S. Geological Survey Scientific Investigations Report 20105065. Technical Report.
Wikle, C., Berliner, L., 2007. A Bayesian tutorial for data assimilation. Physica D: Nonlinear Phenomena 230, 1–16.
Wilcock, P.R., 2001. Toward a practical method for estimating sediment-transport rates in gravel-bed rivers. Earth Surface Processes and Landforms 26, 1395–1408.