Estimating bedload from suspended load and water discharge in sand bed rivers

Thomas Ashley[1], Brandon McElroy[1], Daniel Buscombe[2], Matt Kaplinski[2], and Paul Grams[3]

[1]University of Wyoming
[2]Northern Arizona University
[3]US Geological Survey

November 23, 2022

Abstract

Estimates of fluvial sediment discharge from in situ instruments are an important component of large-scale sediment budgets that track long-term geomorphic change. Suspended sediment load can be reliably estimated using acoustic or physical sampling techniques; however, bedload is difficult to measure directly and can consequently be one of the largest sources of uncertainty in estimates of total load. We propose a physically-informed predictive empirical model for bedload sand flux as a function of variables that are measured using existing acoustic or physical sampling techniques. This model depends on the assumption that concentration and grain size in suspension are in equilibrium with reach-averaged boundary conditions. Bayesian inference is used to fit model parameters to data from eight sand-bed rivers and to simulate bedload flux over the available gage record at one site on the Colorado River in Grand Canyon National Park. We find that the cumulative bedload flux during the nine year period from 2008 to 2016 was 5\% of the cumulative suspended sand load; however, instantaneous bedload flux ranged from as little as 1\% of instantaneous suspended sand load to as much as 75\% of instantaneous suspended sand load due to fluctuations in flow strength and sediment supply. Changes in bedload flux at a constant discharge are indicative of short-term sediment supply enrichment and depletion. Long-term average bedload flux cannot be expected to remain constant in the future as the river adjusts to changes in sediment runoff and the dam-regulated discharge regime.

Estimating Bedload From Suspended Load and Water Discharge in Sand Bed Rivers

T. C. Ashley[1]*, B. McElroy[1], D. Buscombe[2], P. E. Grams[3], and M. Kaplinski[2]

[1]Department of Geology and Geophysics, University of Wyoming, Laramie, WY
[2]School of Earth and Sustainability, Northern Arizona University, Geosciences Division, Flagstaff, AZ
[3]US Geological Survey, Southwest Biological Science Center, Grand Canyon Monitoring and Research Center, Flagstaff, AZ

Key Points:

- Bedload flux is predicted from variables that are measured at acoustic suspended sediment monitoring stations.
- Bayesian modeling extends the utility of this approach to a wide range of conditions and rivers.
- Predicted bedload flux provides an indicator of short-term sediment supply enrichment and depletion.

*1000 E. University Ave, Laramie, WY, USA

Corresponding author: Thomas Ashley, tashley3@uwyo.edu

Abstract

Estimates of fluvial sediment discharge from in situ instruments are an important component of large-scale sediment budgets that track long-term geomorphic change. Suspended sediment load can be reliably estimated using acoustic or physical sampling techniques; however, bedload is difficult to measure directly and can consequently be one of the largest sources of uncertainty in estimates of total load. We propose a physically-informed predictive empirical model for bedload sand flux as a function of variables that are measured using existing acoustic or physical sampling techniques. This model depends on the assumption that concentration and grain size in suspension are in equilibrium with reach-averaged boundary conditions. Bayesian inference is used to fit model parameters to data from eight sand-bed rivers and to simulate bedload flux over the available gage record at one site on the Colorado River in Grand Canyon National Park. We find that the cumulative bedload flux during the nine year period from 2008 to 2016 was 5% of the cumulative suspended sand load; however, instantaneous bedload flux ranged from as little as 1% of instantaneous suspended sand load to as much as 75% of instantaneous suspended sand load due to fluctuations in flow strength and sediment supply. Changes in bedload flux at a constant discharge are indicative of short-term sediment supply enrichment and depletion. Long-term average bedload flux cannot be expected to remain constant in the future as the river adjusts to changes in sediment runoff and the dam-regulated discharge regime.

1 Introduction

Estimates of fluvial sediment load provide an important tool for quantifying large-scale geomorphic change. In a wide range of environments, suspended sediment load can be accurately constrained using acoustic surrogates for sediment concentration [*Topping et al.*, 2004; *Topping & Wright*, 2016], enabling low-cost measurement of suspended load at high temporal resolutions over multi-year timescales [*Dean et al.*, 2016; *Grams et al.*, 2013, 2018]. However, acoustic estimates of flux depend on assumptions about the vertical concentration distribution that are reasonable if not strictly valid in the interior of the flow [*Gray & Gartner*, 2010] but that become increasingly dubious in the near-bed region. Bedload may vary significantly with respect to suspended sediment load due to changes in Rouse conditions [*van Rijn*, 1984].

Existing procedures for measuring bedload separately from suspended load in sand-bedded rivers [*Gray et al.*, 2010; *Holmes*, 2019] are incompatible with the goals and limitations of long-term monitoring. Direct physical sampling is costly and can be inaccurate in large rivers due to undersampling [*Pitlick*, 1988], and existing predictive bedload transport models that might be used in lieu of direct measurements (e.g. *Wong & Parker* [2006]) generally require, at minimum, an estimate of the skin friction component of bed shear stress, which in turn necessitates additional measurements or models each subject to their own logistical limitations and uncertainty. Sediment budgets therefore rely on simplified treatments of bedload flux that can introduce large persistent biases to estimates of total bed material load. For example, bedload is typically estimated either as a constant fraction of suspended load [*Grams et al.*, 2013], a power-law function of water discharge (i.e. a rating curve) [*Ellison et al.*, 2016], or ignored [*Wright et al.*, 2010]. This is problematic because bedload flux can be a substantial fraction of total load in suspension-dominated rivers, particularly at low flow conditions [*Turowski et al.*, 2010]; bedload flux can vary relative to suspended load due changes in suspension conditions, and it can vary with respect to a fixed water discharge due to changes in bed material composition and channel geometry [*Topping et al.*, 2000a,b].

The purpose of this paper is to provide a reliable means for estimating bedload flux in sand-bed rivers when suspended sediment information is available. The rationale behind our approach is that bedload and suspended load are mutually determined by the same causal boundary conditions at the reach-averaged scale. As a result, measured changes in concentration and grain size in suspension can be used to deduce changes in these boundary conditions and estimate bedload flux. This concept was first proposed by *Rubin & Topping* [2001] and underlies an empirical model that expresses bedload flux per unit channel width as a function of unit water discharge, suspended sand concentration, and suspended sand diameter.

Our primary goal is to estimate bedload flux from gage data and propagate uncertainty through estimates of cumulative load. This is accomplished using Bayesian inference, which provides a convenient framework for quantifying uncertainty in sediment transport parameters using numerical Markov chain Monte-Carlo (MCMC) methods [*Schmelter et al.*, 2011; *Schmelter & Stevens*, 2012; *Schmelter et al.*, 2015]. Moreover, Bayesian techniques implemented in the MCMC framework enable rigorous propagation of uncertainty through individual estimates of sediment load and time-integrated mass balance calculations [*Schmelter et al.*, 2012].

Our model can be applied in any sand-bedded river and does not require site-specific calibration. However, our analysis reveals that predictions may be biased on a site-specific basis such that greater predictive accuracy is achieved when the model is fit using only data from one site. This is particularly important when computing sediment budgets because error associated with model bias accumulates over time [*Grams et al.*, 2013]. Unfortunately, site-specific data are not always available; in order to meet the varying needs of different applications, we present three modeling approaches that utilize historical data from seven rivers reported by *Toffaleti* [1968] to varying degrees. The first approach involves pooling all data to estimate model parameters and is acceptable for obtaining single estimates of bedload flux at sites where no direct observations are available. The second approach utilizes only data from the site of interest, and is suitable when extensive site-specific data are available. The third approach involves a hierarchical modeling framework [*Gelman et al.*, 1995; *Christensen et al.*, 2011] that optimizes use of limited site-specific data by using sites with many observations to inform prediction at sites with relatively few observations. Application of all three approaches is demonstrated at one sediment monitoring station on the Colorado river. The statistical procedure presented here ultimately provides a convenient method for tracking changes in bedload flux driven by flow strength and sediment supply limitation over timescales ranging from days to years.

2 Colorado River sediment monitoring

On the Colorado River in Grand Canyon National Park, flux-based sediment budgets inform flow regulation protocols aimed at minimizing the downstream impact of Glen Canyon Dam. The primary management objective is the reversal of long-term depletion of alluvial sand deposits, especially emergent deposits known as eddy sand bars, through the use of controlled floods [*Topping et al.*, 2010; *Wright & Kaplinski*, 2011; *Grams et al.*, 2015]. However, the range of available management solutions is limited; this objective must be accomplished without compromising other economic [*Ingram et al.*, 1991] and ecological [*Minckley*, 1991] objectives. Designing such a protocol requires a detailed understanding of the dynamics of flow and sediment transport through the canyon.

In the dam-regulated Colorado River, the upstream sediment supply is completely independent from water discharge. Undammed tributaries comprise the only resupply of alluvial material to the post-dam river, while the hydrograph is determined by clear water releases from Lake Powell [*Andrews*, 1991; *Topping et al.*, 2000a; *Rubin et al.*, 2002]. Sediment supply and flow fluctuations cause complex morphodynamic interactions as the channel adjusts to accommodate pulses of sediment under the imposed discharge regime. Confinement by bedrock and bouldery debris fans also limits the extent to which flow can modify local slope and hydraulic geometry. As a result, antecedent sedimentary and morphological conditions are as important as water discharge in regulating instantaneous sediment transport [*Rubin & Topping*, 2001]. This condition, known as “supply limitation,” is common in natural rivers, but is particularly pronounced on the Colorado River and other dammed rivers

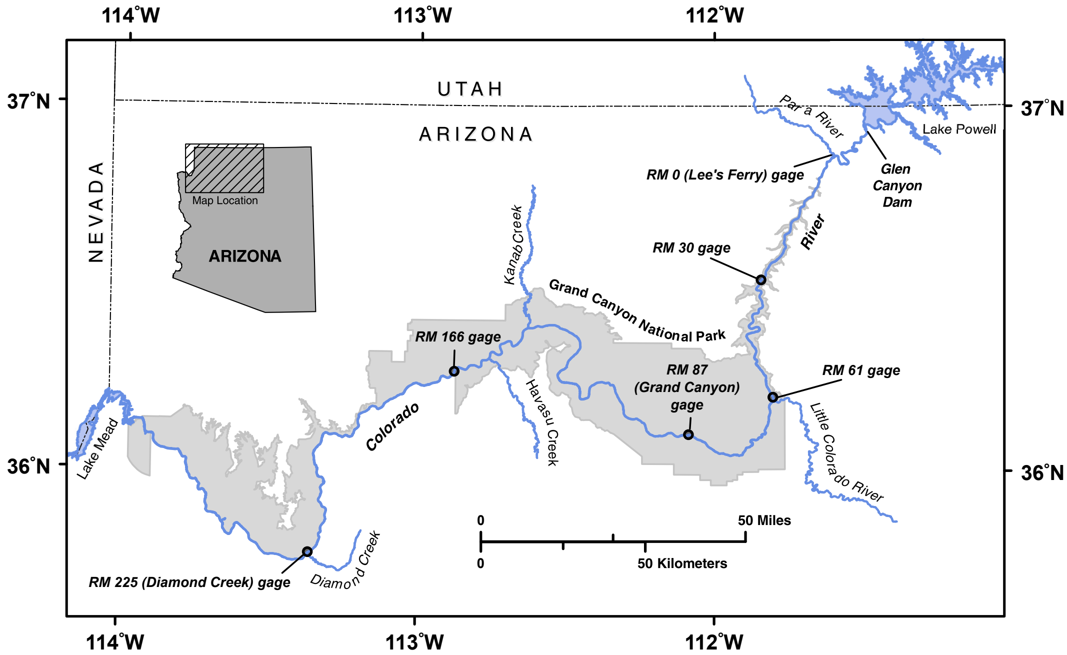


Figure 1. Map of the Colorado River in Grand Canyon National Park, after *Grams et al.* [2013]). Data used in this study come from the reach adjacent to the Diamond Creek gage located at river mile 225.

due to artificial flow regulation and sediment starvation [*Dolan et al.*, 1974; *Schmidt & Graf*, 1990].

Modeling the dynamics of alluvial deposits in supply-limited systems requires substantial physical simplifications and empiricism (e.g. *Wright et al.* [2010]). Changes in stored sediment mass estimated from spatial gradients in sediment flux are a useful metric for evaluating the effects of past flow regimes and for testing predictive models that can be used to determine best-practice scenarios for the future. The canyon is divided into five sediment budget reaches, each bounded by monitoring stations on the main stem and major tributaries that estimate total sand load every fifteen minutes (Figure 1). At the time of writing, these records comprise over a decade of almost uninterrupted suspended sediment data that can be used to quantify morphodynamic trends over a range of timescales: multi-year trends indicate regime-scale adjustment while short-term variability reflects the transient response to individual or seasonal perturbations in flow strength and sediment supply. Data are available online at https://www.gcmrc.gov/discharge_qw_sediment/.

Bedload flux is perhaps the largest source of uncertainty in estimates of total sediment load. At the time of writing, bedload is estimated at all monitoring sites on the Colorado River as a constant 5% of suspended load based on a single set of concurrent measurements of bedload and suspended load [*Rubin et al.*, 2001]. Presently, we aim test this assumption at one site (Figure 2), and reduce bias in estimates of total load by developing and applying a robust statistical methodology for estimating bedload flux from gage data.

3 Methods and data

3.1 Modeling approach

The goal of this paper is to predict total mass bedload flux, Q_b $[MT^{-1}]$, from measurements of water discharge, suspended sand concentration, and suspended sand diameter. To this end, we adopt an empirical power-law equation for bedload flux per unit width

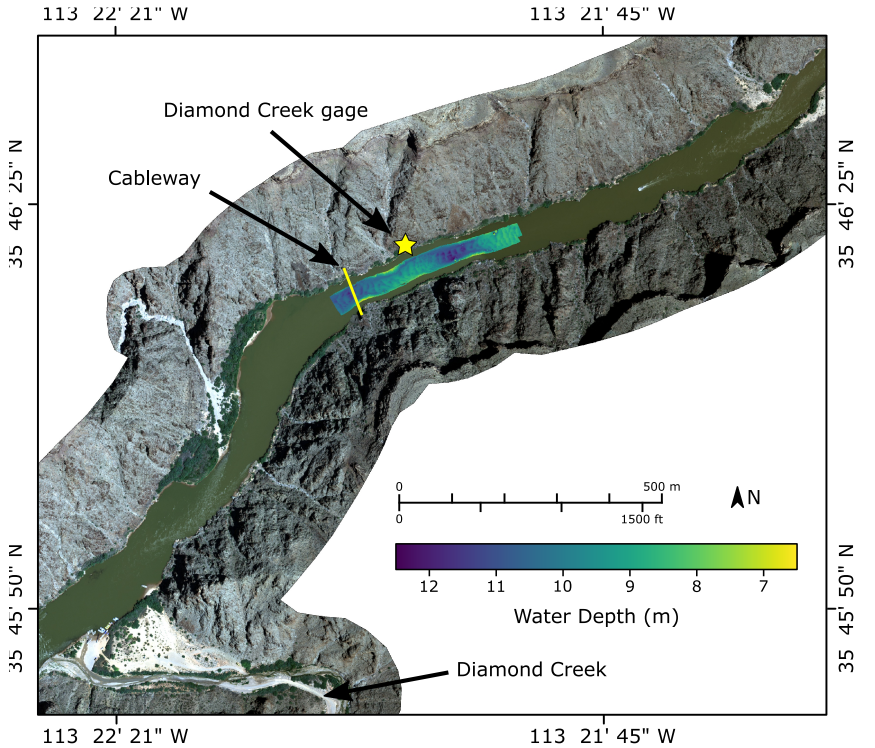


Figure 2. Aerial view of the Diamond Creek study site. One survey of water depth is plotted, illustrating the extent of the sonar mapping area.

q_b $[MT^{-1}L^{-1}]$ given by:

$$q_b = Ae^{\beta_0} q_w^{\beta_1} C_s^{\beta_2} D_s^{\beta_3}. \tag{1}$$

Here, q_w $[L^2/T]$ is the average volumetric water discharge per unit width equal to Q_w/W, where Q_w $[L^3/T]$ is the total volumetric water discharge and W $[L]$ is the surface width of the channel. C_s $[L/L]$ is the discharge-averaged suspended sand concentration, D_s $[L]$ is median diameter of suspended sand and A is a dimensional coefficient expressed in terms of fixed reference values for each variable (denoted by the subscript 0) as $A = q_{b0}/q_{w0}^{\beta_1} C_{s0}^{\beta_2} D_{s0}^{\beta_3}$. Finally, β_0 is an intercept term that is equal to 0 if reference values are chosen so that $q_b = q_{b0}$ when $q_w = q_{w0}$, $C_s = C_{s0}$, and $D_s = D_{s0}$.

Equation (1) is purely empirical; however, we consider the form of this expression in the context of existing theory to (1) facilitate qualitative interpretation of our results and (2) support the notion that in-sample fit will extend to out-of-sample predictive accuracy. Forward models for equilibrium sediment transport [*Einstein*, 1950; *McLean*, 1992; *Molinas & Wu*, 2002; *Wright & Parker*, 2004] encompass the physical interactions that are relevant to this objective, and generally involve several computational steps that incorporate various physical and empirical expressions. As an example *Wright & Parker* [2004] proposed a computational procedure for estimating C_s, D_s, and the Shields' stress due to skin friction τ_{*s} (among other variables) from specified reach-average boundary conditions, which are q_w, slope S $[L/L]$, and bed material grain size D_b $[L]$. Bedload flux can be computed from τ_{*s} using an empirical bedload transport formula [e.g. *Wong & Parker*, 2006]. Additional relevant physical parameters that must be specified are often assumed to be constants. These are gravitational acceleration g $[L/T^2]$, the kinematic viscosity of water ν $[L^2/T]$, and the densities of sediment ρ_s $[M/L^3]$ and water ρ_w $[M/L^3]$. In summary, this model approximates three unknown physical equations of the following functional form:

$$q_b = f(q_w, S, D_b, \rho_s, \rho_w, g, \nu) \tag{2}$$
$$C_s = f(q_w, S, D_b, \rho_s, \rho_w, g, \nu) \tag{3}$$
$$D_s = f(q_w, S, D_b, \rho_s, \rho_w, g, \nu). \tag{4}$$

Each forward equation has eight variables (seven predictor variables and one response variable) and three physical dimensions, and can therefore be reduced to five dimensionless variables (four predictor variables and one response variable) according to the Buckingham Pi theorem [*Gibbings*, 2011]. However, four of the eight physical variables are usually assumed to be constant so one of these dimensionless variables will always be a constant or a linear combination of other variables. Assuming power-law forward equations between dimensionless variables, we assert that any choice of dimensionless variables can be rearranged to obtain the following dimensional equations:

$$q_b = \gamma_1\ q_w^{\alpha_{11}} S^{\alpha_{12}} D_b^{\alpha_{13}} \tag{5}$$
$$C_s = \gamma_2\ q_w^{\alpha_{21}} S^{\alpha_{22}} D_b^{\alpha_{23}} \tag{6}$$
$$D_s = \gamma_3\ q_w^{\alpha_{31}} S^{\alpha_{32}} D_b^{\alpha_{33}} \tag{7}$$

where γ_1, γ_2, and γ_3 are fixed dimensional coefficients that can be expressed in terms of g, ν, ρ_s, and ρ_w. This system of equations can then be solved to obtain equation (1), noting that β_i exponents are simply algebraic combinations of α_{ij} exponents.

Based on these arguments, we offer the following interpretation of equation (1), leaving further discussion to Section 5.2. We assume changes in fluvial sediment transport conditions are driven by changes in q_w, S, and D_b. By measuring one of these variables (q_w) and two variables that directly respond to changes in these variables (C_s and D_s), it is possible to constrain the state of the transport system and predict unknown variables including bedload flux. In this manner, C_s and D_s are viewed as proxies for S and D_b.

As an aside, equation (1) can also be derived by combining simplified relations presented in the canonical sediment transport literature [e.g. *Wong & Parker*, 2006; *Engelund & Hansen*, 1967; *Brownlie*, 1983; *Garcia & Parker*, 1991]; however, many of these relations have empirical origins and thus contain large, unquantifiable uncertainty. Rather than combining a series of existing empirical expressions, we fit β_i parameters and quantify predictive uncertainty directly; this approach minimizes predictive bias assuming that available data sufficiently capture the underlying physical processes.

The majority of this paper focuses on the development and application of a statistical methodology used to estimate empirical scaling parameters in equation (1) and predict bedload flux. We present an example application at our field site on the Colorado River in Grand Canyon National Park, where estimates of bedload flux obtained from repeat bathymetric surveys of dune migration paired with concurrent gage measurements form the observational basis for statistical analyses. Parameter estimation and prediction is conducted using Bayesian inference which facilitates consistent propagation of uncertainty from multiple sources of information and prediction of distributions for quantities of interest [*Schmelter et al.*, 2011; *Schmelter & Stevens*, 2012; *Schmelter et al.*, 2015]. This approach is particularly useful for propagating uncertainty arising from both measurement uncertainty and parameter estimation uncertainty in calculations of cumulative sediment load [*Schmelter et al.*, 2012].

In addition to the data from our site, we also consider data from six other rivers reported by *Toffaleti* [1968] in order to test generality and improve the predictive power of our model. These data cover a much wider range of discharge, slope, and bed grain size conditions than those that are found at the site on the Colorado River. In order to incorporate these data into the predictive model for bedload flux at our site, we consider three statistical models that are distinguished in principle by their assumptions regarding the universality of scaling exponents and in practice by their treatment of groups in the data. These approaches have advantages and disadvantages to each other relative to the specific modeling conditions and objectives, as well as the quantity and quality of data that are available at a site of interest.

3.2 Statistical methods

3.2.1 Bayesian linear regression

The generalized linear model given by equation (1) has four unknown parameters that must be estimated from a large number of observations of model variables. This system is overdetermined and no single solution can fit all of the data simultaneously. As a result, it is necessary to employ regression analysis to handle uncertainty and error. Log-transformed variables enable linear regression, which assumes that the i^{th} observation of the response variable $\log(q_b)_i$ can be expressed as a linear function of the predictor variables $\log(Q)_i$, $\log(C_s)_i$ and $\log(D_s)_i$, plus an error term ϵ_i

$$log(q_b)_i = log(A)_i + \beta_0 + \beta_1 log(q_w)_i + \beta_2 log(C_s)_i + \beta_3 log(D_s)_i + \epsilon_i \tag{8}$$

Perhaps the most common variant of linear regression is Ordinary Least-Squares (OLS), which finds the combination of model parameters β_0, β_1, β_2, and β_3 that minimizes the sum of the sum of the squared error terms. OLS regression leads to an unbiased predictor of the response variable assuming ϵ_i is normally distributed and independent across all samples. However, for the purposes of the present research, this approach has several limitations. OLS regression cannot handle hierarchical organizations of data that potentially violate the assumed independence of ϵ_i, such as when individual observations are grouped by river or site. Additionally, analytical quantification of predictive uncertainty in the OLS framework does not readily allow for propagation of errors through mass-balance calculations.

Bayesian inference provides a convenient framework for overcoming these issues. For a general discussion of Bayesian methods, see *Gelman et al.* [1995]; *Christensen et al.* [2011]. The standard Bayesian approach to linear regression starts with the same assumptions as OLS that are encapsulated by (8). However, we introduce an additional parameter σ that quantifies the standard deviation of the error term, i.e.:

$$\epsilon_i \sim \mathcal{N}(0, \sigma) \tag{9}$$

where the tilde means "distributed as" and $\mathcal{N}(0, \sigma)$ is an independent normal distribution with zero mean and standard deviation σ. Consequently, we aim to draw inference on five parameters: β_0, β_1, β_2, β_3, and σ.

At this point we note for clarity that the term "variables" refers to measurable physical quantities, while the term "parameters" refers to unknown quantities that appear in the data model and are the object of statistical inference. Henceforth, we use θ to refer to the 5×1 vector of model parameters, i.e. $\theta = [\beta_0, \beta_1, \beta_2, \beta_3, \sigma]$. Additionally, we use $\boldsymbol{X}$ to refer to the $4 \times N$ matrix of N observations of model variables q_w, C_s, D_s, and q_b.

While OLS regression seeks estimates of model parameters that minimize the global sum of the squared residuals, Bayesian model fitting embraces uncertainty associated with the fact that small differences in model parameters may fit the data nearly as well as the optimal result. These small differences are quantified by the likelihood function, which exists on the domain of model parameters assuming fixed observational data $\boldsymbol{X}$, and is denoted by $L(\theta|\boldsymbol{X})$. Here, the vertical line denotes conditional dependence, i.e. the likelihood of θ given $\boldsymbol{X}$. The likelihood can be computed for any combination of parameters, where higher likelihoods represent more likely combinations of parameters. Introducing the prior probability distribution $P(\theta)$, we obtain an expression for the posterior probability distribution of model parameters conditional on observational data $P(\theta|\boldsymbol{X})$ through Bayes theorem:

$$P(\theta|\boldsymbol{X}) = \frac{L(\theta|\boldsymbol{X})P(\theta)}{\int L(\theta|\boldsymbol{X})P(\theta)d\theta} \tag{10}$$

Once the posterior probability distribution of model parameters is known, unobserved values of q_b can be estimated from measured values of predictor variables using Bayesian posterior predictive distributions, which efficiently propagate uncertainty through individual estimates of q_b as well as time-integrated mass-balance calculations.

3.2.2 Grouped, ungrouped, and hierarchical model variations

The basis for equation (1) suggests that it is sufficient to predict bedload flux in any sand bed river using a single universal set of scaling parameters. However, some degree of predictive uncertainty is inevitable owing to both measurement error and model bias arising from simplification of physical processes. While measurement error can be considered uncorrelated, systematic biases are caused by a failure of the data model to capture specific physical processes, and are thus likely to be correlated when conditions are similar. As a result, we anticipate persistent site-specific biases using a general model based on data from many rivers. For example, details of channel geometry not explained by width and slope may cause bedload flux to be more or less sensitive to changes in water discharge at one site compared with the central tendency of all sand bed rivers. In this case, better predictive accuracy would be achieved at that site by adjusting the value of β_1 to reflect this difference. In general, we anticipate better predictive performance if model parameters are constrained on a site-specific basis.

This theoretical consideration is at odds with practical limitations: regression analysis requires numerous independent estimates of bedload flux that are expensive and difficult to obtain. Thus, it would be advantageous if existing data from many rivers could be used to help inform bedload prediction at a new site. Optimal model parameters may differ slightly from site to site; however, sand-bed rivers are all governed by the same general physical processes such that it is reasonable to expect that scaling parameters should be similar between rivers. In order to balance theoretical and practical concerns, we consider three distinct generative data models, each of which reflects a different trade-off between observational data requirements and assumptions regarding the generality of scaling parameters.

The first model (the grouped model, Appendix B.1) assumes a single universal set of model parameters $\theta = [\beta_0, \beta_1, \beta_2, \beta_3, \sigma]$. The standard deviation of the error term σ is the same for all data. All observations are therefore treated as independent observations from the same exchangeable group of observations. The advantage of this model is that it can be applied at a new site without collecting any additional data. However, it ignores the possibility of correlated errors by river or site, and is therefore subject to unquantifiable systematic biases when applied at a specific site without local data.

The second model (the ungrouped model, Appendix B.2) assigns different independent scaling parameters $\theta_j = [\beta_{j0}, \beta_{j1}, \beta_{j2}, \beta_{j3}, \sigma_j]$, for $j = 1, ..., m$ and $m = 8$ is the number of data groups (i.e. independent sites). This is equivalent to performing grouped regression independently on a site-specific basis: each site is treated as an independent statistical entity comprising its own exchangeable group of observations. This model is perhaps the most theoretically conservative in that it assumes nothing with regard to physical similarity between sites. However, it is also the least practical in that it requires extensive uncorrelated observations of bedload flux from each monitoring site in order to ensure reliable results, and cannot be applied at a site where bedload has never been measured directly.

The third model (the hierarchical model, Appendix B.3) assigns different regression coefficients to each site, but assumes some degree of physical similarity between sites. Observations are treated as exchangeable on a site-specific basis, and each site comes from an exchangeable group of sites, that is, all sand bed rivers. We aim to draw inference, not only on the behavior of individual sites, but also on the distribution of behaviors that can be observed at different sites. Site-specific coefficients are thus determined partly by data collected at that site, but are also informed by the behavior of other rivers which can reduce issues related to low sample size at one site if sufficient data exists at other sites. Hierarchical organization is implemented through priors for the regression coefficients which are assumed to be normally distributed with a mean and variance that reflects the central tendency and variability of sand-bed rivers. This model lies somewhere between the grouped and ungrouped models in terms of both theoretical assumptions and data requirements. Some data are useful in order to constrain bedload flux at a new site, but limited observations are utilized to greater effect than in the ungrouped model.

3.2.3 Priors

Diffuse (i.e. wide, minimally informative) priors are commonly used to minimize influence on model results, and are employed here for all three model variations. Diffuse priors are effectively constant over the relevant parameter domain, which means that the posterior distribution is essentially reflects a renormalization of the likelihood function, preserving the relative log-likelihoods while ensuring the posterior integrates to 1. Due to the relatively large sample size, our results are not sensitive to the specific choice of diffuse prior.

Grouped and ungrouped regression models were fit using an approximation for Jeffrey's prior, which is an attractive choice due its unique theoretical properties [*Gelman et al.*, 1995; *Christensen et al.*, 2011]]. Jeffrey's prior is a uniform distribution on the domain $(\beta_0, \beta_1, \beta_2, \beta_3, log(\sigma))$, which is an improper prior because it does not integrate to 1. Thus, normal distributions centered on zero with large standard deviations are used to approximate Jeffrey's prior because a normal distribution approaches a uniform distribution as the standard deviation goes to infinity. Jeffrey's prior is also uniform $log(\sigma)$ meaning the prior probability that the parameter is between 0.01 and 0.1 is the same as the probability the parameter is between 0.1 and 1. The inverse gamma distribution approaches a uniform distribution on $log(\sigma)$ as its parameters go to zero.

The hierarchical model structure is implemented through informative, dynamic priors, where the parameters for these priors are referred to as "hyperparameters". Inference is drawn on parameters and hyperparameters simultaneously such that the hyperparameters have their own prior and posterior probability distributions. Priors for hyperparameters, or "hyperpriors" must be specified. Again, we utilized diffuse, minimally-informative hyperpriors, the specific choice of which does not influence model results. For additional details on priors and hyperpriors, see Appendices B.1, B.2, and B.3.

3.2.4 Model fitting

All three models were fit using Markov Chain Monte-Carlo (MCMC) sampling methods. This technique is commonly used to sample the posterior distribution and conduct predictive simulation when analytical alternatives are cumbersome or impossible. For additional details on MCMC sampling, see Appendix B.4 and example workflows [*Ashley*, 2019b].

3.2.5 Model selection

Quantitative comparison of predictive power is accomplished using the Deviance Information Criterion (DIC, *Spiegelhalter et al.* [2002]; Gelman et al. [2014]), Appendix B.6), a generalization of the Akaike Information Criterion that is suitable for comparing the hierarchical and non-hierarchical models used here. DIC includes two two terms: one which quantifies in-sample predictive accuracy and one which corrects for model complexity to approximate out-of-sample predictive accuracy under certain assumptions [Gelman et al., 2014]. As a relative measure of predictive power, models with lower DIC are expected to have lower prediction error than models with higher DIC. However, DIC is not a perfect measure of relative prediction error and is reported here (Table 2) to inform model evaluation rather than as the sole discriminatory factor.

3.3 Field methods

Transport-related data were collected at one field site on the Colorado River in Grand Canyon National Park during three field campaigns in the Spring and Summer of 2015, as well as the Fall of 2016. The site (Figure 2) is located at river mile 225 in the vicinity of USGS monitoring station 09404200 (Colorado River above Diamond Creek near Peach Springs, AZ). Hereafter, we refer to this site informally as "Diamond Creek" or the "Diamond Creek field site". Data include repeat bathymetric surveys of dune migration, ADCP surveys of flow velocity, suspended sediment and bed sediment samples, and bed photographs for optical grain-size analysis [*Buscombe et al.*, 2010]. Concurrent gage measurements of water discharge, suspended sand concentration, and grain size were also collected following standard procedures during this time [*Rantz et al.*, 1982; *Topping & Wright*, 2016].

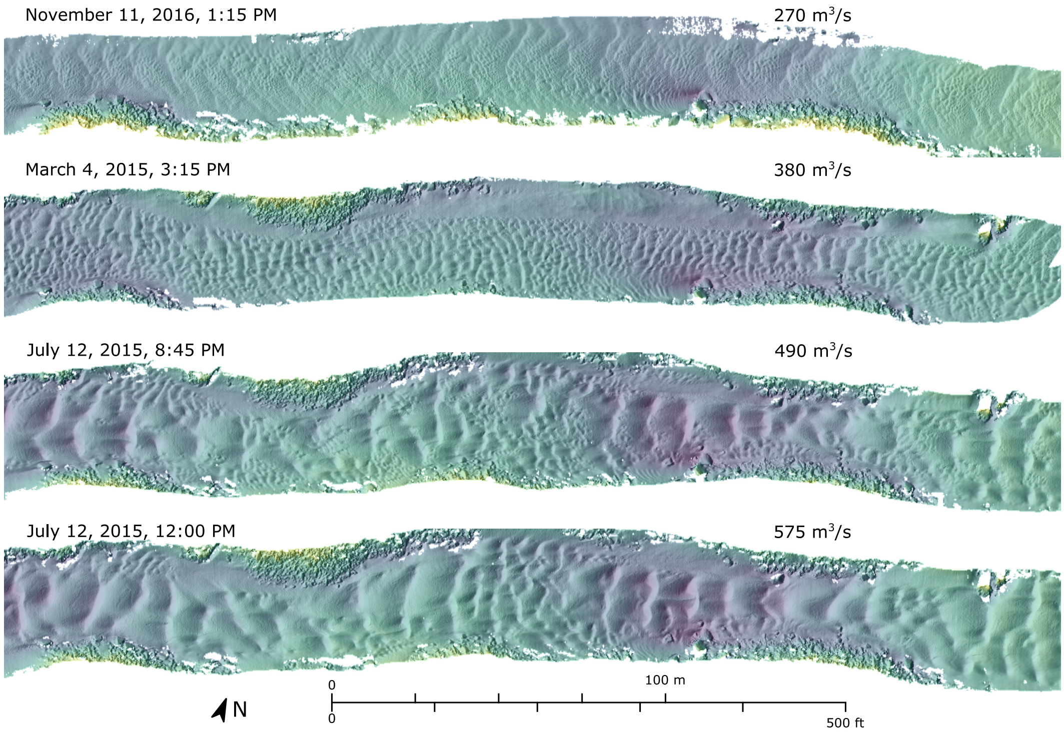
Estimates of bedload flux were obtained using 320 high resolution, full-width bathymetric surveys of an approximately 400 meter reach adjacent to the Diamond Creek gaging station. Surveys were collected using a 400 kHz Reson 7125 multibeam echo sounder (MBES) which produces a swath comprised of 512 beams (each 1 x 0.5 degrees) across a transverse subtended angle of 135 degrees. In order to map sonar returns onto a global coordinate system, the location of the boat was tracked using a robotic Total Station referenced to a fixed position on the bank, and a fiber-optic gyrocompass and inertial sensors were used to calculate heading, roll, and pitch of the sonar head. Patch tests were conducted before the surveys to determine the offset angles and timing latency between the various system components. Bad soundings and sweep misalignments (due to, for example, systematic side-lobe interference; and scattering of soundings by air bubbles, drifting insects and other organic matter in the water) were identified by manual sweep editing and systematically stepping through overlapping sweeps. Quality assurance assessments were performed after the surveys by comparing selected soundings from all surveys over a large, flat-topped rock located along the channel margin. The mean standard deviation of soundings over this feature was 0.015 m and indicate a high level of survey precision. The final, edited surveys used here are ungridded point clouds, where each point corresponds to a valid sonar return from the river bed. More details about acquisition of MBES data with this instrument and configuration are found in *Kaplinski et al.* [2009]; *Kaplinski et al.* [2014], *Grams et al.* [2013, 2018], and *Buscombe et al.* [2014a,b]. Four example surveys are plotted in Figure 3.

Simons et al. [1965] provide the method by which bathymetric data can be used to generate bedload flux estimates. Their expression is given by:

$$q_b = (1-p)V_c\frac{H_c}{2} + C, \tag{11}$$

where q_b $[L^2T^{-1}]$ is the volumetric bedload flux per unit width, p $[-]$ is the bed porosity taken to be a constant 0.35, V_c $[LT^{-1}]$ is a characteristic bedform migration rate, $H_c[L]$ is a characteristic bedform height, and C is a constant of integration assumed to be zero. Measured bedform heights ranged from 0.15 to 0.70 m, and measured migration rates ranged from 0.21 to 1.76 m/hr. Both of these quantities varied predictably with water discharge.

Equation (11) is derived from a statement of mass conservation (the Exner equation, *Paola & Voller* [2005]) combined with a simplified model for dune evolution characterized by translationally invariant migration of triangular or sinusoidal forms. Although it represents substantial simplifications of physical process (for example, by ignoring bedform deformation and variability in bedform migration rate and geometry), flume and field studies find good agreement between (11) and other estimates of bedload flux across a wide range of conditions extending from the threshold of bedform development to suspension-dominated dunes [*Simons et al.*, 1965; *Engel & Lau*, 1980; *van den Berg*, 1987; *Mohrig & Smith*, 1996]. Consequently, we argue that this expression provides a reasonable estimate of bedload transport that is not captured by acoustic estimates of suspended sand load. Equation (11) was used to compute 55 hourly estimates of average bedload flux (Figure 4). Major elements of



381 **Figure 3.** Example bathymetric surveys with shaded relief plotted at 10 cm resolution. Water discharge during the survey is indicated in the upper right corner of each survey. Flow is from right to left. Colors represent water depth, as in figure 2.

this procedure are discussed in Appendix A. Additional details can be found in the documentation of software developed for this purpose [*Ashley*, 2019a]

3.4 Additional data from other rivers

The large river dataset presented by *Toffaleti* [1968] (and derived quantities) is used to supplement limited data from our field site. This dataset comprises a total of 262 concurrent observations of bedload flux Q_b, water discharge Q_w, suspended sand concentration C_s, median suspended sand diameter D_s, and channel width W on the Atchafalaya River ($n = 60$), the Mississippi River in Louisiana ($n = 47$), the Mississippi River in Missouri ($n = 63$), the Red River ($n = 28$), the Rio Grand River ($n = 36$), the Middle Loup River ($n = 9$), and the Niobrara River ($n = 19$). These sites are similar to the Diamond Creek field site in that the predominant bed material is sand; however they are different in that they are all alluvial rivers (whereas the Colorado River in Grand Canyon is a bedrock-confined alluvial river with gravelly and sandy reaches). Our model is based on physical theory describing one-dimensional transport, and assumes nothing about channel form. Consequently, it can be applied in rivers that are not fully alluvial as long as the bed material at the site of interest is sand.

Total suspended sand concentration C_s and median suspended sand grain size D_s were computed from reported grain-size specific suspended sediment concentrations. Bedload flux was computed according to the revised Meyer-Peter & Müller bedload equation [*Wong & Parker*, 2006] with grain stresses estimated using the Einstein drag partition as reformulated by *Garcia* [2008]. This procedure was also used to compute bedload flux at our study site when flow velocity and bed sediment data are available to check approximate correspon-

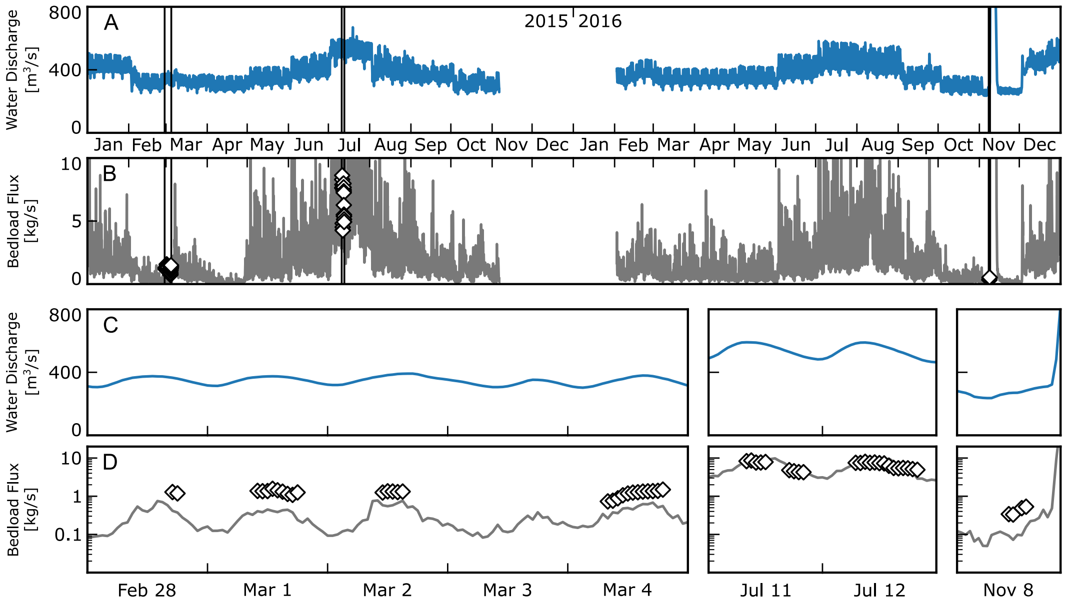


Figure 4. Time series plot of water discharge (A) and bedload flux (B) at the Diamond Creek sediment monitoring station in 20015 and 2016. Grey line shows bedload flux estimated as a constant fraction (5%) of suspended load, and black diamond show hourly average estimates of bedload flux from bedform migration. Insets (C, D) highlight the periods where bedform flux estimates are available.

dence with estimates of flux from dune migration. Note that here, and throughout, "observations" is used as part of the statistical vernacular to refer to independent samples of variables and implies nothing about how those samples were obtained. This distinction is particularly important here because "observations" of bedload flux are actually computed from depth, slope, grain size, and flow velocity using physically-based model. Similarly, observations of bedload flux at Diamond Creek are computed using a physically-based model from dune height and velocity.

3.5 Data treatment

The statistical methods employed here assume errors in observations are uncorrelated. However, the 55 hourly estimates of average bedload flux from the Diamond Creek field site were collected over seven days during which temporal correlation is likely. Unqualified extrapolation of trends in this dataset to the full gage record spanning nearly ten years may therefore produce unrealistic results. In order to mitigate this effect, we use only the first and last measurement from each day ($n = 14$) in order to estimate model parameters.

The full data set used for statistical analysis comprises a total of 276 observations from eight sites. Data were log-transformed to obtain the linear regression variables q_w^*, C_s^*, D_s^*, and q_b^* using fixed reference values of each variable (Figure 5). We chose to use a single reference values for each variable (as opposed to individual reference values for each site) computed as the geometric mean of all 276 pooled observations of each variable, which results in centered (zero mean) log-transformed variables. Other choices may provide additional insight (if for example, different physically important reference values are used on a site-specific basis like mean annual discharge or bankfull discharge); however, such analyses are beyond the scope of this paper. Reference values of model variables are given by: $q_{b0} = 0.039$ kg/s/m, $q_{w0} = 4.35\ m^2/s$, $C_{s0} = 1.07 \times 10^{-4}$, and $D_{s0} = 0.13$ mm. Chan-

Table 1. Summary of variable ranges measured at each site

	Q_w $[m^3/s]$		W $[m]$		$\log_{10}(S)$		D_b $[mm]$	
	min	max	min	max	min	max	min	max
Atchafalaya River	931	14186	314	503	-5.0	-4.3	0.10	0.41
Mississippi @ Tarbert Landing	4248	28827	896	1414	-4.7	-4.4	0.20	0.38
Mississippi @ St. Louis	1512	8778	457	518	-5.0	-3.2	0.20	0.86
Red River	190	2826	130	183	-4.2	-3.1	0.11	0.28
Rio Grande River	35	286	41	198	-3.1	-3.0	0.25	0.45
Middle Loup River	9	14	22	46	-2.9	-2.7	0.34	0.48
Niobrara River	6	21	19	41	-2.9	-2.7	0.30	0.40
Colorado @ Diamond Creek	267	590	59	64	-4.0	-3.7	0.30	0.50
	q_w $[m^2/s]$		C_s $[ppm]$		D_s $[mm]$		Q_b $[kg/s]$	
	min	max	min	max	min	max	min	max
Atchafalaya River	2.9	28.6	4	372	0.08	0.16	0.20	12.5
Mississippi @ Tarbert Landing	4.7	24.2	5	199	0.10	0.18	0.43	6.7
Mississippi @ St. Louis	3.3	17.2	13	307	0.10	0.25	0.63	11.6
Red River	1.2	20.1	8	495	0.09	0.12	0.10	3.3
Rio Grande River	0.3	3.4	373	3177	0.12	0.22	1.5	41.1
Middle Loup River	0.2	0.6	183	1032	0.13	0.18	1.8	7.8
Niobrara River	0.2	0.9	189	1088	0.08	0.18	1.0	11.0
Colorado @ Diamond Creek	4.5	9.1	2	135	0.12	0.22	0.33	8.6

nel widths were computed using an empirical power-law function of water discharge at the Diamond Creek field site. Reported widths were used at other sites.

Here, we emphasize that the full dataset contains observations of bedload flux that were obtained using two very different methods. Bedload was estimated from grain stresses computed using the Einstein drag partition and the Wong & Parker bedload equation for the large river dataset reported by *Toffaleti* [1968], while bedload flux at Diamond Creek was computed using observations of bedform migration. For the purposes of statistical analysis, we assume both methods produce unbiased estimates of bedload flux with comparable uncertainty. Consequently, both methods are treated identically in the context of inference and prediction.

4 Results

4.1 Bedload fluxes at Diamond Creek

Bedload flux computed from bedform migration is similar to bedload flux estimated as a constant 5% of suspended sand load during the July 2015 survey period, corresponding to the highest water discharges observed (450 m^3/s to 600 m^3/s). Bedload fractions are significantly higher during the March 2015 and November 2016 survey periods, corresponding to lower water discharges (275 m^3/s to 400 m^3/s). Bedload flux ranged from 0.33 kg/s to 8.6 kg/s during the various data collection intervals (Figure 4). The bedload fraction is negatively correlated with suspended sand flux, ranging from as little as 3% to as much as 26% of suspended sand flux.

4.2 Inference on model parameters

Kernel density estimates of the marginal posterior distributions of model parameters are are plotted in Figure (6). The statistical effect of each predictor variable is quantified by the value of the β exponent corresponding to that variable. Peaked distributions indicate low

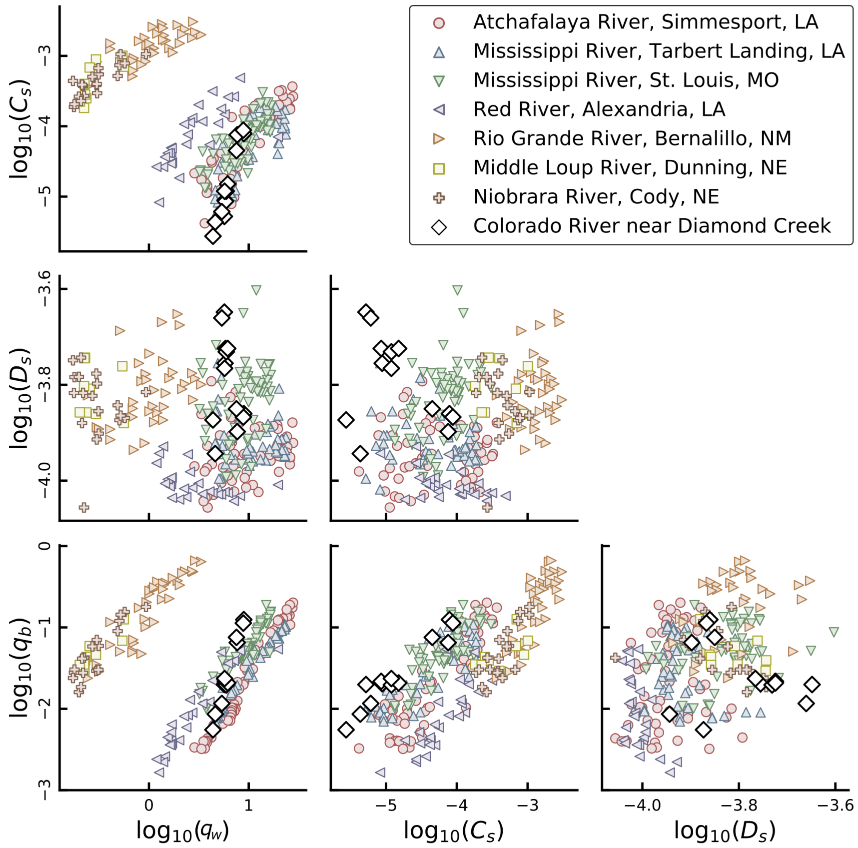


Figure 5. Expanded visualization of regression data. Pale colored markers indicate values of model variables computed from data reported by *Toffaleti* [1968]. Note that predictor variables (q_w, C_S and D_s) cover a wide range of conditions and are only weakly correlated when viewed collectively. Site-specific correlations are evident, especially between C_s and q_w

.

Table 2. Median posterior parameter estimates

Location	β_0	β_1	β_2	β_3	σ
	Grouped Model (DIC = 552)				
All	0.08	0.048	0.68	1.65	0.52
	Ungrouped Model (DIC = 382)				
Atchafalaya River	-1.68	1.65	0.21	-0.24	0.20
Mississippi @ Tarbert Landing	-1.41	1.39	0.09	-0.27	0.16
Mississippi @ St. Louis	-0.58	1.29	0.11	0.02	0.25
Red River	-0.90	1.01	0.41	-0.18	0.24
Rio Grande River	1.07	0.92	0.68	0.01	0.34
Middle Loup River	2.95	0.95	0.02	-0.96	0.26
Niobrara River	2.84	1.08	0.27	-1.10	0.32
Colorado @ Diamond Creek	-2.56	5.04	-0.16	-0.35	0.10
	Hierarchical Model (DIC = 123)				
Atchafalaya River	-1.63	1.61	0.22	-0.19	0.22
Mississippi @ Tarbert Landing	-1.35	1.36	0.11	-0.20	0.22
Mississippi @ St. Louis	-0.51	1.24	0.15	-0.16	0.22
Red River	-0.90	1.04	0.39	-0.19	0.22
Rio Grande River	1.47	1.01	0.57	-0.16	0.22
Middle Loup River	2.45	0.87	0.14	-0.22	0.22
Niobrara River	2.68	1.07	0.31	-0.30	0.22
Colorado @ Diamond Creek	-0.43	2.07	0.36	-0.12	0.22
μ_k	0.19	1.30	0.28	-0.21	
σ_k	1.83	0.44	0.18	0.11	

parameter estimation uncertainty, and wide distributions indicate high uncertainty. Median parameter estimates are reported in Table 2.

Computed DIC values indicate that the hierarchical model has the lowest expected prediction error averaged across all sites. In order to evaluate the effect of each parameter on predictive power, we computed DIC using permutations of each model involving only two predictor variables. The predictive power using the grouped model is significantly reduced using any of the two-variable permutations. However, we find that the predictive power of the ungrouped model is improved by ignoring D_s^* (DIC = 298 compared to 382). This indicates that considering D_s^* does not improve model fit enough to justify the added complexity. Excluding D_s^* has essentially no effect on the predictive power of the hierarchical model (DIC = 112 compared to 123).

4.3 Prediction

Predictive distributions of total mass bedload flux (Appendix B.5) were computed using all three models using hourly-average measurements of Q, C_s, and D_s recorded at the Diamond Creek gage from January 1, 2008 to December 31, 2016. This was accomplished by computing full posterior predictive distributions for each gage measurement of model variables. Median predictions are compared against observational data in Figure (7). The full simulated time series of bedload flux, the ratio of bedload to suspended load, and predictor variables are plotted in Figure (8).

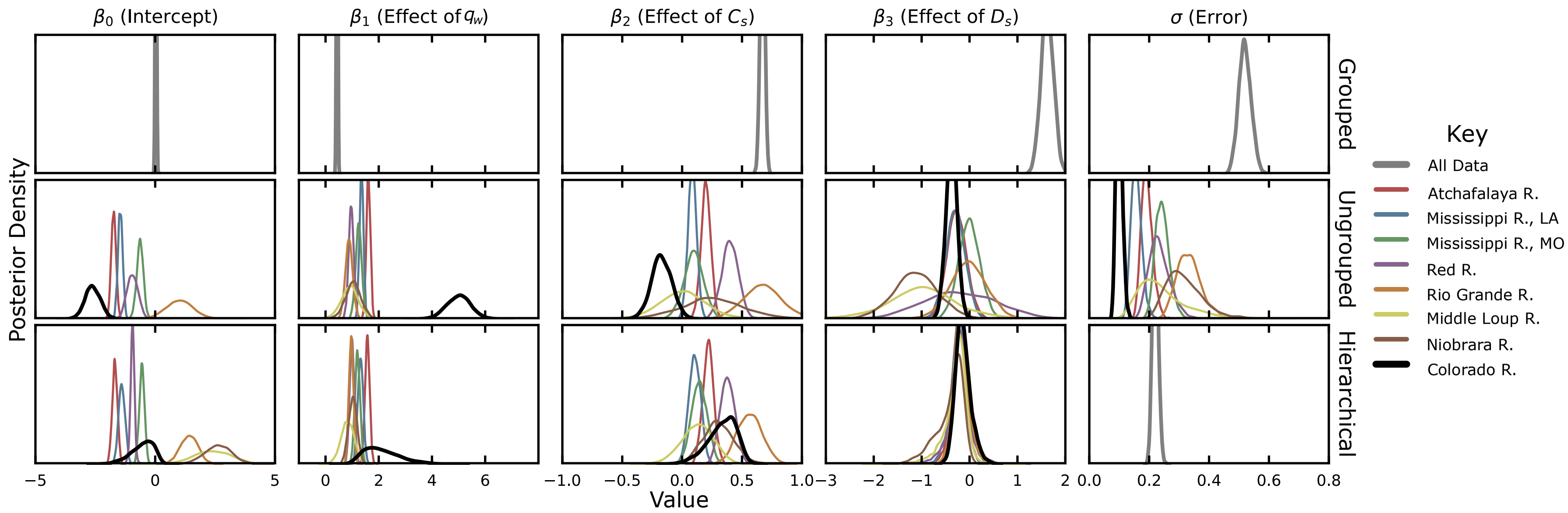


Figure 6. Marginal posterior distributions for model parameters. Distribution widths indicate parameter estimation uncertainty. The grouped model uses a single set of parameters to describe all data, while ungrouped and hierarchical model variations involve fitting separate regression parameters to each site. Note that site-specific posterior distributions are more clustered for the hierarchical model than for the ungrouped model owing to the dynamic, informative priors for the regression coefficients.

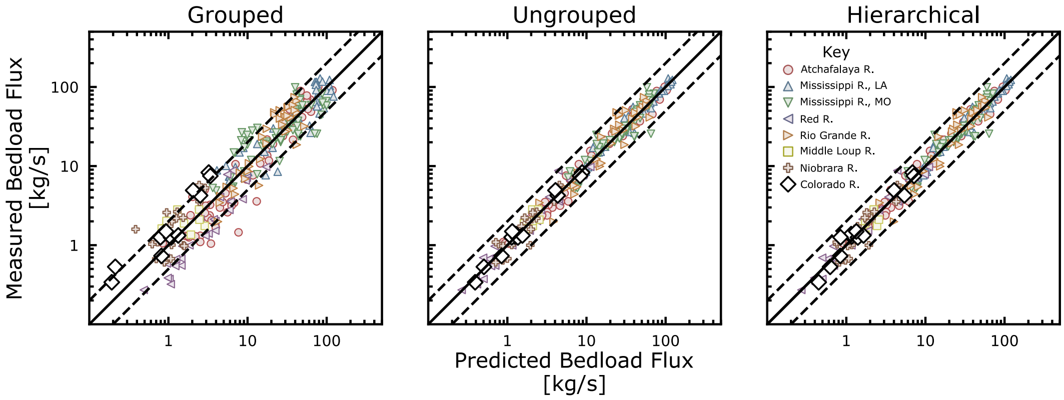


Figure 7. Plot comparing predicted and observed bedload flux. Predictions reflect median parameter estimates. Dashed lines indicate a factor of two deviation between predicted and observed bedload flux. Note that the ungrouped and hierarchical models provide improved fit compared with the grouped model. The hierarchical model leads to more precise estimates of model parameters while providing similar fit to the data when compared with the ungrouped model.

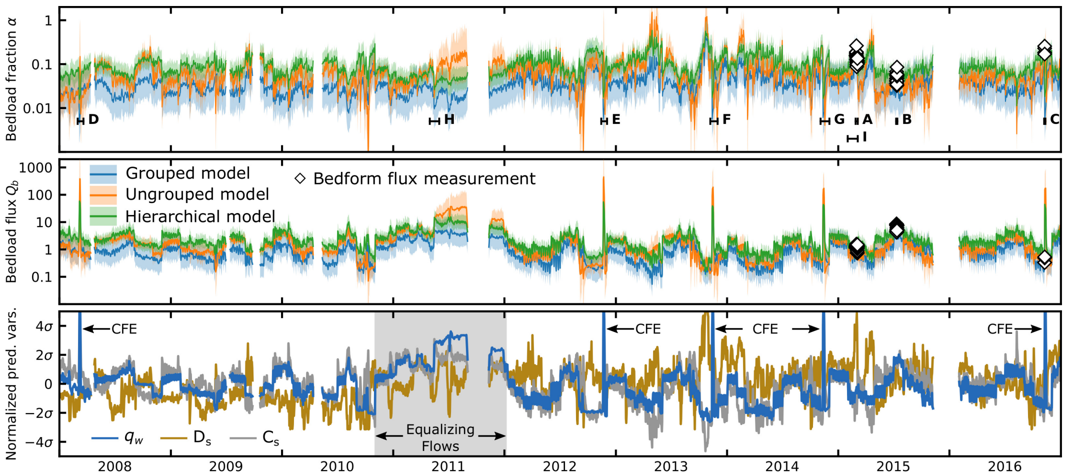


Figure 8. Simulated hourly-averaged bedload fraction (upper panel), bedload flux (middle panel) and transformed predictor variables (lower panel) over the full gage record. Dark lines represent the median of the predictive distribution for bedload flux. Shaded regions represent 95% prediction intervals. Bracketed segments denoted A through I are plotted in Figures (9), (10), and (11). Plotted predictor variables are log-transformed and then normalized by subtracting the mean and dividing by the standard deviation. Controlled flood experiments (CFE's) and elevated "equalizing flows" used to balance reservoir levels are also indicated in the bottom panel.

5 Discussion

5.1 Comparison of model variations

We have presented three variations on our generalized bedload modeling framework that differ in their assumptions, implementation, and interpretation. Here, we compare model variations in the context of the statistical inference and predictions reported in Section 4.

The grouped model most closely encapsulates the physical reasoning presented in Section 3.1, which argues that quasi-universal relationships between transport parameters emerge through the processes governing their interaction and equilibration. These relationships comprise three primary modes of variability driven by water discharge, channel geometry, and bed composition. Three predictor variables improve predictive power compared with any two-parameter model permutation, indicating that that all three modes of variability are represented in the data.

In principle, the grouped model can be applied at any site to predict bedload flux, including new sites that lack direct observational data. However, while individual predictions are unbiased relative to the full dataset, systematic biases exist among groups of measurements that come from a single site; for example, the grouped model under-predicts bedload flux at the Diamond Creek field site (Figure 7). Systematic biases are problematic when computing sediment budgets because they accumulate over time to cause compounded errors.

By considering each site separately, the ungrouped and hierarchical models reduce site-specific systematic biases. They also reflect a restricted scope of physical process: while the grouped model represents quasi-universal physical relationships across many sites, the ungrouped and hierarchical models capture site-specific associations between variables. As a result, we find that two-parameter permutations of the grouped and hierarchical models (ignoring D_s) provide equal or better predictive power than the generalized three-parameter approach. This observation can be explained by the fact that slope is effectively fixed at each site over human timescales in comparison to the differences observed between rivers, reducing the number of principle modes of variability to two. These modes are driven by fluctuations in flow strength and sediment supply, where sediment supply influences fluxes through both "grain size and reach-geometric effects" (*sensu Topping et al.* [2000a,b]). This finding is potentially valuable for sediment monitoring purposes because measurements of C_s are significantly easier to obtain than measurements of D_s. C_s varies by many orders of magnitude and can be measured accurately using single-frequency instruments in a wide range of conditions, while D_s requires well sorted suspended material, two-frequency instrumentation, and is only accurate for a small range of grain sizes [*Topping & Wright*, 2016].

The hierarchical model differs from the ungrouped model in that the site-specific associations between variables are assumed to be similar between sites. Through this assumption, sites with many observations inform prediction at sites with relatively few observations. This effect is most clear at our field site, where few observations ($n = 14$) lead to spurious point estimates of regression parameters (Table 2) and large uncertainty (Figure 6) using the ungrouped model. The hierarchical model produces a slightly poorer fit to the data but yields much more precise and consistent estimates of regression parameters.

In summary, each model has a specific set of assumptions, data requirements, and limitations that must be evaluated in order to be applied to a specific problem. The grouped model reflects quasi-universal physical relationships between variables and can be applied at any site without training data but introduces systematic bias to cumulative bedload estimates. The ungrouped model minimizes site-specific, systematic biases and assumes nothing about similarity between sites but requires extensive observational data to be applied at a given site. The hierarchical model reduces the number of observations needed at a site relative to the ungrouped model under the assumption that sites are similar. Grouped and hierarchical models can potentially be applied using only measurements of Q_w and C_s.

Presently, we aim to compute sediment budgets over the full gage record at the Diamond Creek sediment monitoring station. We argue that the hierarchical model is the best choice for this purpose because it reduces systematic bias but provides efficient use of limited data. Time series predictions made using the hierarchical model are plotted over select intervals in Figures (9), (10), and (11).

5.2 Comparison with existing methods for estimating bedload flux

Prior to this research, the two primary methods for estimating bedload flux from gage data in practical applications are (1) rating curves with discharge [e.g., *Leopold & Maddock*, 1953; *Emmett & Wolman*, 2001] and (2) constant bedload coefficients based on continuous measurements of C_s [e.g., *Rubin et al.*, 2001; *Grams et al.*, 2013]. To highlight the advantages of the model presented here, we compare simulated bedload time series with rating curve and bedload coefficient predictions. Several short example intervals were selected for this purpose and are plotted in Figures (10) and (11).

Both approaches are special cases of our general model (equation 1), wherein certain parameters are fixed. For example, rating curves express bedload flux as a power-law function of water discharge, i.e.:

$$Q_b = kQ_w^m \tag{12}$$

which is similar to equation (1) with null coefficients on suspended sand concentration C_s and diameter D_s:

$$q_b = Ae^{\beta_0} q_w^{\beta_1} C_s^0 D_s^0 W^{1-\beta_1}. \tag{13}$$

Assuming width scales with discharge ($W = aQ_w^b$), this reduces to

$$Q_b = (Ae^{\beta_0} a^{1-\beta_1}) Q_w^{\beta_1+b(1-\beta_1)}. \tag{14}$$

Here, $k = Ae^{\beta_0} a^{1-\beta_1}$ and $m = \beta_1 + b(1-\beta_1)$ are assumed to be constant. For the purposes of comparing rating curve and hierarchical predictions, rating curve parameters (k and m) were found using ordinary least-squares regression applied to concurrent observations of water discharge and bedload flux obtained at the gaging station and from repeat surveys of dune migration, respectively. By specifying $\beta_2 = 0$ and $\beta_3 = 0$, rating curves assume a unique relationship between bed composition, channel geometry, and discharge, which is problematic because sediment supply limitation is known to modify the transport efficiency of a given discharge through reach-geometric and grain size effects [*Topping et al.*, 2000a,b]. Sediment supply variability can thus cause systematic deviations from rating-curve predictions; pulses of fine bed material result in an enriched state characterized by increased bedload flux. Subsequent preferential evacuation of fine material produces a depleted state during which bedload flux is suppressed relative to a hypothetical discharge rating curve prediction (Figure 12). Our modeling approach provides the potential to capture the effects of sediment supply limitation parameterized by C_s and D_s. As a result, we interpret the difference between hierarchical model predictions and rating curve predictions as an indicator of the relative supply-limitation state of the Diamond Creek sediment monitoring reach: a positive difference is indicative of relative enrichment of fine sand whereas a negative difference is indicative of relative depletion.

Such enrichments or depletions are particularly pronounced during and after controlled flood experiments (Figure 10). For example, the period following each controlled flood typically records finer suspended sand grain sizes and elevated suspended sand concentrations relative to antecedent conditions, indicating fine-sediment enrichment [*Rubin & Topping*, 2001]. This is perhaps caused by delivery of fine material accessed above the typical high water line and/or the reworking of existing alluvial deposits in a manner that increases transport efficiency. Hierarchical model predictions are correspondingly elevated relative to rating curve predictions following each controlled flood.

Bedload coefficients are sometimes used to account for the contribution of bedload to total load in scenarios where measurements of suspended flux are available and bedload

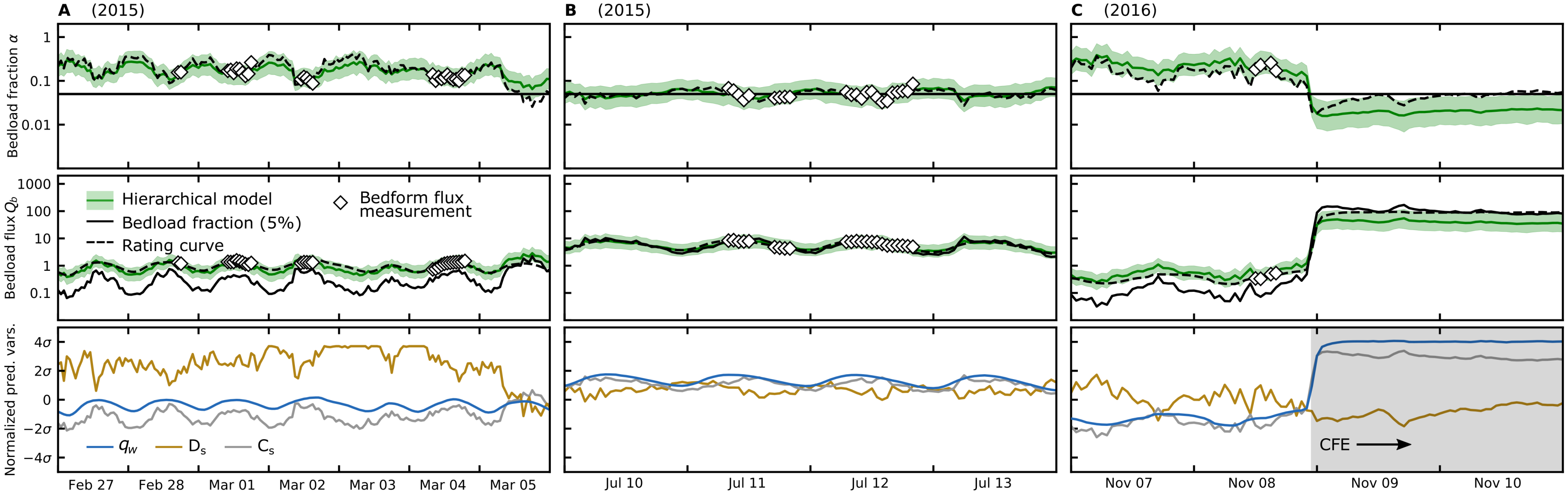


Figure 9. Simulated bedload fraction (upper panel), bedload flux (middle panel) and predictor variables (lower panel) at fifteen-minute resolution during the periods where observations of bedload flux from bedform migration are available. Plotted predictor variables are log-transformed and then normalized by subtracting the mean and dividing by the standard deviation.

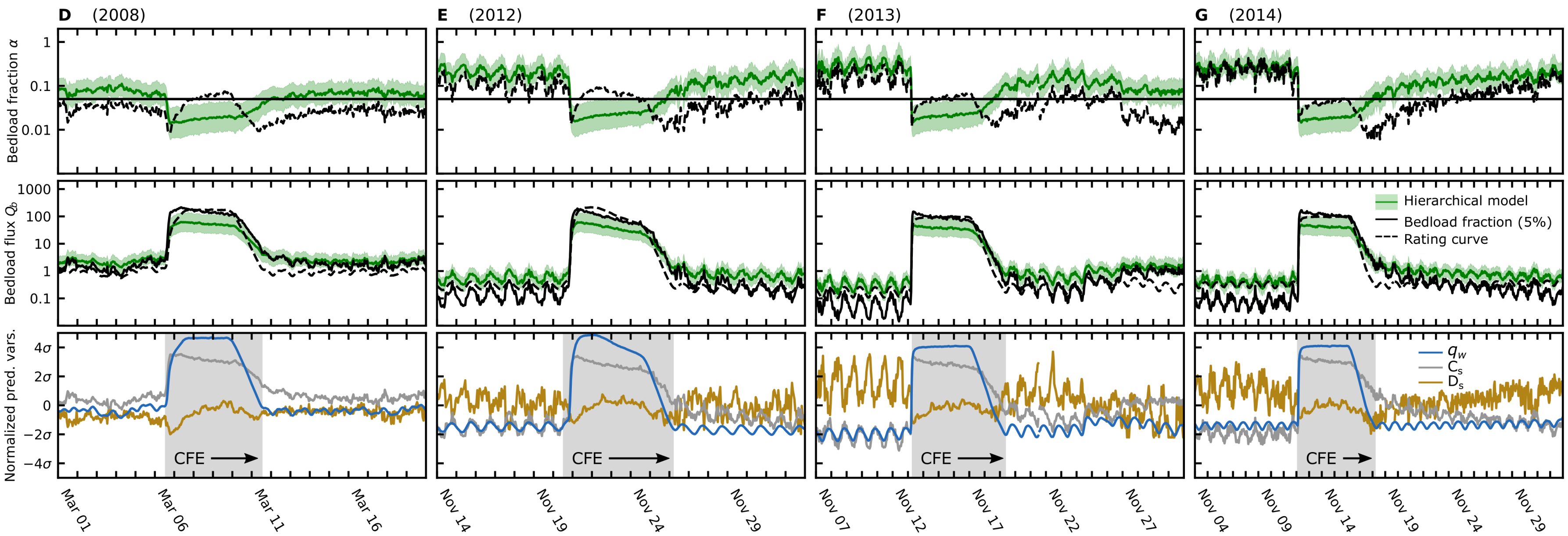


Figure 10. Simulated bedload fraction (upper panel), bedload flux (middle panel) and transformed predictor variables (lower panel) for the periods surrounding controlled flood experiments (CFEs).

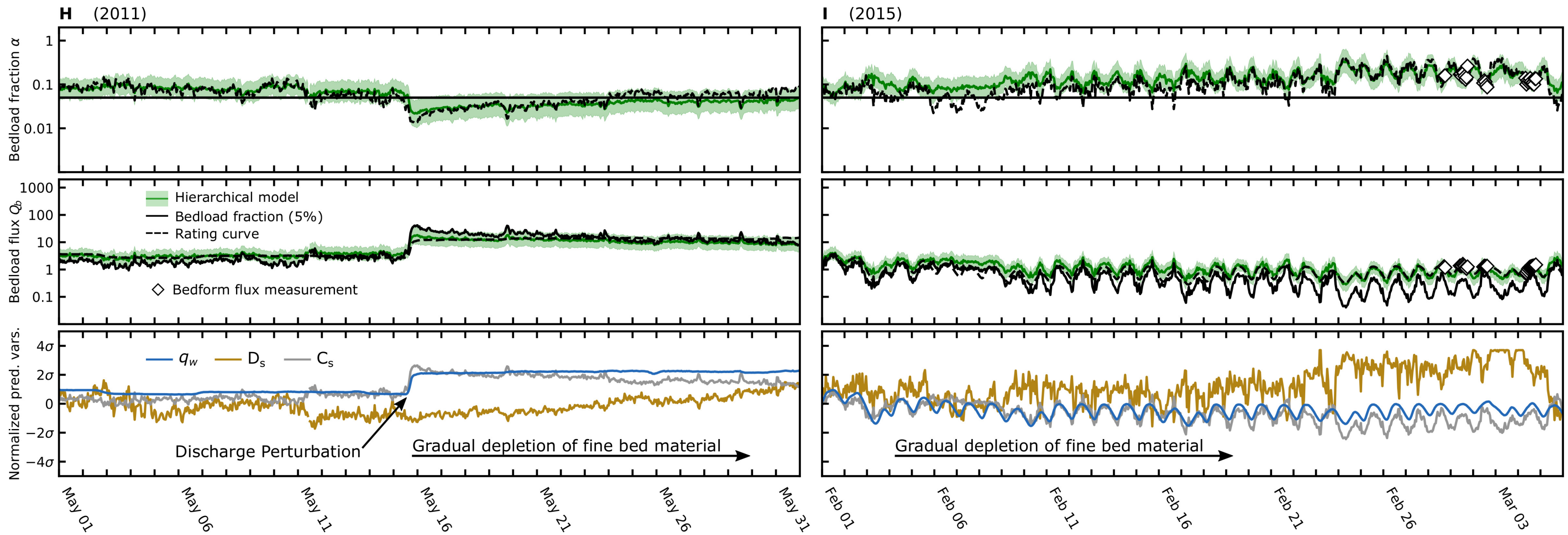


Figure 11. Simulated bedload fraction (upper panel), bedload flux (middle panel) and transformed predictor variables (lower panel) for two periods that record gradual coarsening of bed material under nearly constant (H) and periodically-fluctuating (I) flow conditions.

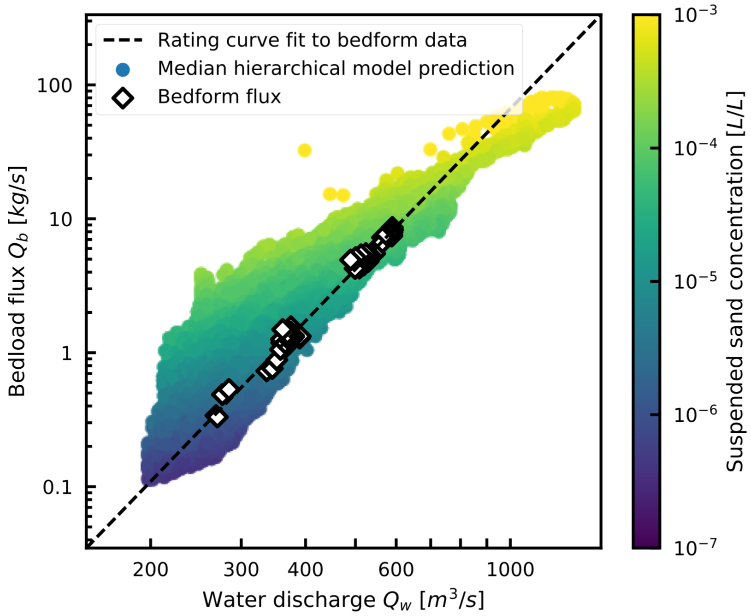


Figure 12. Plot illustrating the advantages of the proposed model over a traditional rating curve approach. Note that predicted bedload flux may vary by over an order of magnitude with respect to a fixed water discharge, an effect that is typically attributed to supply-limitation effects. Suspended sand concentration is connected to the supply-limitation state of a reach; here, elevated suspended sand concentrations indicative of fine-sediment enrichment and amplified bedload flux.

is thought to be small [e.g., *Grams et al.*, 2013]. In order to estimate total load, researchers sometimes apply a universal correction factor $1 + \alpha$ to measurements of suspended sand flux Q_s, which implies

$$q_b = \alpha q_s. \tag{15}$$

Noting that $q_s = q_w C_s$, equation (15) is a special case of the our general bedload model (1) wherein $\beta_1 = 1$, $\beta_2 = 1$, and $\beta_3 = 0$, i.e.:

$$q_b = A e^{\beta_0} q_w^1 C_s^1 D_s^0. \tag{16}$$

Here, $\alpha = A e^{\beta_0}$ is the constant bedload coefficient. In some sense, this expression represents a crude attempt to account for supply limitation effects by assuming bedload and suspended load are equally sensitive to changes in their mutual causal predictors (water discharge, channel geometry, and bed composition). However, suspension conditions (parameterized by the Rouse number, $Z_R = w_s/\kappa u_*$, where w_s is the particle setling velocity, u_* is the basal shear velocity, and κ is von Karman's constant) vary with flow strength and sediment supply, and are the most important predictor of α (*van Rijn* [1984], Equation 45). Insofar as the Rouse number may vary over time at a site, it is unreasonable to expect that the bedload fraction should remain constant. Instead, increasing Z_R should generally cause an increase in α. This may occur due to changes in u_* (as a function of water discharge, channel geometry, and bed roughness), or due to changes in w_s, which is a monotonically increasing function of D_b.

Comparison of hierarchical model and bedload coefficient predictions reveals several expected behaviors. In general, elevated suspended sand fluxes tend to correspond to increased suspension conditions (low Rouse numbers) and low bedload fractions (Figure 13). Bedload flux is a larger fraction of total load when discharge is low, corresponding to higher Rouse numbers due to decreases in u_*. Sediment supply depletion also increases the bedload fraction when discharge is held constant, corresponding to higher Rouse numbers due to

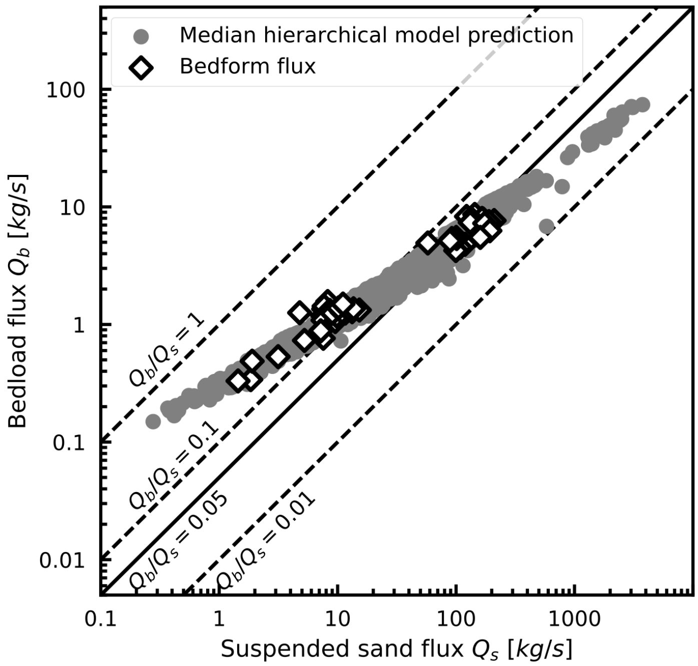


Figure 13. Plot illustrating the advantages of the proposed model over a constant bedload fraction approach. Diagonal lines show contours of constant bedload fraction. The predicted (and to a lesser extent, measured) trend seen here is consistent with the notion that high suspended sand fluxes correspond to elevated suspension conditions and lower bedload fractions.

increases in w_s. Discharge effects are most pronounced before, during, and after controlled flood experiments (Figure 10), which exemplify both the high and low bedload fraction extremes. Supply limitation effects are evident during long periods of nearly constant discharge which tend to be associated with gradual sediment-supply depletion (Figure 11). Gradual depletion causes a concurrent increase in the bedload fraction which is also apparent after the 2014 controlled flood experiment (Figure 10).

5.3 Management implications

Sediment budgets are used to estimate changes in stored sediment mass over a wide range of timescales. Short term effects of interest include perturbations related to dam-regulated water discharge or tributary sand delivery. Serendipitously (in the context of 5% bedload coefficients used by *Rubin et al.* [2001]; *Topping et al.* [2010]; *Grams et al.* [2013]), we find that cumulative bedload discharge was approximately 5% of the cumulative suspended sand discharge over the nine-year record considered here. However, instantaneous bedload flux ranges from less than 1% to as much as 75% of suspended sand load depending on water discharge and the supply-limitation state. As a result, short-term mass-balance fluctuations caused by experimental changes in discharge regime (i.e. controlled floods), transient accommodation of tributary sand pulses, or prolonged periods of constant discharge are not adequately represented using a constant bedload fraction or rating curve model for bedload flux. For example, cumulative bedload during controlled floods is only 2% of cumulative suspended load during the same intervals, whereas the cumulative bedload is 10% of cumulative suspended load during period when flow is below the mean annual discharge. In general, the magnitude of deviations in short-term average bedload fraction from a measured long-term average is a function of averaging timescale (Figure 14)

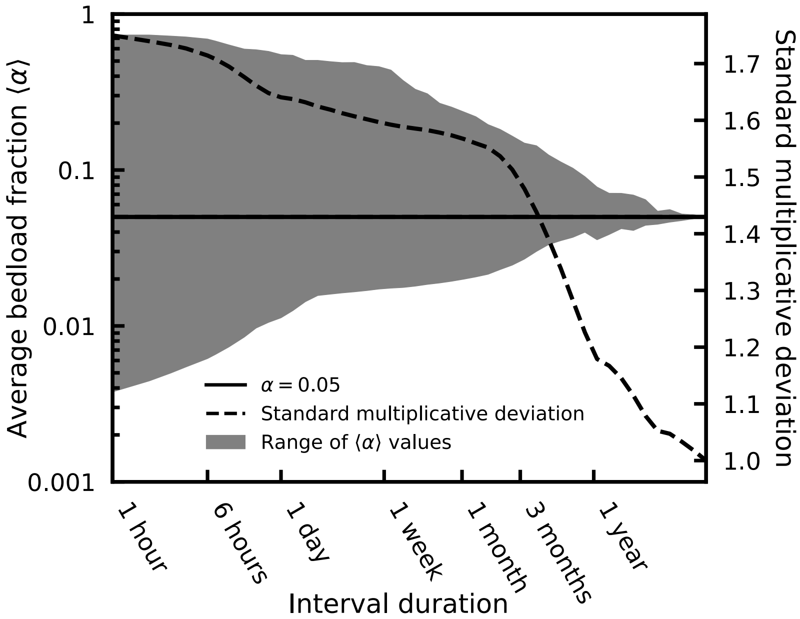


Figure 14. Plot illustrating the timescale dependence of the average bedload fraction, $\langle \alpha \rangle$. Average bedload fraction was computed for every period with the duration indicated on the x axis. The filled gray area spans the full range of average bedload fractions computed for intervals with the specified duration. The dotted gray line shows the standard deviation of the average bedload fraction as a function of interval duration. At the minimum model resolution, bedload fraction may range from 0.004 to 0.74.

Over longer timescales, researchers aim to constrain the effects of changes in the water discharge or sediment delivery regime as dictated by dam protocols, climate, and land use in the upper Colorado River basin [e.g., *Andrews*, 1991; *Grams et al.*, 2013; *Mueller et al.*, 2014; *Grams et al.*, 2015; *Kasprak et al.*, 2018; *Mueller et al.*, 2018]. In particular, the dam-regulated water discharge regime is the primary tool for enacting management decisions aimed at balancing ecological, social, and economic goals. Nearly three decades of Grand Canyon research suggests that a return to a more natural, seasonal discharge regime would induce a desirable geomorphic response. Actionable proposals like the "Fill Mead First" plan [*Schmidt et al.*, 2016] are designed to balance this and other management objectives by changing the annual cycle of dam releases, and flux-based sediment budgets are critical for accurately evaluating the effects and effectiveness of such plans. However, the intended geomorphic response will necessarily involve changes in channel geometry and bed composition, affecting sediment flux in a manner that cannot be tracked using traditional rating curve or bedload fraction approaches. Measurements of q_w, C_s and D_s are indicative of changes in q_b and such that it is possible to resolve short-term morphodynamic adjustment and evaluate the effects of future changes in the water discharge and sediment supply regime.

5.4 Other applications of modeling approach

Bedload has historically been difficult to measure directly. As a result, its role in governing large-scale river organization poorly understood. Although this paper focuses on estimating bedload on the Colorado River, the modeling approach presented herein will enable improved estimates of bedload flux in any sand-bedded river. Our model can be applied retroactively to innumerable historical measurements of suspended sediment concentration and grain size, providing a new approach for connecting bedload transport to continent- and basin-scale river dynamics.

This work also supports a more general principle that extends beyond the problem of estimating bedload flux. We have argued that our bedload model provides reliable predictions because it approximates quasi-universal relationships between transport parameters emerge through the processes governing their interaction and equilibration. In this view, first order changes in flow and transport conditions including bedload flux, suspended sand concentration, and suspended sand diameter are driven by three variables: water discharge, slope, and bed material grain size. This implies that any relevant variable can be estimated from measurements of three other variables, providing a general formula for constructing predictive empirical relations in sandy fluvial systems. This strategy may prove useful for reconstructing hydraulic and transport conditions in scenarios where certain variables are difficult or impossible to measure, for example in applications involving remotely sensed river data or measurements of fluvial sedimentary rocks.

6 Conclusions

The modeling approach presented here was developed to estimate reach-averaged bedload flux from measurements of water discharge, concentration, and grain size in suspension. This approach is based on the assumption that most of the variability in sand-bed rivers can be reduced to three principle modes of variation that are causally attributed to water discharge, slope, and bed grain size. Measurements of concentration and grain size in suspension provide reliable proxies for the effect of slope and bed material grain size on bedload flux.

Bayesian hierarchical modeling assumes similarity between rivers to ensure efficient use of limited data. This approach reduces in-sample bias compared with a fully grouped regression, and it improves parameter estimation precision compared with the ungrouped regression. However, we anticipate that the general modeling approach presented here may prove useful in other contexts for which grouped or ungrouped generative data models may be preferable.

We find that predicted bedload flux during the period from 2008 to 2016 averaged over the full gage record at Diamond Creek is approximately 5% of the measured suspended sediment load. However, instantaneous values deviate significantly from 5% depending on flow strength and sediment supply conditions. Notably, changes in bedload flux at a constant water discharge are indicative of short-term sediment supply enrichment and depletion. Using the median prediction from the hierarchical model, we find that bedload flux ranges from as high as 75 % of suspended sand load (during fine-sand depleted, low-discharge periods) to less than 1% (during fine-sand enriched floods). The decade-average bedload fraction is expected to deviate systematically from 5% in the future if bed composition and channel geometry evolve due to changes in tributary sand supply or the dam-regulated discharge regime. In order to ensure accurate quantification of fluctuations in sediment storage over a range of timescales, it is critical to account for deviations in the ratio of bedload to suspended load driven both by individual events (for example, high flow experiments or tributary floods) and long-term evolution of channel geometry and bed composition.

A: Estimating bedload flux from repeat bathymetric surveys of dune migration

Bedload flux estimates at our site were computed from point clouds of bed topography obtained at approximately six-minute intervals. This was accomplished using the following procedure:

1. Flow direction is determined by inspection and point clouds are transformed to streamwise and cross-stream coordinates.

2. An upstream and downstream extent is chosen to bracket a region of the bed used for computation of flux. The region used here is largest region where the margins of the bedform field are parallel and bedform geometry appears to be uniform in all surveys.
3. Point clouds are divided by cross-stream coordinate into streamwise oriented transects spaced at 25 cm.
4. Ungridded points that fall within each 25 cm-wide transect are gridded at a 10 cm streamwise resolution using a locally-weighted nonparametric filter.
5. Transects are detrended using a high-pass Fourier filter. The filter wavelength used here is three times the largest dune length determined by inspection.
6. Characteristic bedform height is estimated as $2\sqrt{2} * \sigma_\eta$ where σ_η is the root mean squared detrended bed elevation [*McElroy*, 2009]
7. A matrix of dune displacements (determined from the maximum of the cross-correlation function) is computed for each transect using every pair of surveys. Valid displacements are retained to calculate migration rate according to the following criteria: (a) temporal separation is not greater than one hour, (b) displacement is not greater than 20 percent of the bedform length, determined from the spectral centroid of the detrended bed profile [*Van der Mark & Blom*, 2007], (c) the maximum of the cross correlation function is not less than 0.8, and (d) the implied migration rate (displacement divided by temporal separation) is not greater than 3 meters per hour and not less than 0.3 meters per hour. These criteria optimize temporal resolution and stability of the bedload flux calculation, and reliably discriminate transects with active dune evolution from plane-bed topography.
8. Bedform migration rate is computed for each transect using ordinary least-squares regression forced through the origin with all valid displacements.
9. Volumetric bedload flux per unit width is computed for each streamwise transect using the bedform bedload equation [*Simons et al.*, 1965].
10. Total bedload mass flux was computed for each transect by multiplying unit bedload flux by the transect width (25 cm) and the density of quartz (2650 kg/m^3), then summed.

We find that the bedform migration rate regression using displacements forward and backward in time is necessary to ensure stable results. However, this means that bedload flux estimates are derived from overlapping data. Down sampling is thus necessary to ensure that each reported value of bedload flux is computationally independent: we consider a maximum temporal resolution of one hour. Results are plotted in Figure 4.

B: Bayesian regression

Here, we provide additional details on the statistical techniques employed in this paper. In order to make this explanation more clear, we adopt notation that is common in statistical literature [e.g. *Gelman et al.*, 1995; *Christensen et al.*, 2011]. We consider the problem of predicting a continuous response variable y from a vector of predictor variables $\bar{\mathbf{x}} = [1, x_1, x_2, x_3]$. The relationship between predictor variables and response variables is studied using a probabilistic model with parameters θ for the data generating process.

Physical variables of interest are log-transformed and normalized to obtain linear predictor and response variables such that $y = \log(q_b/q_{b0})$, $x_1 = \log(Q/Q_0)$, $x_2 = \log(C_s/C_{s0})$ and $x_3 = \log(D_s/D_{s0})$ and The subscript 0 denotes the geometric mean of all observations, which is equivalent to subtracting the arithmetic mean of log-transformed variables and results in centered response and predictor variables. This is a convention that facilitates interpretation of the intercept term β_0. The subscript i denotes a specific observation such that y_i and $\bar{\mathbf{x}}_i$ are the i^{th} of n observations of response and predictor variables, respectively. A capital X is short hand for all observations of model variables, i.e. $X = (\bar{\mathbf{x}}_0, ..., \bar{\mathbf{x}}_n, y_0, ..., y_n)$. Finally, we use $\tilde{x}_i$ to denote a vector of observations of predictor variables for which we intend to predict an unobserved value of the response variable, $\tilde{y}_i$.

B.1 Grouped model

The grouped model ignores potential correlations that may exist on a site specific basis. All data is pooled into a single normal linear regression analysis. Regression coefficients and errors are assumed to be equivalent at all sites. Here, $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2, \beta_3]$ is a 1×4 vector of regression coefficients. The i^{th} observation of the response variable y_i is modeled as a linear function of predictor variables plus a normally-distributed independent error term (e.g. equation 8). This is equivalent to specifying that the y_i follows a normal distribution with mean $\boldsymbol{\beta}\bar{\mathbf{x}}_i$ and standard deviation σ. Formally, the probability of observing y_i given $\bar{\mathbf{x}}_i$, $\boldsymbol{\beta}$, and σ is given by:

$$p(y_i|x_i, \beta, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \beta x_i)^2}{2\sigma^2}\right] \tag{B.1}$$

and the likelihood of model parameters $\theta = (\boldsymbol{\beta}, \sigma)$ conditional on all observational data $X = (x_0, ..., x_n, y_0 ... y_n)$ is simply the product of the probabilities of each individual observation:

$$L(\theta|X) = \prod_{i=1}^{n} (y_i|x_i, \boldsymbol{\beta}, \sigma). \tag{B.2}$$

For the grouped model, we employ the following independent priors for model parameters:

$$\begin{aligned}
\beta_0 &\sim \mathcal{N}(0, 100) & \text{(B.3)}\\
\beta_1 &\sim \mathcal{N}(0, 100) & \text{(B.4)}\\
\beta_2 &\sim \mathcal{N}(0, 100) & \text{(B.5)}\\
\beta_3 &\sim \mathcal{N}(0, 100) & \text{(B.6)}\\
\sigma &\sim \Gamma^{-1}(0.001, 0.001). & \text{(B.7)}
\end{aligned}$$

Here, $\mathcal{N}(\mu, \sigma)$ denotes a normal distribution with mean μ and standard deviation σ, and $\Gamma^{-1}(\alpha_1, \alpha_2)$ denotes the inverse gamma distribution with shape parameter α_1 and scale parameter α_2. Since the marginal priors are independent, $p(\beta, \sigma) = p(\beta_0)p(\beta_1)p(\beta_2)p(\beta_3)p(\sigma)$. These priors approximate Jeffrey's prior for normal linear regression which is a uniform distribution on $(\boldsymbol{\beta}, \log(\sigma))$ [*Gelman et al.*, 1995; *Christensen et al.*, 2011].

The posterior probability distribution of model parameters θ given data X is proportional to the product of the likelihood function and the prior:

$$P(\theta|X) = \frac{L(\theta|X)P(\theta)}{\int L(\theta|X)P(\theta)d\theta}, \tag{B.8}$$

where the constant of proportionality $[\int L(\theta|X)P(\theta)d\theta]^{-1}$ ensures that the posterior integrates to 1.

B.2 Ungrouped model

The ungrouped model involves fitting separate regression models for each site. Henceforth, the subscript $j = 1, ...m$ denotes the j^{th} of $m = 8$ sites. $\boldsymbol{\beta}_j = [\beta_{0j}, \beta_{1j}, \beta_{2j}, \beta_{3j}]$ is thus the vector of regression coefficients corresponding to site j, and σ_j is the standard deviation of the error term at site j. The full data model thus contains $4 \times m$ regression coefficients and m error terms, totaling 40 parameters compared with the 5 parameters used in the ungrouped model.

Each site has a different number of observations, n_j. The probability of observing i^{th} of n_j observations of the response variable at site j, $y_{i,j}$ given $x_{i,j}$, β_j, and σ_j is given by

$$p(y_{i,j}|x_{i,j}, \boldsymbol{\beta}_j, \sigma_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[-\frac{(y_{i,j} - \boldsymbol{\beta}_j x_{i,j})^2}{2\sigma_j^2}\right] \tag{B.9}$$

and the likelihood function of model parameters $\theta = (\boldsymbol{\beta}_0, ..., \boldsymbol{\beta}_m, \sigma_0, ..., \sigma_m)$ conditional on all observational data X is given by the product of the probabilities of all observations:

$$L(\theta|X) = \prod_{j=1}^{m}\prod_{i=1}^{n_j}(y_{i,j}|x_{i,j}, \boldsymbol{\beta}_j, \sigma_j) \tag{B.10}$$

Separate indpendent priors are used for each site, i.e:

$$\begin{aligned}
\beta_{0j} &\sim \mathcal{N}(0, 100) & \text{(B.11)}\\
\beta_{1j} &\sim \mathcal{N}(0, 100) & \text{(B.12)}\\
\beta_{2j} &\sim \mathcal{N}(0, 100) & \text{(B.13)}\\
\beta_{3j} &\sim \mathcal{N}(0, 100) & \text{(B.14)}\\
\sigma_j &\sim \Gamma^{-1}(0.001, 0.001). & \text{(B.15)}
\end{aligned}$$

B.3 Hierarchical Model

Like the ungrouped model, the hierarchical model involves fitting separate regression coefficients for each site. However, unlike the ungrouped model, these regression coefficients are assumed to come from a common distribution that encompasses the range of parameters that exist in sand bed rivers. Additionally, there a single error term σ is applied at all sites. Instead of using separate, diffuse priors with fixed parameters for the regression coefficients at each site, informative, dynamic priors are used, i.e.:

$$\begin{aligned}
\beta_{0j} &\sim \mathcal{N}(\mu_{\beta_0}, \varsigma_{\beta_0}) & \text{(B.16)}\\
\beta_{1j} &\sim \mathcal{N}(\mu_{\beta_1}, \varsigma_{\beta_1}) & \text{(B.17)}\\
\beta_{2j} &\sim \mathcal{N}(\mu_{\beta_2}, \varsigma_{\beta_2}) & \text{(B.18)}\\
\beta_{3j} &\sim \mathcal{N}(\mu_{\beta_3}, \varsigma_{\beta_3}) & \text{(B.19)}\\
\sigma &\sim \Gamma^{-1}(0.001, 0.001). & \text{(B.20)}
\end{aligned}$$

Here, $\psi = (\mu_{\beta_0}, \mu_{\beta_1}, \mu_{\beta_2}, \mu_{\beta_3}, \varsigma_{\beta_0}, \varsigma_{\beta_1}, \varsigma_{\beta_2}, \varsigma_{\beta_3})$ are known as hyperparameters; μ terms are the mean of the prior on the regression coefficients and represent the central tendency of sites in our data set (as a proxy for sand bed rivers), while ς terms are the standard deviation of the priors and represent the variability present across sites in our dataset. Because the priors depend on dynamic hyperparameters, the posterior probability takes a slightly different form:

$$p(\theta|X) = \frac{L(\theta|X)P(\theta|\psi)P(\psi)}{\int[L(\theta|X)P(\theta|\psi)P(\psi)]d\theta d\psi}, \tag{B.21}$$

where $P(\theta|\psi)$ is the prior probability distribution for model parameters θ given hyperparameters ψ, and $P(\psi)$ is the prior probability distribution for ψ, or the hyperprior. Reported results were obtained using the following diffuse, independent hyperpriors:

$$\begin{aligned}
\mu_k &\sim \mathcal{N}(0, 100) & \text{(B.22)}\\
\varsigma_k &\sim \Gamma^{-1}(0.001, 0.001) & \text{(B.23)}
\end{aligned}$$

for $k = 0, 1, 2, 3$. The grouped and ungrouped models can be framed as special cases of the hierarchical model with informative hyperpriors. Specifically, the grouped model is a case where $\varsigma_k \sim \delta(0)$, where δ is the dirac delta function. This leads to to $\beta_k = \mu_k$ for all sites. The ungrouped model is a case where $\mu_k \sim \delta(0)$ and $\varsigma_k \sim \delta(100)$ such that the hyperpriors exert minimal influence on β_k.

B.4 MCMC sampling

Posterior distributions for model parameters were constructed using the No-U-Turn sampling (NUTS) algorithm [*Hoffman & Gelman*, 2014], as implemented in the open source

Python package, PyMC3 [*Salvatier et al.*, 2016]. The sampler was initiated using the automatic differentiation variational inference algorithm [*Kucukelbir et al.*, 2016]. Three chains were used, and 1000 burn-in steps were more than sufficient to achieve convergence. The posterior distribution of model parameters was approximated using 5000 steps without thinning.

B.5 Prediction

Once the posterior probability distribution of model parameters is known, unobserved values of the response variable $\tilde{y_i}$ can be estimated using Bayesian posterior predictive distributions. The posterior predictive density $P(\tilde{y_i}|\tilde{x_i}, X)$ is found by integrating the sampling distribution of $\tilde{y_i}$ given a specific set of parameters, $p(\tilde{y_i}|\tilde{x_i}, \theta)$, against the posterior distribution of model parameters, $P(\theta|X)$:

$$P(\tilde{y_i}|\tilde{x_i}, X) = \int P(\tilde{y_i}|\tilde{x_i}, \theta)P(\theta|X)d\theta. \tag{B.24}$$

This distribution is straightforward to compute numerically using MCMC techniques. In addition to predicting single unobserved values of q_b, it is possible to obtain a simulated predictive distribution for any conceivable quantity that can be expressed as a function of model parameters (for example, time-integrated bedload flux).

B.6 Deviance Information Criterion

The Deviance Information Criterion (DIC) is a measure of relative predictive power that reflects the trade-off between goodness of fit and parameter estimation precision [*Spiegelhalter et al.*, 2002; Gelman et al., 2014]. It is used here instead of other more well-known model selection criteria like the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) because unlike AIC, it is suitable for comparing the hierarchical and non-hierarchical models considered here, and unlike BIC, its intended use is for comparing expected out-of-sample predictive accuracy under the assumption that the data model is correct.

DIC uses the log-likelihood $\log L(\theta|X)$ of different models to compare expected out of sample predictive accuracy. Models that achieve higher values of the likelihood function provide better in-sample fit. The log-likelihood of the posterior mean parameter estimate $\log L(\bar{\theta}|X)$ is used here to quantify model fit. For clarity, $\bar{\theta} = E(\theta|X)$ is the posterior mean parameter estimate.

More complex models may lead to higher log-posterior densities and better in-sample fit at the cost of parameter estimation precision. In other words, a much wider range of model parameters provide a good fit to the data such that it is difficult to select optimal values. For models that are too complex, predictive uncertainty is primarily related to uncertainty in model parameters rather than being directly quantified by the noise term (σ in the models presented here). It is thus necessary to introduce a correction factor that accounts for parameter estimation uncertainty. Here, the effective number of parameters $p_{DIC} = 2var_{post}(\log L(\theta|X)$ is framed in terms of the posterior variance in the log-likelihood, and can be computed by taking the variance of MCMC sampled log-likelihoods.

The expected log predictive density is given by $elpd = \log L(\bar{\theta}|X) - p_{DIC}$. Assuming predictive error is normally distributed, the expected log predictive density is proportional to the mean squared error. DIC is a related to the expected log posterior density by a factor of -2 due to convention:

$$DIC = -2 \log L(\bar{\theta}|X) + 2p_{DIC} \tag{B.25}$$

For additional details on the derivation and interpretation of DIC, see *Spiegelhalter et al.* [2002]; Gelman et al. [2014].

Acknowledgments

We thank Bob Tusso, Dan Hamill, and Erich Mueller their help in the field. U.S. Geological Survey studies in Grand Canyon are supported by the Grand Canyon Dam Adaptive Management Program administered by the U.S. Department of the Interior Bureau of Reclamation. Additional support was provided by the University of Wyoming School of Energy Resources. Data and software supporting the analysis and conclusions presented here are available online through the SEAD data repository and github [*Leary*, 2018; *Ashley*, 2019a,b]. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Ashley, T. (2019) qDune version 1.0 [Computer Software] https://doi.org/10.5281/zenodo.3515483

Ashley, T. (2019) qbStats version 1.0 [Computer Software] https://doi.org/10.5281/zenodo.3515485

Andrews, E. D. (1991). Sediment transport in the Colorado River basin. In *Colorado River Ecology and Dam Management: Proceedings of a Symposium* (pp. 54–74). Washington, DC: National Academies Press.

Brownlie, W. R. (1983). Flow depth in sand-bed channels. *Journal of Hydraulic Engineering, 109*(7), 959–990.

Buscombe, D., Rubin, D. M., & Warrick, J. A. (2010). A universal approximation of grain size from images of noncohesive sediment. *Journal of Geophysical Research: Earth Surface*, 115, F02015. https://doi.org/10.1029/2009JF001477

Buscombe, D., Grams, P. E., & Kaplinski, M. A. (2014a). Characterizing riverbed sediment using high-frequency acoustics: 1. Spectral properties of scattering. *Journal of Geophysical Research: Earth Surface*, 119, 2674–2691. https://doi.org/10.1002/2014JF003189

Buscombe, D., Grams, P. E., & Kaplinski, M. A. (2014b). Characterizing riverbed sediment using high-frequency acoustics: 2. Scattering signatures of Colorado Riverbed sediment in Marble and Grand Canyons. *Journal of Geophysical Research: Earth Surface*, 119, 2692–2710. https://doi.org/10.1002/2014JF003191

Christensen, R., Johnson, W., Branscum, A., & Hanson, T. E. (2011). *Bayesian ideas and data analysis: An introduction for scientists and statisticians*. Boca Raton, FL: Chapman Hall.

Dean, D. J., Topping, D. J., Schmidt, J. C., Griffiths, R. E., & Sabol, T. A., (2016). Sediment supply versus local hydraulic controls on sediment transport and storage in a river with large sediment loads, *Journal of Geophysical Research: Earth Surface 121*, 82–110, https://doi.org/10.1002/2015JF003436.

Dolan, R., Howard, A., & Gallenson., A. (1974), Man's impact on the Colorado River in the Grand Canyon: The Grand Canyon is being affected both by the vastly changed Colorado River and by the increased presence of man. *American Scientist*, *62*(4), 392–401.

Einstein, H. A. (1950). The bed load function in open channel flows. *U.S. Dept. of Agriculture Technical Bulletin No. 1026.*

Einstein, H. A., & Chien, N. (1953). Can the rate of wash load be predicted from the bedâĂŘload function?. *Eos, Transactions American Geophysical Union*, *34*(6), 876-882.

Ellison, C. A., Groten, J. T., Lorenz, D. L., & Koller, K. S. (2016). Application of dimensionless sediment rating curves to predict suspended-sediment concentrations, bedload, and annual sediment loads for rivers in Minnesota. *USGS Scientific Investigations Report No. 2016-5146*.

Emmett, W. W., & Wolman, M. G. (2001). Effective discharge and gravel-bed rivers. *Earth Surface Processes and Landforms*, *26*(13), 1369–1380. https://doi.org/10.1002/esp.303

Engel, P., and Lau, Y. L. (1980), Computation of bed load using bathymetric data. *Journal of the Hydraulics Division of the Ameican Society of Civil Engineers, 106*(HY3), 369âĂŞ-380.

Engelund, F., & Hansen, E. (1967). A monograph on sediment transport in alluvial streams. *Technisk Forlag, Copenhagen, Denmark*.

Gaeuman, D., & Jacobson, R. B. (2007). Field assessment of alternative bed-load transport estimators. *Journal of Hydraulic Engineering, 133*(12), 1319-1328.

Garcia, M. (2008), Sediment transport and morphodynamics. In *Sedimentation Engineering: Process, Management, Modeling and Practice* (pp. 21–163). Reston, VA: ASCE.

Garcia, M., & Parker, G. (1991). Entrainment of bed sediment into suspension. *Journal of Hydraulic Engineering, 117*(4), 414–435.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and computing*, *24*(6), 997–1016. https://doi.org/10.1007/s11222-013-9416-2

Gibbings, J. C. (2011). The Pi-Theorem. In *Dimensional Analysis* (pp. 55-82). Springer, London.

Grams, P. E., Topping, D. J., Schmidt, J. C., Hazel, J. E., & Kaplinski, M. (2013). Linking morphodynamic response with sediment mass balance on the Colorado River in Marble Canyon: Issues of scale, geomorphic setting, and sampling design. *Journal of Geophysical Research: Earth Surface*, *18*(2), 361–381. https://doi.org/doi:10.1002/jgrf.20050

Grams, P. E., Schmidt, J. C., Wright, S. A., Topping, D. J., Melis, T. S, & Rubin, D. M. (2015). Building sandbars in the Grand Canyon. *EOS: Transactions of the American Geophysical Union*, *96*, 1–11. https://doi.org/doi:10.1002/jgrf.20050

Grams, P. E., Buscombe, D., Topping, D. J., Kaplinski, M., & Hazel, J. E. (2018). How many measurements are required to construct an accurate sand budget in a large river? Insights from analyses of signal and noise. *Earth Surface Processes and Landforms, 44*, 160-178. https://doi.org/10.1002/esp.4489

Gray, J. R., Gartner, J. W., Barton, J. S., Gaskin, J., Pittman, S. A., & Rennie, C. D. (2010). Surrogate technologies for monitoring bed-load transport in rivers. *Sedimentology of Aqueous Systems, 2*, 45–79.

Gray, J. R., & Gartner, J. W. (2009). Technological advances in suspended-sediment surrogate monitoring. *Water resources research*, *45*(4). https://doi.org/10.1029/2008WR007063

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). Model assessment and selection. In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, (pp. 219–260). New York, NY: Springer.

Hoffman, M. D., & Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593–1623. https://arxiv.org/pdf/1111.4246

Holmes, R. R. (2010). Measurement of bedload transport in sand-bed rivers: A look at two indirect sampling methods. *US Geological Survey Scientific Investigations Report, 5091*, 236-252.

Ingram, H., Tarlock, A. D., & Oggins, C. R. (1991), The law and politics of the operation of Glen Canyon Dam. In *Colorado River Ecology and Dam Management: Proceedings of a Symposium* (pp. 10–27), Washington, DC: National Academies Press.

Minckley, W. L. (1991). Native fishes of the Grand Canyon region: an obituary. In *Colorado River Ecology and Dam Management: Proceedings of a Symposium* (pp. 124–177), Washington, DC: National Academies Press.

Kaplinski, M., Hazel, J. E., Parnell, R., Breedlove, M., Kohl, K., & Gonzales, M. (2009). Monitoring fine-sediment volume in the Colorado River ecosystem, Arizona; Bathymetric survey techniques *U.S. Geological Survey Open-File Report 2009-1207*.

Kaplinski, M., Hazel Jr, J. E., Grams, P. E., & Davis, P. A. (2014). Monitoring fine-sediment volume in the Colorado River ecosystem, Arizona: Construction and analysis of digital elevation models *U.S. Geological Survey Open-File Report 2014-1052*. https://dx.doi.org/10.3133/ofr20141052

Kasprak, A., Sankey, J. B., Buscombe, D., Caster, J., East, A. E., & Grams, P. E. (2018). Quantifying and forecasting changes in the areal extent of river valley sediment in response to altered hydrology and land cover *Progress in Physical Geography: Earth and Environment, 42*(6). 739–764. https://doi.org/10.1177/0309133318795846

Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., & Blei, D. M. (2017). Automatic differentiation variational inference. *The Journal of Machine Learning Research*, *18*(1), 430–474. https://arxiv.org/pdf/1603.00788.pdf

Leary, K. (2018). Diamond Creek Repeat Multibeam Data, *SEAD Internal Repository*, http://doi.org/10.5967/M02J6904

Leopold, L. B., & Maddock, T. (1953). The hydraulic geometry of stream channels and some physiographic implications *U.S. Geological Survey Professional Paper 252*.

McElroy, B. J. (2009). Expressions and implications of sediment transport variability in sandy rivers (PhD Dissertation). http://hdl.handle.net/2152/15117

McLean, S. R. (1992). On the calculation of suspended load for noncohesive sediments. *Journal of Geophysical Research: Oceans*, *97*(C4), 5759–5770. https://doi.org/10.1029/91JC02933

Meyer-Peter, E., & Müller, R. (1948). Formulas for bed-load transport. In *IAHSR 2nd meeting*, Stockholm: IAHR.

Mohrig, D., & Smith, J. D. (1996). Predicting the migration rates of subaqueous dunes. *Water Resources Research, 32*(10), 3207–3217.

Molinas, A., and Wu, B. (2000). Comparison of fractional bed-material load computation methods in sand-bed channels. *Earth Surface Processes and Landforms*, *25*, 1045–1068.

Mueller, E. R., Grams, P. E., Schmidt, J. C., Hazel, J. E., Alexander, J. S., & Kaplinski, M. (2014). The influence of controlled floods on fine sediment storage in debris fan-affected canyons of the Colorado River basin. *Geomorphology*, *226*, 65–75. https://doi.org/10.1016/j.geomorph.2014.07.029

Mueller, E. R., Grams, P. E., Hazel, J. E., & Schmidt, J. C. (2018). Variability in eddy sandbar dynamics during two decades of controlled flooding of the Colorado River in the Grand Canyon. *Sedimentary Geology*, *363*, 181–199. https://doi.org/10.1016/j.sedgeo.2017.11.007

Paola, C., & Voller, V. R. (2005). A generalized Exner equation for sediment mass balance. *Journal of Geophysical Research: Earth Surface*, *110*(F4). https://doi.org/10.1029/2004JF000274

Pitlick, J. (1988). Variability of bed load measurement. *Water Resources Research*, *24*(1), 173–177. https://doi.org/10.1029/WR024i001p00173

Rantz, S. E., et al. (1982). Measurement and computation of streamflow, *U.S. Geological Survey Water Supply Paper 2175*.

Rubin, D. M., Tate, G. B., Topping, D. J., & Anima, R. A. (2001). Use of rotating side-scan sonar to measure bedload. *Proceedings of the Seventh Federal Interagency Sedimentation Conference* (pp. 139–144). Reno, NV: U.S. Geological Survey

Rubin, D. M., & Topping, D. J. (2001). Quantifying the relative importance of flow regulation and grain size regulation of suspended sediment transport α and tracking changes in grain size of bed sediment β. *Water Resources Research*, *37*(1), 133–146. https://doi.org/10.1029/2000WR900250

Rubin, D. M., Topping, D. J., Schmidt, J. C., Hazel, J., Kaplinski, M., & Melis, T. S. (2002). Recent sediment studies refute Glen Canyon Dam hypothesis. *Eos, Transactions American Geophysical Union*, *83*(25), 273–278. https://doi.org/10.1029/2002EO000191

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* 2:e55. https://doi.org/10.7717/peerj-cs.55

Schmelter, M. L., Hooten, M. B., & Stevens, D. K. (2011). Bayesian sediment transport model for unisize bed load. *Water Resources Research, 47*(11).

Schmelter, M. L., & Stevens, D. K. (2012). Traditional and Bayesian statistical models in fluvial sediment transport. *Journal of Hydraulic Engineering, 139*(3), 336-340.

Schmelter, M. L., Erwin, S. O., & Wilcock, P. R. (2012). Accounting for uncertainty in cumulative sediment transport using Bayesian statistics. *Geomorphology, 175*, 1-13.

Schmelter, M., Wilcock, P., Hooten, M., & Stevens, D. (2015). Multi-fraction Bayesian sediment transport model. *Journal of Marine Science and Engineering, 3*(3), 1066-1092.

Schmidt, J. C., and J. B. Graf (1990). Aggradation and degradation of alluvial sand deposits, 1965 to 1986, Colorado River, Grand Canyon National Park, Arizona, *US Geological Survey Professional Paper 1493*.

Schmidt, J. C., Kraft, M., Tuzlak, D., and Walker, A. (2016). Fill mead first: A technical assessment. *White Paper No. 1*, Utah State University Quinney College of Natural Resources, Center for Colorado River Studies.

Simons, D. B., Richardson, E. V., & Nordin, C. F. (1965). Bedload Equation for Ripples and Dunes. *U.S. Geological Survey Professional Paper 462-H*.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353

Toffaleti, F. B. (1968) A procedure for computation of the total river sand discharge and detailed distribution, bed to surface. *Corps of Engineers Committee on Channel Stabilization Technical report number 5*.

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*(2), 515–542. https://doi.org/10.1029/1999WR900285

Topping, D. J., Rubin, D. M., Nelson, J. M., Kinzel, P. J.,& Corson, I. C. (2000). Colorado River sediment transport: 2. systematic bed-elevation and grain-size effects of sand supply limitation. *Water Resources Research*, *36*(2), 543–570. https://doi.org/10.1029/1999WR900286

Topping, D. J., Rubin, D. M., Grams, P. E., Griffiths, R. E., Sabol, T. A., Voichick, N. et al. (2010). Sediment transport during three controlled-flood experiments on the Colorado River downstream from Glen Canyon Dam, with implications for eddy-sandbar deposition in Grand Canyon National Park, *U.S. Geological Survey Open-File Report 2010–1128*.

Topping, D. J., & Wright, S. A. (2016). Long-term continuous acoustical suspended-sediment measurements in rivers – Theory, application, bias, and error. *USGS Professional Paper 1823*. https://doi.org/10.3133/pp1823

Topping, D. J., Melis, T. S., Rubin, D. M., & Wright, S. A. (2004). High-resolution monitoring of suspended-sediment concentration and grain size in the Colorado River in Grand Canyon using a laser-acoustic system. In *Proceedings of the Ninth International Symposium on River Sedimentation* (pp. 2507–2514). Yichang, China.

Topping, D. J., Wright, S. A., Melis, T. S., & Rubin, D. M. (2007). High-resolution measurements of suspended-sediment concentration and grain size in the Colorado River in Grand Canyon using a multi-frequency acoustic system. In *Proceedings of the Tenth International Symposium on River Sedimentation* (Vol. 3). Moscow, Russia.

Turowski, J. M., Rickenmann, D., & Dadson, S. J. (2010). The partitioning of the total sediment load of a river into suspended load and bedload: a review of empirical data. *Sedimentology, 57*(4), 1126–1146.

van den Berg, J. H. (1987). Bedform migration and bed-load transport in some rivers and tidal environments. *Sedimentology, 34*(4), 681–698.

van der Mark, C. F., & Blom, A. (2007). A new and wdely applicable bedform tracking tool. *Technical Report*, University of Twente, Faculty of Engineering Technology, Department of Water Engineering and Management.

van Rijn, L. C. (1984). Sediment transport, part II: Suspended load transport. *Journal of Hydraulic Engineering*, *110*(11), 1613–1641. https://doi.org/10.1061/(ASCE)0733-

9429(1984)110:11(1613)
Wong, M., & Parker, G. (2006). Reanalysis and correction of bed-load relation of Meyer-Peter and Müller using their own database. *Journal of Hydraulic Engineering*, *132*(11), 1159–1168. https://doi.org/10.1061/(ASCE)0733-9429(2006)132:11(1159)
Wright, S. A., & Kaplinski, M. (2011). Flow structures and sandbar dynamics in a canyon river during a controlled flood, Colorado River, Arizona. *Journal of Geophysical Research: Earth Surface*, *116*(F1). https://doi.org/10.1029/2009JF001442
Wright, S., & Parker, G. (2004). Flow resistance and suspended load in sand-bed rivers: simplified stratification model. *Journal of Hydraulic Engineering*, *130*(8), 796–805. https://doi.org/10.1061/(ASCE)0733-9429(2004)130:8(796)
Wright, S. A., Topping, D. J., Rubin, D. M., & Melis, T. S. (2010). An approach for modeling sediment budgets in supply-limited rivers. *Water Resources Research*, *46*(10). https://doi.org/10.1029/2009WR008600