Reconstruction of sediment supply and transport behind the Elwha River dams

Claire F. Beveridge[1], Christina Bandaragoda[1], Christopher Curran[2], Erkan Istanbulluoglu[1]

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA
[2]United States Geological Survey, Tacoma, WA

Keywords: sediment transport, hydrology, reservoir sedimentation, Elwha

Abstract: The Elwha River watershed in Washington State's Olympic National Park is predominantly a steep mountainous landscape where dominant geomorphic processes include landslides, debris flows, and gullying. The removal of the two massive hydropower dams along the Elwha River in 2013 marked the largest global dam removal in history. Over the century-long lifespan of the dams, approximately 21 million cubic meters of sediment was impounded behind them. In this study, we test different approaches to reconstructing sediment supply and delivery to the Elwha dams throughout their lifespan. Available observational data that aid in this approach include DEM, channel morphology, meteorology, and streamflow and sediment discharge. We analyze the performance of three numerical modeling constructs — Cases I, II, and III — in predicting historic reservoir sediment yields in the Elwha River watershed. In Case I, we use a sediment rating curve approach and improve upon a published sediment rating curve (Curran et al., 2009) for the Elwha River using a modified, calibrated version of the Wilcock & Crowe (2003) sediment transport equation. In Cases II and III we use a network model that we have developed which connects processes of sediment supply from hillslopes with sediment detachment and transport at stream links of a channel network. The model is driven by streamflow and sediment forcings which can be of varying complexity. In Case II, the network model is driven by stochastic streamflow forcings that require simple input data and parameterization. In Case III, the network model is driven by physically-based model outputs of streamflow from a physically-based distributed model, DHSVM. Using these three cases, we strive to answer the questions: Is using a watershed network model warranted in predicting dam sedimentation? If so, what level of complexity is needed in the driving streamflow and sediment forcings? Is a rating curve approach adequate and robust enough for predicting dam sedimentation, particularly with expected changes in environmental conditions? Cases I, II, and III in this report require further calibration before these questions can fully be answered, however are valuable "proof-of-concept" models. We show that our uncalibrated modeling results along with the insights we have developed provide a promising path in moving forward to answer these questions.

1.0 Introduction

In 2013, the removal of the two massive hydropower dams along the Elwha River in Washington State's Olympic National Park (Figure 1) marked the largest dam removal in history. One metric by which it is the largest dam removal is the volume of accumulated sediment that was available for release – approximately 21 million cubic meters (East et al., 2015; Magirl et al., 2015). Reservoir sedimentation, such as that which occurred to the Elwha River dams, poses a major threat to global hydropower dams and their downstream river reaches. Sediment deposition in reservoirs reduces hydropower capacity and can cause dam operations and maintenance issues, while downstream sediment starvation can be detrimental to ecosystem health and fluvial regimes (East et al., 2015). Given the growing concerns of increased sediment loading with climate and land use/land cover change, there is a need to develop and test geomorphically relevant models for forecasting of reservoir sedimentation in mountainous watersheds. Robust tools and techniques that can aid in predicting and understanding reservoir sedimentation and its impacts can also be valuable to improving the efficiency and sustainability of reservoir and river systems.

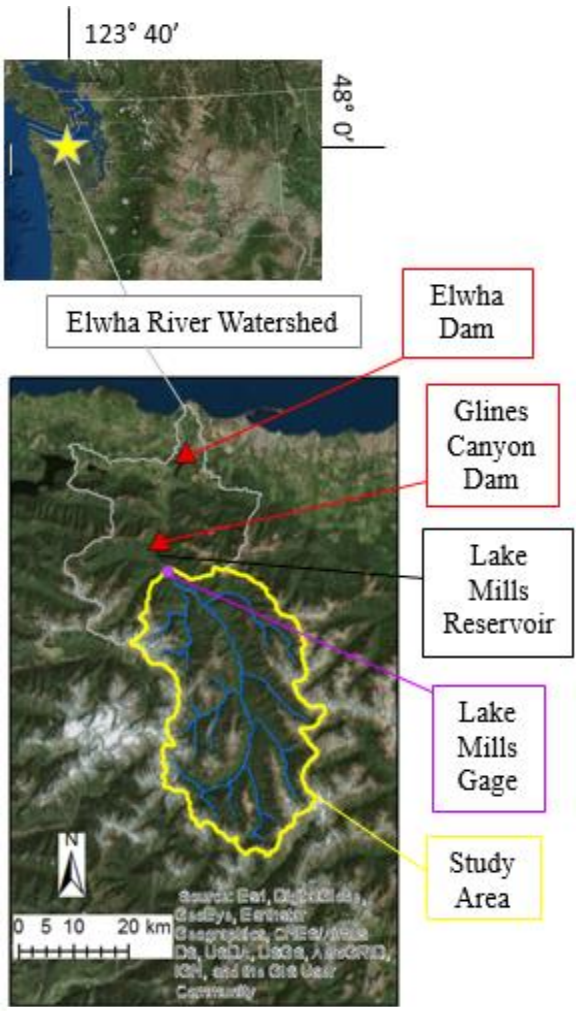


Figure 1: Study location map

Despite the prominence of dam sedimentation challenges, most existing sediment supply and transport models used for dam planning and operations poorly represent the complex conditions in most river and reservoir systems (United States Society on Dams, 2015). These models are based on simplifying assumptions that are especially deficient in steep mountainous regions, where a significant portion of hydropower dams are located. Common models used in engineering application provide either mean annual or annual amounts of sediment supply from hillslopes based on empirical methods such as the Universal Soil Loss Equation (USLE) and its variants (United States Society on Dams, 2015). However, USLE lacks robustness and is poorly suited for the watershed scale, which is most relevant to dam sediment budgeting (Shen et al., 2009). More sophisticated, mechanistic models are adapted from agriculture erosion models such as the Watershed Erosion Prediction Project (WEPP), or developed specifically for scour deposition in reservoirs such as HEC-6 (United States Society on Dams, 2015). However, even these more complex models do not account for climate variability and extremes, have limited representation of sediment transport mechanisms (e.g., wash and rainsplash, which are typical of low-relief agricultural settings), and often require more information about a watershed than what is readily available (i.e., are overparameterized).

On the other hand, relevant models developed for academic research purposes are typically domain-specific (e.g., hillslope geomorphology, hydrology, river network processes) and lack in integrating the appropriate complexity of other domains that are relevant to sediment supply and transport processes. Academic research models are also typically not developed for the appropriate timescales of engineering decision making– decadal to centuries– but are rather either too short (e.g., sub-daily timescale) or too extensive (e.g., millennial). For example, the watershed-scale sediment model of Czuba & Foufoula-Georgiou (2014) provides robust characterization of hillslope and fluvial processes, however assumes it uniform hydrology. This simplification of hydrologic processes undermines the significant role that dynamic hydrologic events (e.g., floods, low flows) play in sediment transport. Hence, there is a need for sediment supply and transport models which bridge the gaps between academic research, engineering application, and relevant sub-domains in each of these sectors (Figure 2) in order to better aid in dam sedimentation planning and management.

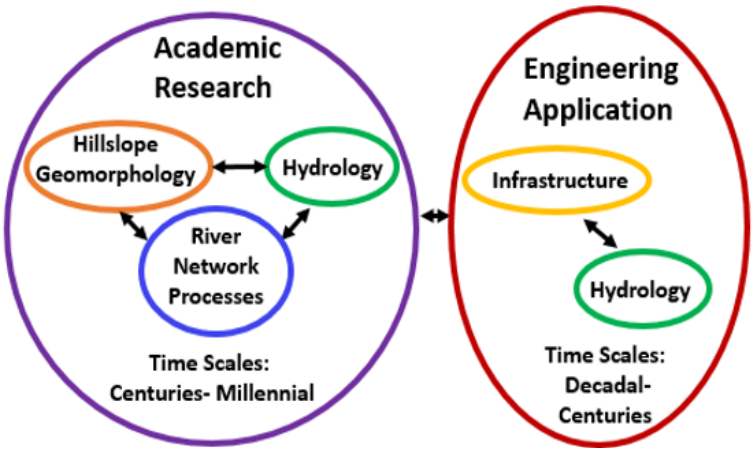


Figure 2: Overarching motivation for this study

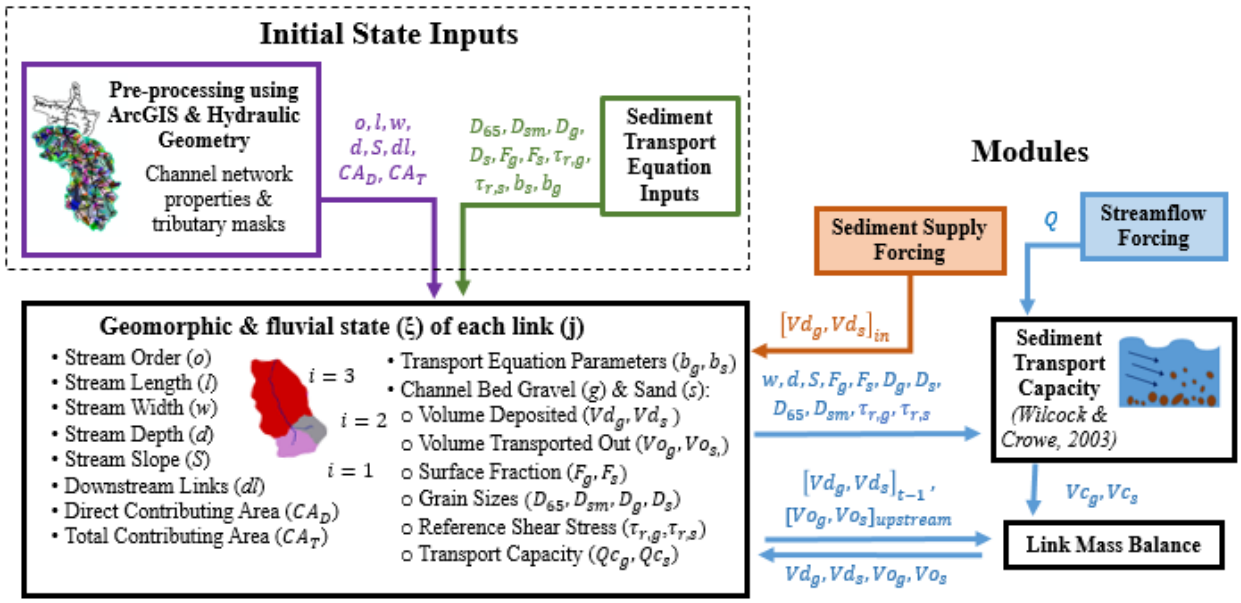
Historical data on Elwha River hydrology and dam sedimentation provides a unique opportunity to test sediment supply and transport models that can be used for sustainable planning and management of dams. In this study, we present the performance of three numerical modeling constructs — Cases I, II, and III — in predicting historic coarse sediment yields in the upstream Elwha River dam throughout its lifespan (Table 1). From Case I to Case III, the model construct becomes increasingly complex. In Case I, we use a sediment rating curve approach and improve upon a published sediment rating curve for the Elwha River using a modified, calibrated version of the Wilcock & Crowe (2003) sediment transport equation. In Cases II and III we use a network model (Figure 3) that we have developed which connects processes of sediment supply from hillslopes with sediment detachment and transport at stream links of a channel network. The network model is driven by streamflow and sediment supply forcings (i.e., inputs to the model which come from an external source and drive the processes in the model) which can be of varying complexity. In both cases, the sediment supply forcings are conducted stochastically. In Case II, the network model is driven by streamflow forcings which are also stochastic and require simple input data and parameterization. In Case III, the network model is driven by streamflow outputs from a physically-based hydrology model (Distributed Hydrology Soils Vegetation Model (DHSVM; Wigmosta, Vail, & Lettenmaier, 1994). Due to time limitations, we were not able to fully calibrate each of the three model constructs nor address all of the technical issues that arose. Therefore, in this paper we present "proof-of-concept" models along with preliminary results and inferences. We provide the main issues we encountered and recommendations for next steps.

Table 1: Three model constructs considered in this analysis

Case	Streamflow Forcing	Sediment Forcing
I: Sediment Rating Curve (No Network Model)	Streamflow rating curve at the Lake Mills gage.	None. Assumes bedload transport is always at bedload transport capacity.
II: Network-Based Model with *Stochastic* Streamflow Forcing	Streamflow rating curve at the Lake Mills gage is scaled across entire river network under assumption of uniform hydrology. Daily streamflow for each year (i.e., 365 values) are randomly sampled from rating curves, and then the 5 highest daily streamflow values are used as model forcing.	Each stream network has a direct[1] tributary area that accumulates sediment at the assumed denudation rate. Each tributary area has randomly occuring "mass wasting" events where sediment that has accumulated since the last mass wasting event is deposited into the stream.
III: Network-Based Model with *Physically-Based* Streamflow Forcing	A physically-based distributed hydrology model (DHSVM) is run for the entire river network at sub-daily timestep. Results are aggregated to daily streamflow which are then used as model forcing.	

Notes

1. The direct tributary area in this context is the total contributing area of a stream links minus the contributing area of all upstream links.



Note: In this application, all "Initial State Inputs" have static values for all timesteps (t) but may vary for each stream link (j). The parameters exchanged by the modules are for a given j and t unless: (1) they have subscript *t-1*, which indicates that they are values for a given j at the previous t; or (2) they have subscript *upstream*, which indicates that they are values for a given t at upstream j.

Figure 3: Network Model Overview

In comparing the methods and results for each model construct, we strive to answer the questions: Is using a watershed network model warranted in predicting dam sedimentation? If so, what level of complexity is needed in the driving streamflow and sediment supply forcings? Is a rating curve approach adequate and robust enough for predicting dam sedimentation, particularly with expected changes in environmental conditions? We are not able to fully answer these questions in this study. However, with our proof-of-concept models, we are able to demonstrate when and why a network model would be necessary for robust and reliable predictions of reservoir sedimentation.

1.1 Study Area

The Elwha watershed drains the northern slopes of the Olympic Mountains of western Washington State and flows northward into the Strait of Juan de Fuca (Figure 1). The basin encompasses an area of 833square kilometers (km^2) and reaches up to 2,250 meters (m) in elevation. The Elwha watershed has a maritime climate with dry summers and wet winters. Mean annual precipitation ranges from 1,400 millimeters (mm) at low elevations to 5,600 mm at high elevations. Construction of the two dams, Glines Canyon and Elwha (Figure 1), was completed in years 1913 and 1927, respectively. Removal of both dams began in 2011 and was completed in 2013. The upstream Glines Canyon Dam – which we focus on in this study – was a 64 m tall concrete arch dam and had the capacity for 50 million cubic meters (m^3) of storage water in the Lake Mills reservoir at the time of construction. At the time of removal, 16.1 ± 2.4 million m^3 of sediment was stored in the Lake Mills reservoir, of which 44% (by volume) was estimated as fine-grained (silt and clay) material and 56% was estimated as coarse-grained (sand-sized and coarse) material (Warrick et al., 2015).

This study focuses on modeling sediment supply and transport upstream of the United States Geological Survey (USGS) gaging station no. 12044900, Elwha River above Lake Mills (Lake Mills gage) over the 84-year lifetime (1927 to 2011) of the Glines Canyon dam. The Lake Mills gage was located approximately 400 m upstream of the former Lake Mills reservoir and 4.3 river kilometers (rkm) upstream of the former Glines Canyon Dam (Curran et al., 2009). The Glines Canyon dam was located 21.6 river km upstream of the Elwha River mouth and had a tributary area of 636 km^2 (approximately 76% of the entire Elwha watershed), whereas the tributary area of former Lake Mills gage is approximately 513 km^2 (62% of the Elwha River basin; Curran, et al., 2009). Although Lake Mills gage tributary area is ~81% of the Glines Canyon dam tributary area, we focus on modeling the Lake Mills gage tributary area since hydrology and sediment observations have historically been collected at the Lake Mills gage and allow for clear comparison (Section 2.1) between observations and model outputs.

The study area lies within the Olympic Subduction Complex, which is primarily composed of metasedimentary phyllites and schists (Warrick et al., 2011). The Olympic Mountains, and hence the Elwha watershed, are reported to be in a tectonic steady-state, although rock uplift rates vary spatially (Montgomery & Brandon, 2002). Over geologic time, denudation rates in the study area range from 0.6 millimeters per year (mm/yr) in the mountainous uplands and decrease to 0.3 mm/yr at the coast (Batt et al., 2001; Brandon et al., 1998). The modern erosion rates of the study area based on an inventory of erosion rates is 0.18 mm/yr, whereas the erosion rate based on the volume of sediment accumulated behind the Glines Canyon dam over 61 years (prior to

1998) is >0.11 mm/year (Brandon et al., 1998). Channel morphology at the Elwha River headwaters is strongly influenced by the steep slopes, which are maintained by the uplift rates and the dominant erosional processes that include rockfalls, shallow and deep landslides, and debris flows (Warrick et al., 2011). These fast erosional processes occur mostly during storm events and supply the majority of sediment to headwater valleys (Buffington & Woodsmith, 2003). As the Elwha River continues downslope, channel morphology is increasingly dominated by fluvial processes. The Elwha River alternates between narrow bedrock canyons and wider alluvial reaches for much of its length (Randle et al., 2014). This alternating boundary composition leads to differential behavior in sediment transport along the length of the river. In the study area, the 50^{th} percentile daily flow (i.e., flow that is exceed 50 percent of the time) is 29 cubic meters per second (cms), 5^{th} percentile daily flow is 91.7 cms, and 95^{th} percentile daily flow is 8.86 cms (USGS, 2016). Since the Elwha watershed is within Olympic National Park, it is minimally impacted by humans (e.g., land use changes).

2.0 Methods

The methods conducted in this analysis are primarily data analysis and computational modeling of bedload transport. Bedload is the thin layer of particles that frequently maintain contact with the bed of the stream channel, and mostly consists of sand and gravel. The other components of sediment transported in rivers are suspended load and washload, which primarily remain in suspended in the stream water due to turbulent motions, and primarily consist of silts and clays. While suspended load typically dominates the transport in rivers, we focus on bedload in this study since we expect bedload to play an important role in sediment transport and channel morphology in the Elwha River, particularly in the upper reaches where the study area is. In addition, modeling of bedload can inform our understanding of the Elwha River system geomorphology and inform the future inclusion of suspended load and washload in the model. The ability to model suspended load and washload is also limited since these sediment loads are primarily dominated by sediment supply rather than transport capacity. By contrast, bedload in mountainous watersheds is typically influenced by the interplay between supply and transport capacity, which can more readily be modeled. Therefore, significant gains can be made in understanding reservoir sedimentation using bedload modeling.

The hydrologic and sediment data used in this study was obtained from the USGS and publications, as described in Section 2.1. These data inform the selection and calibration of bedload transport equations as well as other hydrologic modeling parameters. The parameterization of stream channel geometry is discussed in Section 2.2. The development of the bedload rating curve (Case I) is discussed in Section 2.3. An overview of the network model development is discussed in Section 2.4. The application of stochastic hydrology with the network model (Case II) is discussed in Section 2.5 and the application of physically-based hydrology with the network model (Case III) is discussed in Section 2.6.

2.1 Data and Previous Studies

The data used in this analysis was collected at the Lake Mills gage, which was in operation from March 26, 1994 to November 12, 2015 (USGS, 2016). The Lake Mills gage was located within a

bedrock-confined river reach known as Rica Canyon, which has a length of approximately 1.9 rkm and widths ranging from 12 to 40 m (Childers et al., 2000; Duda et al., 2008).

Historical data collected at the Lake Mills gage that are relevant to this study include measurements (collected at varying temporal frequencies) of channel geometry (i.e., channel width, depth, cross-sectional area); streamflow velocity; streamflow discharge; bedload discharge and grain size distribution; and grain size distributions of the wetted channel surface, subsurface, and a gravel bar located nearby. Figure 1 shows the location of the Lake Mills gage and Table A1 of the Appendix provides details of historical data used in this analysis (sample dates, number of samples, sampling equipment/techniques and additional notes) that are mentioned below.

Throughout the period of the Lake Mills gage operation, streamflow discharge was monitored by the USGS daily from March 26, 1994 to May 31, 1998 and February 18, 2004 to September 30, 2011. Streamflow channel measurements (i.e., channel width, depth, cross-sectional velocity) were also collected intermittently from May 11, 1994 to October 13, 2011. These USGS data are publicly available online (USGS, 2016). Streamflow and sediment data were collected for two studies conducted by the USGS in anticipation of dam removal. The first study, "Hydrologic Data Collected During the 1994 Lake Mills Drawdown Experiment, Elwha River, Washington" (Childers et al., 2000) describes a reservoir drawdown analysis conducted at the Lake Mills reservoir in 1994 which aimed to determine the erodability and size gradation of sediments in Lake Mills reservoir, the impact of lake drawdown on sediment transport, and the impact of sediment erosion on water quality (Childers et al., 2000). This study largely focused on measuring sediments accumulated in the Lake Mills reservoir, however streamflow, bedload, channel bed surface sediments, and gravel bar sediments were also measured at the Lake Mills gage. The second study, "Estimates of sediment load prior to dam removal in the Elwha River, Clallam County, Washington" (Curran et al., 2009), describes an analysis which aimed to estimate the sediment load behind the two Elwha River dams prior to their removal. To achieve this, Curran et al. (2009) developed equations using empirical ordinary least squares regression analysis for bedload (as well as suspended load) as a function of streamflow. Bedload data used in this analysis was collected by USGS in water years (WY) 1994 (Childers et al., 2000) and 1995-98 (Curran et al., 2009). The resulting empirical bedload rating curve equation is

$$Q_B = 1.13\ x\ 10^{-2} x\ Q_W^{2.41} \qquad (1)$$

where $\boldsymbol{Q_B}$ is the bedload discharge in megagrams per day (Mg/day) and $\boldsymbol{Q_W}$ is water discharge in cms.

Randle et al. (2014) provide estimates of reservoir sedimentation volumes in July 2010, which is just prior to the start of dam removal. They estimated this volume by developing digital elevation models (DEMs) of the predam valley bottoms and pre-removal reservoir surface, and then taking the difference of the DEMs in the full pool areas and upstream river segment impacted by reservoir sedimentation. They also determined percentages of fine (silt and clay <0.062 mm) and coarse (sand, gravel, and cobble >0.062 mm) sediment in reservoirs based on gradation data from Gilbert and Link (1995). The estimated volume of Lake Mills reservoir sedimentation in the coarse fraction was $9.0 \pm 1.9 \times 10^6$ m^3. The coarse sediment volume is greater than the

volume of bedload deposition for two main reasons: (1) The coarse sediment volume includes some of the sand-size suspended sediment load; and (2) The coarse sediment volume also includes inputs from the stream reaches downstream of Lake Mills gage but upstream of the Glines Canyon Dam, however we only model bedload discharged at the Lake Mills gage in this study. Thus, we cannot directly compare the coarse sedimentation volume to the outputs of each model construct. However, we can still use the estimate from Randle et al. (2014) as an upper threshold for our model outputs as well as an order-of-magnitude comparison. Randle et al. (2014) also provide an estimate of the bulk density of coarse sediment measured in the dams as 1.71 grams per cubic centimeter (g/cm^3), which is used in this study to convert the volume of bedload transported to the volume of bedload reservoir sedimentation (Section 2.3).

2.2 Hydraulic Geometry

Hydraulic geometry relates stream channel geometry (e.g., width, depth) to discharge. Hydraulic geometry can be used in sediment transport modeling when estimates of channel geometry are needed for inputs to sediment transport equations, however they are unknown for the given discharge and/or location of interest. Hydraulic geometry relationships come in two forms: (1) "at-a-station" hydraulic geometry which relates temporal variations in streamflow and channel geometry at a single cross-section, and (2) "downstream" hydraulic geometry which relates spatial variation in channel geometry along a stream network at a reference discharge (Knighton, 1998). In order to develop the sediment rating curve for the Lake Mills gage in Case I, "at-a-station" hydraulic geometry relationships were developed using channel geometry and streamflow data collected at the Lake Mills gage from 1994-1998. Power regression, which is commonly used for at-a-station hydraulics geometry, was used to develop relationships between streamflow discharge and depth as well as streamflow discharge and velocity. However, the streamflow channel width only varied slightly since Lake Mills gage is located in a bedrock canyon with stable banks. Therefore, a relationship between streamflow discharge and width was developed using a linear regression (which allows width to change much less rapidly than a power regression relationship). Channel geometry measurements collected below a discharge of 19 cms were not included in the width, depth, and velocity regressions since they reduced the goodness-of-fit of the hydraulic geometry relationship, particularly for high flows. The resulting at-a-station hydraulic geometry parameterizations for channel depth (d), velocity (u), and width (w) at the Lake Mills gage are

$$d = 0.03\ x\ Q_W^{0.37} \qquad (2a)$$
$$u = 0.08\ x\ Q_W^{0.63} \qquad (2b)$$
$$w = 0.003\ x\ Q_w + 40.7 \qquad (2c)$$

Downstream hydraulic geometry relationships are used for streams of the study area network (further discussed in Section 2.4) for Cases II and III. The relationships are taken from Magirl and Olsen (2009) which presents hydraulic geometry parameterization based on their analysis of navigable rivers throughout Washington State. The hydraulic geometry relationships are provided for a stream's mean annual discharge ($\boldsymbol{Q_{w,mean}}$, in units of cms**)** and w as well as d as

$$d = 0.26\ x\ Q_{W,mean}^{0.4} \qquad (3a)$$
$$w = 7.35\ x\ Q_{W,mean}^{0.5} \qquad (3b)$$

$\boldsymbol{Q_{w,mean}}$ for streams throughout the study area channel network are computed differently for Cases II and III in order to be consistent with each case (further discussed in Sections 2.5 and 2.6). For Case II, $\boldsymbol{Q_{w,mean}}$ was computed for the Lake Mills gage using the observed streamflow record, and converted to units of millimeters per day (mm/day) with respect to the area of the Lake Mills gage tributary. Then, under the assumption of uniform hydrology, the mean annual discharge was scaled for each stream link with respect to its contributing area (see detailed approach in Section 2.6). For Case III, $\boldsymbol{Q_{w,mean}}$ was computed at each stream link in the network from the modeled DHSVM data, which included daily streamflow from years 1915-2011 at each link in the network. We tested equations (3a) and (3b) at the Lake Mills gage using $\boldsymbol{Q_{w,mean}}$ calculated for Cases II and II and compared the results to observed channel width and depth. We found that our approach for using the Magirl and Olsen (2009) parameterization and $\boldsymbol{Q_{w,mean}}$ was acceptable and results for Cases II and III were within reasonable difference of each other. We computed the channel cross-sectional area ($\boldsymbol{A}$) assuming a rectangular channel ($A=d$ x w).

2.3 Case I: Bedload Rating Curve (No network model)

As an alternative to the observation-based rating curve of Curran et al., 2009 (equation [1]), mechanistic sediment transport equations can be used to predict bedload transport capacity. For this analysis, we parameterize sediment transport equations using observations of channel geometry, channel bed surface sediments, streamflow, and bedload (discussed in Section 2.1). The sediment transport equations and explanation for their implementation are discussed below, and the values input to the constants and parameters in the equations are presented in Table 2.

Table 2: Constants and parameters used in bedload transport capacity equations at Lake Mills gage

Parameter	Units	Value	Source/Basis For Value
Model Constants			
Density of water (rho)	kg/m^3	1000	Property of water at 20 deg C
Gravitational acceleration (g)	m/s^2	9.81	Physical constant
Specific gravity of sediment (sg)	--	2.65	Specific gravity of quartz
slope (S)	--	0.014	Computed using ArcGIS and DEM
river channel shape	--	Rectangular	Childers et al., 2000
Drainge Area of Lake Mills streamflow gage	km^2	513	USGS, 2016; verified in ArcGIS
Bulk density of coarse sediment (sand, gravel, cobbles)	g/cm^3	1.71	Determined by Randle et al., 2014 for Elwha River and reservoir sediments based on literature and data
Lifespan of Glines Canyon dam	years	84	Glines Canyon dam was constructed in years 1925-1927 (Randle et al., 2014) and removed in years 2011-2013, and thus 84 years of sediment accumulation were assumed.
Computed Parameters			
65th percentile of channel bed grain size (D_{65})	mm	25.8	Average from 5 sampling points for the single wetted channel bed surface sample collected on April 9, 1994 and two gravel bar surface samples collected on June 21, 1994 (Childers et al., 2000).
Mean channel bed grain size (D_{sm})	mm	21.6	
Median gravel grain size on channel bed (D_g)	mm	22.2	
Median sand grain size on channel bed (D_s)	mm	1.18	
Fraction of sand on channel bed (F_s)	%	34	

We use a "surface-based" sediment transport modeling approach, which is common for gravel-bedded rivers such as the Elwha River. The underlying theory of surface-based sediment transport models for gravel-bed rivers is that bedload transport rate of mixtures should be based on the availability of each sediment grain size range on the channel surface layer (Parker, 1990). This differs from transport models that depend on a single representative grain size (e.g., Meyer-Peter & Müller, 1948). Wilcock & Kenworthy (2002) describe the influence of grain size on transport rate as a competition between absolute and relative grain sizes of sediment. The effect of absolute size is that mobility of sediment grains decreases with increasing grain size, whereas the effect of relative size is that the transport of larger grains increases relative to the transport of smaller grains. A popular surface-based transport formula is that of Wilcock & Crowe (2003) for mixed sand/gravel sediment, which computes transport rate for each grain size fraction $\boldsymbol{i}$ on the channel bed surface. The formula is

$$\boldsymbol{q_{bi}} = \frac{\boldsymbol{W^*_i F_i u_*^3}}{\boldsymbol{(sg-1)g}} \tag{4a}$$

$$\boldsymbol{W^*_i} = \begin{cases} \boldsymbol{0.002\phi^{7.5}} & \boldsymbol{\phi < 1.35} \\ \boldsymbol{14\left(1-\frac{0.894}{\phi^{0.5}}\right)^{4.5}} & \boldsymbol{\phi \geq 1.35} \end{cases} \tag{4b}$$

where, for each grain size fraction $\boldsymbol{i}$ on the channel bed surface: $\boldsymbol{q_{bi}}$ is the volumetric bedload transport rate; $\boldsymbol{W^*_i}$ is the dimensionless bedload transport parameter; $\boldsymbol{F_i}$ is the proportion of the grain size on the channel bed surface; and $\boldsymbol{\phi = \tau/\tau_{ri}}$, where $\boldsymbol{\tau_{ri}}$ is the reference shear stress of grain size fraction $\boldsymbol{i}$, specifically defined as the value of $\boldsymbol{\tau}$ at which W^*_i is equal to a reference value of 0.002 (Parker et al., 1982; Wilcock, 1997). $\boldsymbol{sg}$ is the ratio of sediment to water density and $\boldsymbol{g}$ is the acceleration of gravity. $\boldsymbol{u_*}$ is shear velocity, $\boldsymbol{u_* = [\tau/\rho]^{0.5}}$, where $\boldsymbol{\rho}$ is the density of water and $\boldsymbol{\tau}$ is the portion of the boundary shear stress that is acting on bed particles and contributing to sediment transport.. Wilcock & Crowe (2003) present the equations for computing $\boldsymbol{\tau_{ri}}$ as:

$$\frac{\tau_{r_i}}{\tau_{r_{sm}}} = \left(\frac{\boldsymbol{D_i}}{\boldsymbol{D_{sm}}}\right)^{\boldsymbol{b}} \tag{5a}$$

$$\boldsymbol{b} = \frac{\boldsymbol{0.67}}{\boldsymbol{1+\exp\left(1.5-\frac{D_i}{D_{sm}}\right)}} \tag{5b}$$

$$\boldsymbol{\tau^*_{r_{sm}} = 0.021 + 0.015\exp(-20\,F_S)} \tag{5c}$$

$$\boldsymbol{\tau^*_{r_{sm}}} = \frac{\boldsymbol{\tau_{r_{sm}}}}{\boldsymbol{(sg-1)\rho g D_{sm}}} \tag{5d}$$

where $\boldsymbol{D_i}$ is the median grain size for each size fraction *i*, $\boldsymbol{D_{sm}}$ is the mean grain size on the channel surface, $\boldsymbol{\tau_{r_{sm}}}$is the reference shear stress for the mean grain size, $\boldsymbol{\tau^*_{r_{sm}}}$is the dimensionless $\tau_{r_{sm}}$, and $\boldsymbol{F_S}$ is the fraction of sand on the channel bed surface.

The equation for $\boldsymbol{\tau}$ is not prescribed by Wilcock & Crowe (2003). Therefore, in this study, $\boldsymbol{\tau}$ is computed using the equation presented by Wilcock (2001)

$$\boldsymbol{\tau = 0.052\ \rho\ (gSD_{65})^{0.25} U^{1.5}} \qquad (6)$$

where $\boldsymbol{S}$ is the slope of the channel bed, $\boldsymbol{U}$ is the streamflow velocity, and $\boldsymbol{D_{65}}$ is the 65th percentile grain size of sediment on the channel bed.

The Wilcock & Crowe (2003) equations (4a-b) and (5a-d) can be applied over a wide range of divided grain size fractions. While applying models on an extensive, finely divided grain-size distribution can capture the variation and interaction among many transport rates, it requires substantial effort in computational and field efforts (Wilcock & Kenworthy, 2002). A simplified approach is to use a two-fraction model that only differentiates between sand (represented by "s" in place of subscript "i" in equations) and gravel (represented by "g" in place of subscript "i" in equations). Wilcock & Kenworthy (2002) describe the two-fraction model as being of practical significance because it represents the difference in the behavior of sand and gravel, similarity of transport rates of different size fractions within the sand and gravel fractions, consistency of variation of sand and gravel with the bed sand content, and the observation that the fines content of a river bed tends to be more transient than the gravel/cobble framework of the bed. Hence, for this study, the Wilcock & Crowe (2003) equations (4a-b) and (5a-d) are applied using the two-fraction approach.

To compute $D_{65}, D_{sm}, D_g, D_s, F_g,$ and F_s at the Lake Mills gage, we computed each of these values at the 5 sampling points for the single wetted channel bed surface sample collected on April 9, 1994. We also computed these values for the two gravel bar surface samples collected on June 21, 1994, since we assume that at high flows (when most sediment is transported) the gravel bars are likely to be submerged and part of the channel bed surface. Then, for each respective parameter of $D_{65}, D_{sm}, D_g, D_s, F_g,$ and F_s, we use the average of the seven samples as the input values for the Wilcock & Crowe (2003) equations.

Although Wilcock & Crowe (2003) prescribes equations (5a-d) to compute τ_{ri}, there have been alternative methods and rationale proposed in literature for computing τ_{ri} (e.g., Wilcock & Kenworthy, 2002; Gaeuman et al., 2009; Gasparini et al., 2004). Of the alternative methods that we tested, we found the equations of Wilcock & Crowe (2003) and Gasperini et al. (2004) to perform the best when compared to bedload observations. The Gasperini et al. (2004) formulation is not detailed in this report because of its complexity, however it can be found in the respective publication. We also tested equations (4a-b) and (5a-d) with manually calibrated τ_{ri} values. The calibration was conducted by adjusting values of τ_{ri} in equation (4a-b) and comparing to bedload observations (visually in plots as well as using error metrics) until we found what we considered to be the best fit. We computed bedload predictions using the streamflow and channel geometry measurements collected at the time of the bedload observations. If channel measurements were not reported for an observation, we computed the channel measurements using our derived at-a-station hydraulic geometry relationships. The parameters used for computing the bedload observations and predictions are shown in Table 2.

We also develop bedload rating curves by applying equations (4a-b) and (5a-d) with each of the three approaches for computing τ_{ri} to the Lake Mills streamflow rating curve. We assumed that the channel bed $D_{65}, D_{sm}, D_g, D_s, F_g, F_s$, and τ_{ri} did not vary with streamflow. We then integrated the rating curve for the duration of the Glines Canyon dam lifespan (84 years) to compute the total estimated amount of coarse sediment transported. Finally, we converted this estimated transported volume ($\boldsymbol{V_{transported}}$) to the estimated volume of the reservoir filled by coarse sediment ($\boldsymbol{V_{reservoir}}$) as

$$\boldsymbol{V_{reservoir}} = \frac{\boldsymbol{BD_{coarse,reservoir}\ x\ V_{transported}}}{\boldsymbol{sg}} \qquad (7)$$

Where $\boldsymbol{BD_{coarse,reservoir}}$ is the estimated bulk density of coarse sediment measured in the dams, 1.71 g/cm^3 (Randle et al., 2014). As done in standard practice for sediment transport computations, the density of $V_{transported}$ is assumed to be the specific gravity of quartz (***sg***) which is 2.65, times the density of water (1 g/cm^3).

2.4 Network Model

2.4.1 Overview

The network model is designed so that it can be applied in any watershed, however in this study we focus solely on model parameterization and application in the study area. The network model is written in Python programming language, and thus is open source. An overview of the network model framework as it is applied in this study is shown in Figure 3.

A key component of the network model is the "state matrix" that stores the geomorphic and fluvial state (***ξ***) of each stream network link (***j***) at a given timestep (***t***). The ξ matrix is initially developed using ArcGIS outputs (described below) as well as hydraulic geometry relationships (Section 2.2) to get, for each *j*, the stream order (***o***), channel length (***l***), *w*, *d*, *S*, downstream links (***dl***), total contributing area ($\boldsymbol{CA_T}$), and direct contributing area ($\boldsymbol{CA_D}$; i.e., $\boldsymbol{CA_T}$ of *j* minus $\boldsymbol{CA_T}$ of all streams upstream of *j*). In this application of the model, these initial state inputs remain static for all *t* but may vary between *j*. The ξ matrix is also initially developed with computed inputs to sediment transport equations (described below). While the model is running, the ξ matrix is updated by two components: (1) the "sediment supply forcing" component which deposits sediment from hillslopes into the stream channel ($\boldsymbol{\left[Vd_g, Vd_s\right]_{in}}$); and (2) the "link mass balance" component which computes, for each *j* and *t*, the volume of sand and gravel deposited on the channel bed ($\boldsymbol{Vd_g, Vd_s}$) as well as the volume of sand and gravel transported out (i.e., downstream) ($\boldsymbol{Vo_g, Vo_s}$). The link mass balance component depends on inputs from the $\boldsymbol{\xi}$ matrix of $\boldsymbol{Vd_g}$and $\boldsymbol{Vd_s}$ in the previous *t* ($\boldsymbol{[Vd_g, Vd_s]_{t-1}}$) as well as $\boldsymbol{Vo_g}$and $\boldsymbol{Vo_s}$ from upstream *j* ($\boldsymbol{[Vo_g, Vo_s]_{upstream}}$). The link mass balance also depends on inputs of gravel and sand transport capacity ($\boldsymbol{Vc_g, Vc_s}$) for a given *j* and *t* which come from the "sediment transport capacity" component. The sediment transport capacity component depends on the streamflow (***Q***) for each *j* and *t* from the "streamflow forcing" component as well as inputs of *w*, *d*, *S*, F_g, F_s, D_g, D_s, D_{65}, D_{sm}, $\tau_{r,g}$, and $\tau_{r,s}$ for a given *j* from the ξ matrix. In this application, we use the equations of Wilcock & Crowe (2003) (4a-b, 5a-d) to compute sediment transport capacity for each *j* and *t*.

The sediment supply forcing, streamflow forcing, sediment transport capacity, link mass balance, and initial state input setup are all "modular" pieces that can be individually modified such that they use different approaches, models, levels of complexity, etc. The network model can also be run for different lengths of *t* (e.g., annually, daily, sub-daily). In this study, we demonstrate this feature of the model by using different approaches for streamflow forcing in Cases II and III. Although the remaining components are consistent between the two cases, they can similarly be varied.

Note that the current version of the model does not consider the exchange of sediment between the channel surface and bedload substrate and it results in D_{65}, D_{sm}, D_s, D_g, b_g, b_s, F_g, F_s, $\tau_{r,g}$, and $\tau_{r,s}$ remaining static over time for all *j*. In future versions of the model it would be ideal to incorporate exchange between the channel bed and bedload substrate in the link mass balance component such that there would be dynamic evolution of D_{65}, D_{sm}, D_s, D_g, b_g, b_s, F_g, F_s, $\tau_{r,g}$, and $\tau_{r,s}$ and more insightful results.

2.4.2 Initialization

For the first step of ArcGIS work and Python pre-processing, a 30-meter DEM of the Puget Sound region was downloaded online from Finlayson et al. (2000) and the study area was delineated. The study area stream network along with the *o*, *l*, *S*, and *dl* for each *j* were derived using a script from Pacific Northwest National Laboratory (2017). This script, which uses Python and ArcGIS, was developed for processing DEM inputs for the DHSVM model. It requires an input of a minimum stream contributing area as well as soil depth range, which we set as 1 km^2 and 0.75 km to 3 km, respectively. Then, we used the stream network output- which has 293 links- to compute CA_T and CA_D for each *j*. *w* and *d* of the links were computed differently for Case II and III, as discussed in Section 2.2.

The application of the Wilcock & Crowe (2003) equation as a two-fraction model (Section 2.3) requires the following for each link of the network: D_{65}, D_{sm}, D_g, D_s, F_g, and F_s which are used to compute b_s and b_g (5a, 5b) as well as $\tau_{r,g}$ and $\tau_{r,s}$ (5c, 5d). Although estimates of these properties are available at the Lake Mills gage (Section 2.3), we had to estimate these properties for all *j* in the channel network. We used the classic Sternberg's law as a basis for our estimations, which approximates fining of sediments while moving downstream as

$$\boldsymbol{D = D_0\, exp(-\alpha L)} \qquad (8)$$

where $\boldsymbol{D_0}$ is the upstream grain diameter, $\boldsymbol{D}$ is the downstream grain diameter, $\boldsymbol{L}$ is the distance in kilometers (km) between upstream and downstream grains (with diameters D_0 and *D*, respectively), and $\boldsymbol{\alpha}$ (1/km) the rate of change of the grain size diameter. Surian (2002) presents a range of 0.001-1 for α in gravel-bedded rivers, and based on lithology of the study area as well as our own testing we approximate α as 0.02 for this study. Since D_g and D_s are assumed to be known at the Lake Mills gage from observations ($\boldsymbol{D_{i,LM}}$), we approximate upstream and downstream D_g and D_s based on distance of a given *j* from the Lake Mills gage ($\boldsymbol{L_{LM,j}}$) using the following relationships:

Streams upstream of Lake Mills gage: $\boldsymbol{D_{i,j} = \frac{D_{i,LM}}{exp(-0.02\, L_{LM,j})}}$ (9a)

Streams downstream of Lake Mills gage: $\boldsymbol{D_{i,j} = D_{i,LM}\, exp(-0.02\, L_{LM,j})}$ (9b)

We also assume that F_i, b_i, and τ_{ri} follow the exponential form of Sternberg's law since these parameters are related to grain sizes of the channel bed. Hence, the form of (9a) and (9b) are used in computing D_{65}, D_{sm}, D_s, D_g, b_g, b_s, F_g, F_s, $\tau_{r,g}$, and $\tau_{r,s}$ for all *j*. While this is a relatively simple and reasonable approach to parameterization of the channel network, it would be ideal to refine this approach for future applications of the model. The sediment supply and streamflow forcings were then initialized following the steps discussed in Sections 2.5, 2.6, and 2.7, which then allowed for the model to be ready to run.

2.4.3 Run Sequence

When the model is run, it starts a given timestep by first running the sediment forcing component for each *j*. Hence, $Vd_{g,in}$ and $Vd_{s,in}$, if any, is deposited onto each *j* and added to Vd_g and Vd_s in the ξ matrix. Then, the streamflow, sediment transport capacity, and network mass balance sequences starts at the lowest *o* (i.e., at the uppermost *j* of the watershed) and runs the transport capacity and link mass balance for all *j* of the same *o* before moving to the next highest *o*. This is done for all *o* before moving to the next *t*. The streamflow, sediment transport capacity, and link mass balance starts with the *Q* for a given *j* and *t* being input to the sediment transport capacity component. This sediment transport capacity component outputs Vc_g and Vc_s which are fed into the link mass balance along with $Vd_g, Vd_s, Vo_{g,upstream}$ and $Vo_{s,upstream}$ from the ξ matrix. For all *o* greater than one, $Vo_{g,upstream}$ and $Vo_{s,upstream}$ is deposited onto the stream bed before Vo_g and Vo_s are computed for a given *j*. Vo_g and Vo_s cannot be more than Vc_g and Vc_s nor Vd_g and Vd_s. Therefore, Vo_g is the minimum of Vc_g and Vd_g, and Vo_s is the minimum of Vc_s and Vd_s. Vo_g and Vo_s are then transported to the next *j* in the network and any remaining Vd_g and Vd_s will remain on the channel bed of *j* until the next *t*.

2.5 Stochastic Sediment Supply Forcing

As described above, the network model can be applied with sediment supply forcing of varying complexity. However, in this analysis, only stochastic sediment supply forcing is considered. The premise of the stochastic sediment supply forcing here is that the Olympic Mountains are considered to be in topographic steady-state with published estimates of denudation rate ranges (Section 1.1). We assume a denudation rate ($\boldsymbol{rate_d}$) of 0.2 mm/year as it falls within the published ranges. Hence, each stream tributary (i.e., CA_D) for all *j* accumulates sediment at the rate of 0.2 mm/year. Each stream tributary is also initialized with a time-lag until the next mass wasting event ($\boldsymbol{t_L}$). t_L is randomly sampled from an exponential distribution of

$$\boldsymbol{f\left(x;\frac{1}{\beta}\right) = \frac{1}{\beta}\, \exp(-\frac{x}{\beta})} \qquad (10)$$

where β is the mean time between mass wasting events and is arbitrarily assumed to be 20 years in this analysis. Hence, when t_L for a given *j* has passes and a mass wasting event occurs, Vd_{in} is

$$\boldsymbol{Vd_{in} = rate_d x\, CA_d x\, t_L} \qquad (11)$$

In the current version of the model we assume that Vd_{in} is always 50 percent sand and 50 percent gravel, however this is a simplifying assumption that can be improved in future versions of the model. After a mass wasting event occurs in a given *j*, a new t_L is sampled from the exponential distribution of (10) for the given *j*.

While this approach to sediment deposition assumes a constant $rate_d$, the stochastic mass wasting events emulate the processes of landslides, debris flows, and gullying that dictate geomorphology of the Elwha watershed and produce sediment supply that is dynamic in both time and space. This simple stochastic approach allows the ability to test the watershed response to different values of $rate_d$ and β, and further calibration of the values can improve model performance.

2.6 Case II: Network Model with Stochastic Streamflow Forcing

The streamflow applied in this case is observation-based and stochastically models the largest flows each year, as these drive the majority of sediment supply and transport. This case runs on sub-annual timesteps–with the exact timestep depending on the user-specified number of largest flows each year ($\boldsymbol{N_{large\ flows}}$)–and requires relatively simple inputs that can be varied to test system wide response. Since streamflow observations are only available at the Lake Mills gage, the Lake Mills gage rating curve is randomly sampled each year for 365 daily streamflow values (i.e., one year). The top $N_{large\ flows}$values are then extracted from the 365 values and we arbitrarily choose $N_{large\ flows}$=5 in this analysis. We assume uniform hydrology for this approach and therefore the five values of streamflow each year at the Lake Mills gage are scaled to all streams in the network by converting the Lake Mills gage flow values to units of mm/day (with respect to the contributing area of Lake Mills gage) and using the equation

$$\boldsymbol{Q_{j,t} = Q_{LM,t}\,(mm/day) x\, CA_{T,j}} \qquad (12)$$

where $\boldsymbol{Q_{LM,t}}$ is one of the five largest flows sampled from the Lake Mills streamflow rating curve for a given time-step $\boldsymbol{t}$, and $\boldsymbol{Q_{j,t}}$ is the flow at *j* for the same *t.*

We ran the Case II model for 134 years, as this is the time length that the Glines Canyon Dam was installed (84 years) plus a model "spin-up" period of 50 years. Hence, the number of timesteps of the model run was 5 large flows per year times 134 years, or 670 timesteps. After the model run was complete, we summed the Vo_g and Vo_s over the final 84 years of the model run at the stream link that the Lake Mills gage was located. We then converted this transported volume to a reservoir sedimentation volume using equation (7).

2.7 Case III: Network model with Streamflow from Hydrology Model

The streamflow forcing in Case III is the output from an open-source, sophisticated, physically-based hydrology model called DHSVM (Wigmosta et al., 1994). DHSVM combines spatially explicit physical characteristics of the watershed with meteorological data and has been extensively used in the Pacific Northwest (Cuo et al. , 2011; Frans, 2015; Wigmosta et al., 1994). A list of DHSVM inputs and their sources are included in Table 3. The meteorological forcing data input was manually adjusted for this study and further described in Section 2.7.1.

Table 3: DHSVM Input Sources

Model Input	Description	Source
Digital Elevation Model	Puget Sound Digital Elevation Model 30m	https://www.ocean.washington.edu/data/pugetsound/psdem2000.html
Soil Texture	NRCS STATSGO2	http://www.nrcs.usda.gov/
Vegetation	National Land Cover Database (NLCD 2011)	http://www.mrlc.gov/nlcd2011.php
Historical Meteorological Data	Mauger et al. (2016)	https://cig.uw.edu/news-and-events/datasets/hydrology-in-the-chehalis-basin/
	Livneh et al. (2013)	https://www.esrl.noaa.gov/psd/data/gridded/data.livneh.html
1981-2010 Precipitation Normals	800 m, PRISM Climate Group	http://www.prism.oregonstate.edu/

DHSVM was run at a 3-hour timestep from years 1915-2011 (97 years) following the procedures on the model website (https://dhsvm.pnnl.gov/). Streamflow was output at each of the 293 stream links of the study area river network at the 3-hour interval. The DHSVM model was minimally calibrated beyond the adjustment of the meteorological forcing data (Section 2.7.1). Minor adjustment of the soil texture parameters was conducted to improve the shape of modeled hydrographs relative to the observed hydrographs at the Lake Mills gage. However, further refinement of the DHSVM model would require time beyond the scope of this study. We considered the DHSVM model performance to be sufficient for the purposes of simulating reasonably realistic streamflow throughout the study area river network.

The 3-hourly DHSVM streamflow outputs were aggregated to daily values for the streamflow forcing component of the network model. The network model was run on a daily time step, which for the 97 years results in 35,429 time steps (~53 times more time steps than Case II). After the model run was complete, we summed Vo_g and Vo_s from 1927 to 2011 (i.e., the 84 years that the Glines Canyon Dam was installed) at the stream link that the Lake Mills gage was located. We then converted this transported volume to a reservoir sedimentation volume using equation (7).

2.7.1 Meteorological Forcing Data

Hydrology models such as DHSVM typically require inputs of meteorological forcing data, i.e., time-series inputs of temperature, precipitation, solar radiation, etc. The quality of meteorological forcing data is a critical factor in the performance of a hydrology model, as the meteorology is the main driver in the timing and amount of streamflow. In this study, we develop a meteorological forcing input as a hybrid of two published datasets: Livneh et al. (2013) and (Mauger et al., 2016). Neither of these individual datasets are suitable to use in the Elwha watershed in their published form, however our hybrid product improves upon their issues in the Elwha watershed and produces a more suitable dataset to use for this application.

The Livneh et al. (2013) meteorological dataset is a widely used statistically derived dataset, gridded at ~6 km resolution for the continental United States (CONUS). It provides daily precipitation, minimum and maximum temperature at 2 m, and wind speed from 1/1/1915 to 12/31/2011. Precipitation and temperature of the Livneh et al. (2013) dataset are derived from National Climatic Data Center (NCDC) Cooperative Observer (COOP) stations across CONUS, and wind data are from the National Centers for Environmental Prediction - National Centers for Atmospheric Research (NCEP-NCAR). Gridded daily precipitation are also scaled to match the 1961-1990 long-term monthly means of the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 1994). Livneh et al. (2013) is one of the few meteorological datasets that spans the entire time period that the Elwha dams were installed, and thus was ideal to use in this study. However, the amount of precipitation in the dataset is relatively high compared to observations of precipitation and streamflow in the Elwha watershed. The dataset also uses a constant temperature lapse rate of -6.5°C/km, while Minder et al. (2010) show that Olympic Mountain lapse rates are considerably higher (i.e., less negative) than -6.5°C/km and Currier et al. (2017) determined the mean lapse rate in the Olympic Mountains during precipitation events to be -4.5°C/km.

The Mauger et al. (2016) dataset is derived from the Weather and Research Forecasting Model (WRF; Skamarock et al., 2008). WRF simulates atmospheric dynamics and contains a cloud microphysical scheme which parameterizes processes that control precipitation. WRF does not require surface gauge observations, however it is sensitive to various model decisions (Currier et al., 2017). The Mauger et al. (2016) dataset is both statistically and physical-based, as it is based on a simulation driven by the NCEP-NCAR Reanalysis Project (NNRP; Kalnay et al., 1996). The Mauger et al. (2016) uses the same grid as the Livneh et al. (2013) dataset (~6 km resolution), allowing for simple comparison between the two datasets. However, the Mauger et al. (2016) dataset spans the year 1950- 2010, and thus excludes the years of 1927-1949 and 2011 which are needed to model the hydrology of the Elwha River during the period of Glines Canyon Dam installation. The Mauger et al. (2016) dataset performs better than the Livneh et al. (2013) dataset when the two datasets are compared to observations of precipitation and streamflow in the study area. We expect that the superior performance of Mauger et al. (2016) is due in part to its use of variable physically-based temperature lapse rates, rather than a constant lapse rate as used in the Livneh et al. (2013) product.

Since we desired to use a meteorological forcing dataset which had the performance of the Mauger et al. (2016) dataset and the temporal extent of the Livneh et al. (2013) dataset, we used

a simple approach to develop a hybrid product using the two datasets. First, we computed the long-term monthly means of temperature and precipitation at each gridded meteorological data point in the study area for both gridded datasets. Then, we computed the differences in monthly mean temperature and the ratios between monthly mean precipitation between the two datasets. We then used these differences (for temperature) and ratios (for precipitation) to adjust the monthly means of the Livneh et al. (2013) dataset. The equations for these computations are

Temperature adjustment:

$$\boldsymbol{T_{hybrid}} = (\boldsymbol{T_{Mauger_{ltmm}}} - \boldsymbol{T_{Livneh_{ltmm}}}) + \boldsymbol{T_{Livneh}} \quad (13a)$$

Precipitation adjustment:

$$\boldsymbol{P_{hybrid}} = (\frac{\boldsymbol{P_{Mauger_{ltmm}}}}{\boldsymbol{P_{Livneh_{ltmm}}}})\boldsymbol{x}\ \boldsymbol{P_{Livneh}} \quad (13b)$$

where ***ltmm*** is long term monthly mean. T or P with subscript $\boldsymbol{Mauger_{ltmm}}$ is the ltmm for the temperature or precipitation of a given month from the Mauger et al. (2016) dataset. Similarly, T or P with subscript $\boldsymbol{Livneh_{ltmm}}$ is the ltmm for the temperature or precipitation of a given month from the Livneh et al. (2013) dataset. T or P with subscript ***Livneh*** is the daily temperature or precipitation in a given month (corresponding to the month of the ltmm) from the Livneh et al. (2013) dataset. T or P with subscript ***hybrid*** is the daily temperature or precipitation in a given month (the same as the ltmm) that is adjusted by the ltmm difference or ratio and used as a daily meteorological forcing input for this study. The final output hybrid dataset was a daily forcing dataset spanning the period 1915-2011 that had ltmm means consistent with the Mauger et al. (2016) dataset.

The hybrid daily meteorological forcing dataset was disaggregated from a daily timestep to 3-hour timestep and incoming longwave and shortwave radiation and relative humidity were estimated using the Mountain Microclimate Simulation Model (MTCLIM) algorithms (Thornton & Running, 1999) as implemented in Bohn et al. (2013). During the DHSVM model run, precipitation was also further spatially distributed from the Livneh et al. (2013) grid cell centroid location to each 30m model grid cell based on the relative distributions of the 1981-2010 PRISM precipitation normals at each 30m model grid cell (Frans, 2015).

3.0 Results

Figures 4a and 4b as well as Table 4 compare the results of Case I, which include the Wilcock & Crowe (2003) formula application (equations 4a-b) with τ_r in its default form (i.e., equations 5a-d) as well as with the two variations of τ_r that we tested (i.e., one following the method of Gasperini et al. (2004) and one manually calibrated). Table 5 compares sediment transport and reservoir accumulation over the Glines Canyon Dam lifespan between Cases I, II, and III. Figures 6a through 6f show network model outputs of Cases II and III at three *j* that are indicated in Figure 6g. Figures 7a and 7b show our DHSVM modeled streamflow predictions compared to streamflow observations at Lake Mills gage.

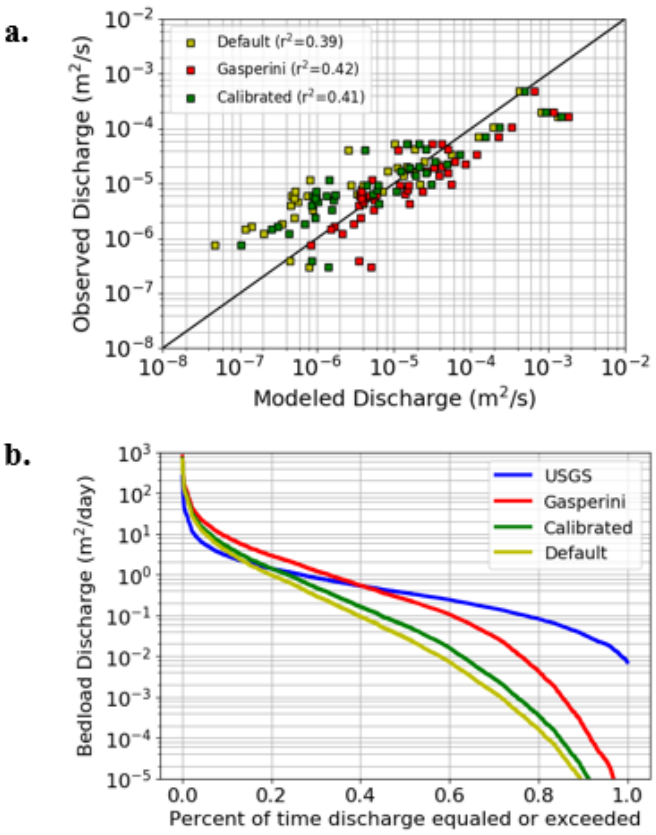


Figure 4. Comparison of Bedload Predictions
(a) Observed versus modeled bedload discharge using three variations of τ_r in Wilcock & Crowe (2003); (b) Bedload discharge rating curves: one from USGS (Curran et al., 2009) and three using different variations of τ_r in Wilcock & Crowe (2003)

Table 4. Comparison of Bedload Predictions

Estimation Method for Coarse Sediment (Bedload)	τ_r (*Pa*)		Dam Lifetime Coarse Sediment Volume (m^3)
	Sand	Gravel	
USGS Rating Curve	--	--	3.6 x 10^6
W-C, Default τ_r	5.06	7.39	6.8 x 10^6
W-C, Calibrated τ_r	4.56	6.65	8.0 x 10^6
W-C, Gasperini τ_r	3.24	5.76	11 x 10^6
Measured[1]	--	--	9.0 ± 1.9 x 10^6

Notes
W-C= Wilcock & Crowe (2003)
1. From Randle et al., 2014

3.1 Case I: Bedload Rating Curve (No Network Model)

Figure 4a shows the one-to-one comparison of observed and modeled bedload discharge for each approach to computing τ_r. Figure 4b shows the bedload rating curves for each approach to computing τ_r along with the rating curve developed by the USGS (Curran et al., 2009). Table 4 shows the values of τ_r for sand and gravel using each of the three approaches as well as the predicted bedload accumulation in the Lake Mills reservoir over the lifetime of the dams (84 years) using each rating curve in Figure 4b. Table 4 also shows the volume of coarse sediment that Randle et al. (2014) measured in the Lake Mills reservoir. Table 5 shows the Case I results of predicting sediment transport and deposition over 84 years using the Wilcock & Crowe (2003) formula application with τ_r in its default form. The default τ_r is used in order to properly compare the Case I results to Cases II and III which also use this transport capacity formulation. Table 5 shows that over the lifespan of the dams in Case I, the total predicted sediment transported is 6.83 x 10^6 m^3, the total predicted dam sedimentation volume is 10.9 x 10^6 m^3, and the overall percentages are 53% gravel and 47% sand.

Figure 4a shows bedload discharge observations versus the predictions using each of the three approaches for τ_r. Figure 4a shows that the Wilcock & Crowe (2003) with default τ_r tends to underpredict low discharge values and overpredict high discharge values for this application in the Elwha River. When τ_r is computed using the approach of Gasperini et al. (2004), the value of τ_r decreases for both sand and gravel (i.e., less τ is required to initiate

movement of sediment particles). This, in turn, leads to a systematic increase in the predicted discharge. Our manually calibrated values of τ_r fall in between the default and Gasperini et al. (2004) values, and as a result so do the bedload discharge predictions. These slight variations of the three approaches show relatively small changes in the discharge predictions and r^2 values. The largest changes of up to two orders of magnitude occur for the lowest discharge values, however these are actually relatively small changes in discharge.

Table 5. Comparison of outputs for three model constructs

Case	I	II	III	Measured[1]
Total Sediment Transported (m^3)	4.41 x 10^6	9.62 x 10^6	7.95 x 10^6	--
Total Volume of Reservoir Filled by Coarse Sediment (m^3)	6.83 x 10^6	14.9 x 10^6	12.3 x 10^6	9.0 ± 1.9 x 10^6
Percent Gravel	53%	51%	50%	--
Percent Sand	47%	49%	50%	--

Notes
All cases use Wilcock & Crowe (2003) parameterization for sediment transport capacity
1. From Randle et al., 2014

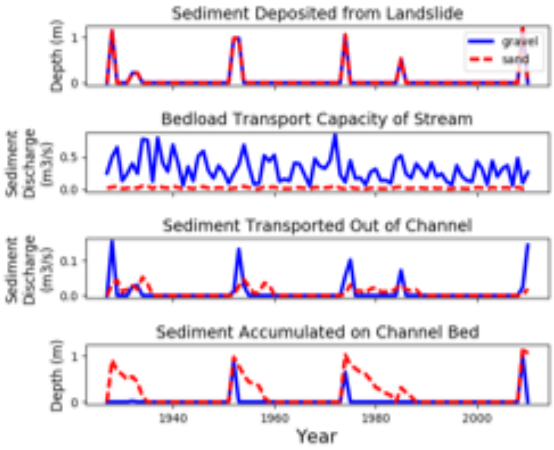
The magnitude of bedload discharge curves (Figure 4b) and total reservoir sedimentation volume (Table 4) predicted by each of the Wilcock & Crowe (2003) formula with varying τ_r corresponds to the trends discussed above, i.e., the approach using the Gasperini et al. (2004) τ_r produces the highest bedload discharge, followed by the calibrated τ_r approach and then the default approach. The approach of Gasperini et al. (2004) for τ_r produces a volume of bedload discharge which exceeds the coarse sediment measured in Lake Mills reservoir at the time of removal (~9.0 x 10^6 m^3). The approaches using the calibrated and default values for τ_r as well as the USGS curve all fall below the measured coarse sediment. We expect that the bedload reservoir sedimentation would fall below the measured coarse volume as these latter three predictions to, since some of the measured coarse sediment volume may also be sand that was carried by the suspended sediment load (which is not considered in this study). Figure 4b also shows that the rating curves for the Wilcock and Crowe (2003) equation and variations of τ_r have a different shape than the power rating curve developed by the USGS. Specifically, the Wilcock & Crowe (2003) derived equations predict higher bedload discharges for higher volume (lower probability) flows and predict lower discharges for lower volume (higher probability) flows. Hence, Table 4 and Figure 4b show that slight calibrations of τ_r cause up to ~30% difference in the predicted volume of dam sedimentation, and furthermore that the values of τ_r have significant control on the magnitude and shape of sediment discharge rating curves.

3.2 Case II: Network model with Stochastic Streamflow Forcing

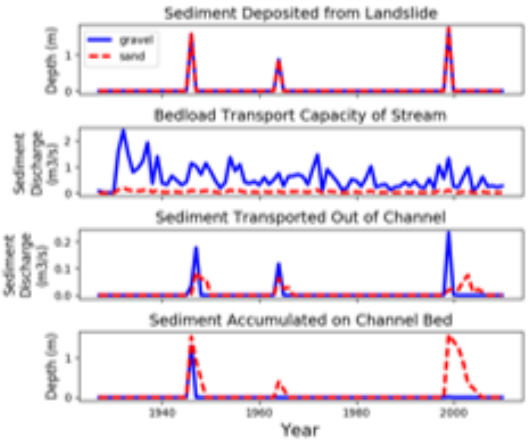
Table 5 shows that over the lifespan of the dams in Case II, the total predicted sediment transported is 9.62 x 10^6 m^3, the total predicted dam sedimentation volume is 14.9 x 10^6 m^3, and the overall percentages are 51% gravel and 49% sand. The total predicted dam sedimentation is larger than the measured volume from Randle et al. (2014) and thus there are significant improvements needed in the model parameterization and/or model design. However, the results provide insight to the dynamics of the physical system and help us to understand how it can be more realistically represented. The ability to view outputs at each individual *j* are particularly helpful in diagnosing the model.

Figure 6 (a-f). Comparison of Case II and Case III Model Outputs

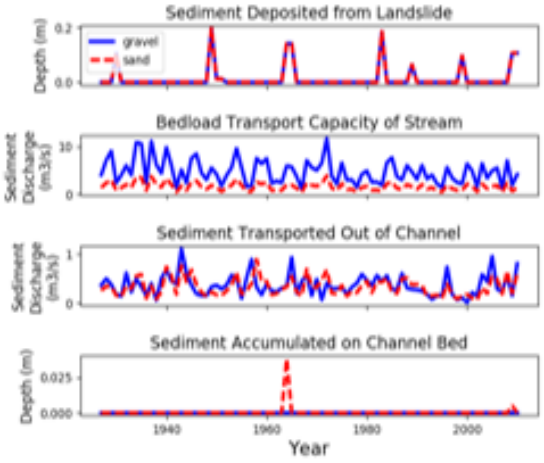
a. Case II: Headwater Reach



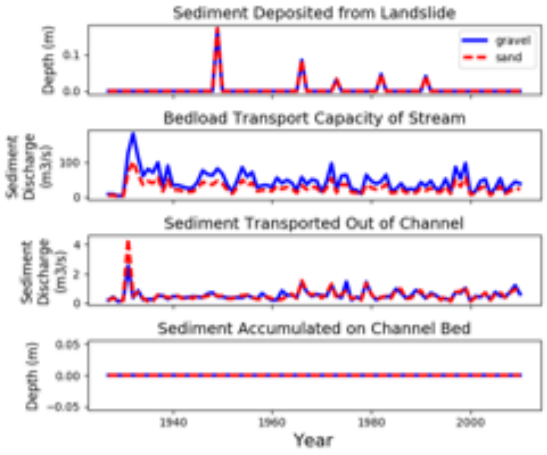
d. Case III: Headwater Reach



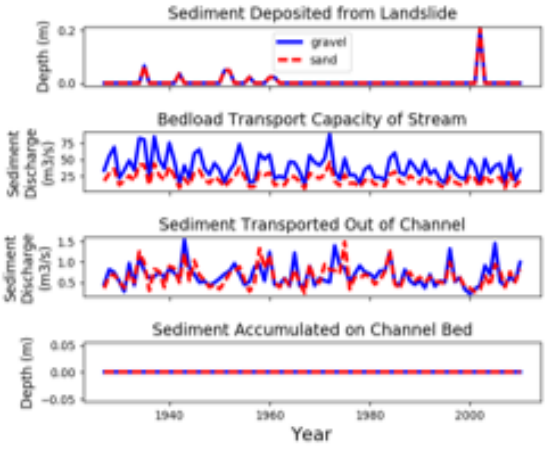
b. Case II: Lower Mainstem Reach



e. Case III: Lower Mainstem Reach



c. Case II: Lake Mills Gage Reach



f. Case III: Lake Mills Gage Reach

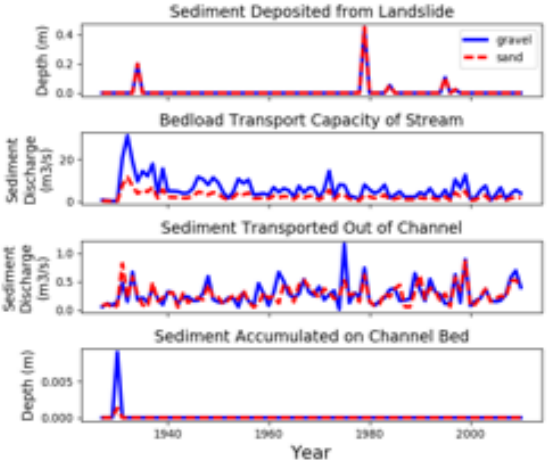
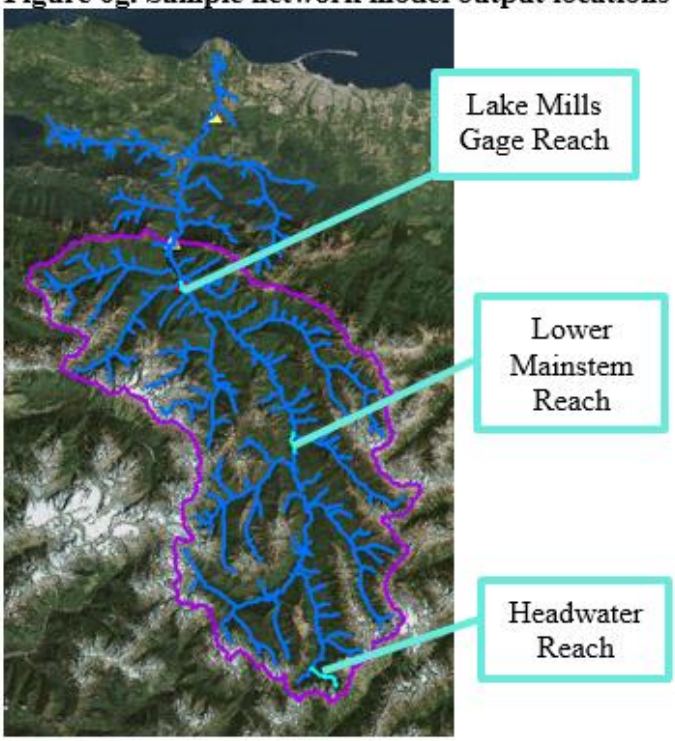


Figure 6g. Sample network model output locations



Figures 6a, 6b, and 6c show example outputs for Case II of the sediment mass balance at three locations in the study area channel network: a headwater reach (Figure 6a), a lower mainstem reach (Figure 6b), and at the Lake Mills gage (Figure 6c). The locations of these streams are indicated in Figure 6g. Each mass balance includes four plots: The depth of sediment deposited due to a stochastic mass wasting event each year (top panel); the bedload transport capacity for each of the three largest floods each year (second panel from the top); the bedload transport rate out of the channel for each of the three largest floods each year (second panel from the bottom); and the depth of sediment accumulated on the channel bed (bottom panel).

The side-by-side comparison of the mass balance plots demonstrates the different dynamics in the study area channel network. The sediment deposition (top panel) demonstrates the randomness of the stochastic wasting model and is independent of the location of the stream in the channel network. The bedload transport capacity of the stream (second panel from the top) shows that bedload transport capacity is lowest in the headwater reach and increases downstream. It can also be seen how variable the bedload transport capacity of the stream can be both intra- and inter-annually. The bedload transport rate of the channel (second panel from the bottom) demonstrates how the sediment transported out of the channel is related to the sediment deposited from mass wasting events, the incoming sediment from the upstream channels, and the bedload transport capacity of the stream. In the headwater reach, the sediment transported out of the channel can only come from the sediment deposited from mass wasting events. However, in the downstream reaches there is also sediment being transported out that was transported into the stream from upstream reaches. The sediment transported out of the Lake Mills gage reach is greater than the other reaches in large part because it has larger streamflow and stream power. The sediment accumulated on the channel bed is most dynamic in the headwater reach, where the bedload transport capacity is not able to remove all the sediment immediately after it was deposited since the stream power is relatively low. However, in the downstream reaches, it is rare for sediment to remain deposited on the channel bed expect when there is a large pulse that comes in from upstream or a mass wasting.

Generally, the model shows that gravel has higher sediment transport capacity and movement than sand, which we do not expect based on observations that show sand transport as being relatively higher. We suspect that this may be due to the arbitrary distribution of gravel and sand in the stochastic sediment supply forcing. Also, in reality, a layer of sediment would remain on

the channel bed of the majority of streams for most of the time, and there would be more dynamic evolution of the sediment accumulation on the channel bed. We suspect that the modeled τ is too high, which may be due to artificially high steepness of slopes that were computed during DEM processing. In addition, we expect sediment movement from other relatively large flow events beyond the top five events as used in this study which could change the behavior seen in the mass balance. Hence, there are various improvements that can be made to the stochastic sediment model such as calibrating $N_{large\ flows}$, $rate_d$, and β. However, our approach here provides a clear proof-of-concept model as well as helpful insights to the study area sediment dynamics.

3.3 Case III: Network model with Streamflow from Physically-based Hydrology Model

3.3.1 DHSVM Results

A comparison of observed versus modeled at the Lake Mills gage streamflow (during the periods Lake Mills streamflow was recorded) is shown in Figures 7a and b. The r^2 value for the modeled versus observed streamflow is 0.45. Figure 7a shows that the model captures the general behavior of the observed streamflow. One noticeable discrepancy is that the modeled streamflow is not able to reach the peak magnitudes in observed streamflow discharge. Capturing peak streamflows is a common challenge in watershed modeling, however it is particularly difficult in the Elwha watershed because of the complex hydroclimatology. The Olympic Mountains are subject to atmospheric river events that bring large amounts of rain, which typically is not captured by large-scale gridded meteorological products. In addition, it is difficult to properly model the complex rainfall, snowfall, and snowmelt patterns of the Elwha watershed that drive the peak events. For example, the model may produce too much precipitation as snow rather than rain in the winter, which leads to peak flows that are too low in

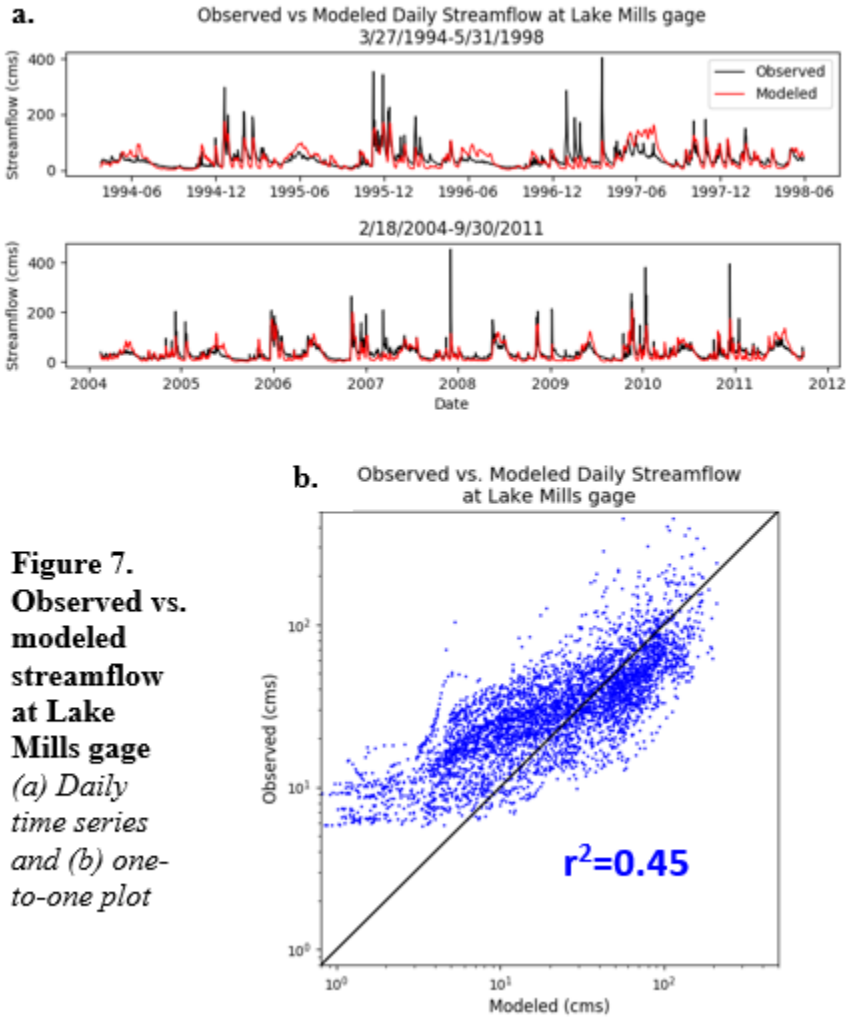


Figure 7. Observed vs. modeled streamflow at Lake Mills gage *(a) Daily time series and (b) one-to-one plot*

the winter (because there is not enough rain running off) and too high in the spring (when the snow melts). Figure 7b also shows that the modeled streamflow is mostly too low, particularly for the lowest and highest streamflow values. There is a clear bias for the streamflow values <20 cms which we expect may be related to the DHSVM model processes. The modeled streamflow is higher than observed streamflow for the middle-to-high values of streamflow, which we infer are mostly during spring snowmelt.

3.3.2 Network Model Results

Table 5 shows that over the lifespan of the dams in Case III, the total predicted sediment transported is 7.95 x 10^6 m^3, the total predicted dam sedimentation volume is 12.3 x 10^6 m^3, and the overall percentages are 50% gravel and 50% sand. As with Case II, the total predicted dam sedimentation is larger than the measured volume from Randle et al. (2014) and thus there are significant improvements needed in the model parameterization and/or model design. However, the lower volume of sediment predicted for Case III relative to Case II may be related to the lower magnitude of extreme events. The exact reason for this difference warrants further investigation.

Figures 6d, 6e, and 6f show example outputs for Case III of the sediment mass balance at the same three locations in the study area channel network that were shown for Case II (Figures 6a, 6b, and 6c) and are indicated in Figure 6g. Figures 6d, 6e, and 6f show similar dynamics as Figures 6a, 6b, and 6c that were discussed in Section 3.2. However, the sediment supply dynamics are inherently different in Case II and Case III since each uses a different stochastic sediment model run. The bedload transport capacity as well as sediment transported out of the channels are of comparable magnitude on an annual scale for Cases II and III, with the exception of a large streamflow and sediment pulse for Case III in the early 1930s and then relatively lower streamflow and sediment values in later years. This is a clear artifact of the physical basis of Case III which represents realistic interannual weather patterns (e.g., Pacific Decadal Oscillation), versus Case II which is purely stochastic. We further compared the spatial outputs of Cases II and III and could clearly see that spatial patterns of the hydroclimatology (e.g., rain shadow) impacted the spatial distribution of stored and transported sediment in the basin.

Besides the differences we point out between Cases II and III, we see similar issues in the two model constructs that may have similar root causes. These issues include the dominance of gravel transport capacity relative to sand as well as the shortage of sediment accumulation on the channel bed. We suspect similar root causes relating to the stochastic sediment forcing as well as artificially high τ related to overly steep slopes.

4.0 Discussion

A primary goal of this analysis is to balance simplicity and robustness in determining the level of model complexity that is necessary for predicting dam sedimentation in steep mountainous watersheds. The approaches presented in this report have various practical constraints including time, budget, modeling expertise, and data availability. Also, there can be complex uncertainties in results as well as difficulty in communicating the model and results to non-experts in watershed modeling and related fields. These constraints and complexities generally increase

from Case I to Case II to Case III. However, advanced models such as our network model and DHSVM are highly utilized in geophysical research as they allow for holistic understanding of hydrologic and geomorphic systems as well as the capacity to test alternate scenarios and conditions. They also provide insights to physical system processes that allow the model user to know whether the right results are being achieved for the right or wrong reasons.

Using a sediment rating curve for understanding sediment dynamics and predicting dam sedimentation, as in Case I of this study, can be sufficient for rough estimates of dam sedimentation if there are no major changes expected in environmental conditions and land use/land cover of a watershed. While in Case I we found that the application of Wilcock and Crowe (2003) using a calibrated value of τ_r matched the observed bedload observations and coarse sediment volume deposited in the Lake Mills reservoir most closely, it is important to point out that the different equations and parameterizations predicted dam lifetime sedimentation estimates within an order of magnitude of each other (~ 4 x 10^6 m^3 to ~ 1 x 10^7 m^3) and the measured volume of coarse sediment (~ 9 x 10^6 m^3). Hence, for preliminary estimates of dam sedimentation, the rating curve approach may be the most ideal because of the limited information and computing resources that are needed. If this approach is applied in practice, we recommend testing various reasonable equations and parameterizations such as we did in Case I to achieve a robust range of expected dam sedimentation.

However, our study demonstrates limitations of a rating curve approach for predicting dam sedimentation beyond ballpark estimates. The rating curve shape and magnitude as well as the predicted dam lifetime sedimentation are sensitive to geomorphic and hydrologic properties of a river reach, such as values for τ_r. A rating curve is based off historical data at a single stream reach and hence is extremely limited in its ability to capture how upstream dynamics may evolve the local geomorphic and hydrologic properties. Our Case II and Case III model results show that sediment does not accumulate along the Lake Mills gage reach. While this is from uncalibrated model runs and may not be realistic, it does begin to reflect system sediment dynamics and suggest that sediment may have relatively low residence time along the Lake Mills gage stream reach. If that is the case, the Elwha watershed and perhaps similar watersheds may have reservoir sedimentation rates that are largely influenced by sediment dynamics of upper reaches, which would warrant better understanding of the entire river network dynamics. In addition, the rating curve is particularly poor when considering hydrologic changes that are expected to occur in the future – such as more extreme events – that would alter streamflow and geomorphic regimes.

In consideration of these limitations of the rating curve approach, our network model (or a similar modeling framework) is an ideal alternative. Using even the simplest application of the model, the user can gain a broader understanding of the system such as (1) the relative amount of sediment supply and/or transport capacity in different streams and regions of the watershed; (2) the spatial and temporal variability of hydrology and geomorphology for the river network; and (3) the sensitivity of dam sedimentation rates to changes in environmental conditions both system-wide and in isolated areas of the watershed. This rich additional information that comes from the network model can be set up simply using a DEM, ArcGIS, and Python code that we have developed. The flexibility to vary sediment and streamflow forcing that we have integrated is ideal for users that have different levels of information, time, expertise, etc., or that would like to test different levels of model complexity in their watershed.

As we demonstrate in the application of Cases II and III, the network model can be applied using either a stochastic or deterministic modeling approach as well as with a range of model complexity. The ideal approach for watershed modeling is largely dependent on the characteristics of the watershed as well as the goals and capabilities of the model user. The Elwha watershed has various complexities such as steep mountainous topography, a strong precipitation gradient from the southwest, complex fluvial geomorphology, snow and glacier dynamics, and significant mass wasting and erosion processes. These complicate the hydrology and geomorphology of the river basin such that the assumption of uniform hydrology used in Case II is too simplistic and does not capture critical system dynamics. Hence, using a physically-based hydrology model, DHSVM, in Case III is a worthwhile enhancement to network model application. However, as we discussed, hydrology models such as DHSVM typically do a poor job at capturing extreme hydrologic events such as atmospheric rivers and rain-on-snow events. One reason for this is that gridded meteorological forcing data products are derived on regional scales that do not incorporate localized climate phenomena. The hydrologic models therefore are a better representation of "average" hydrology of a basin. Considering this, Case II is more ideal than Case III in its ability to simulate the impacts of extreme hydrologic events. While the hydrology is based on historic records, the approach case easily be modified if alternative scenarios are desired to be tested such as increasing the magnitude of peak annual flood(s).

However, because we are fortunate to have estimates of the observed coarse reservoir sedimentation from Randle et al. (2014), we are able to see that both Case II and Case III overpredict bedload reservoir sedimentation in this application. We suspect that in both cases, this relates to the parameterization of the stochastic sediment supply forcing as well as the inputs to τ (i.e., slope) which make its value artificially high. Out network model setup and output capabilities allow us to make this diagnosis and believe that issues such as these can be resolved with further model calibration and testing. We emphasize that our network model application is a proof-of-concept, and has made progress towards our overarching goals of (1) comparing various modeling approaches and (2) determining an adequately simplistic and robust model for dam sedimentation prediction.

We believe that further model development and testing is needed before absolute conclusions can be drawn about the most ideal modeling approach for us in engineering application. The necessary next steps that we would recommend include: (1) Using the Landlab modeling framework (Hobley et al., 2016) to incorporate a more realistic sediment forcing, such as one that integrates both basin-wide diffusive erosion processes and stochastic mass wasting events that are derived from Elwha watershed data; (2) Incorporating exchange and grain size evolution of sediment on the channel surface and in the bedload substrate; (3) Improve channel geometry (e.g., hydraulic geometry) and slope estimations; (4) Testing alternative bedload transport equations (e.g., Parker, 1990); (5) Integrating suspended sediment supply and transport; and (6) Improving the representation of peak flow events in DHSVM (or similar hydrology forcing model). We hypothesize that a "hybrid" approach for streamflow forcing of Cases II and III may also be an avenue to explore.

5.0 Conclusions

Our proof-of-concept modeling study has allowed us to understand and present the complexities of watershed hydrology and sediment model development as well as what factors are most important to consider when predicting dam sedimentation. Our testing of Cases I, II, and III provides insights to the level of model complexity that is warranted when predicting dam sedimentation. In Cases I, II, and III we also combine and compare approaches used in the academics fields of hillslope geomorphology, hydrology, and river network processes along with the engineering practices in the realms of hydrology and infrastructure (Figure 2). We believe that strides in integrating cross-disciplinary approaches used in both academia and practice are critical, such as we have begun to do, are critical to making a positive impact in water and infrastructure management. Furthermore, out study demonstrates how retrospective analysis of the rich history of hydropower can enhance knowledge on the life-cycle of hydropower dams moving forward.

6.0 Acknowledgements

The information, data, or work presented herein was funded in part by the Office of Energy Efficiency and Renewable Energy (EERE), U.S. Department of Energy, under Award Number DE-EE0006506 and the Hydro Research Foundation. We are very grateful for this funding as well as the opportunity to work with these entities.

The information, data or work presented herein was funded in part by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes and warranty, express or implied, or assumes and legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

7.0 Works Cited

Batt, G. E., Brandon, M. T., Farley, K. A., & Roden-Tice, M. (2001). Tectonic synthesis of the Olympic Mountains segment of the Cascadia wedge, using two-dimensional thermal and kinematic modeling of thermochronological ages. *Journal of Geophysical Research: Solid Earth*, *106*(B11), 26731–26746. https://doi.org/10.1029/2001JB000288

Bohn, T. J., Livneh, B., Oyler, J. W., Running, S. W., Nijssen, B., & Lettenmaier, D. P. (2013). Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agricultural and Forest Meteorology*, *176*, 38–49. https://doi.org/10.1016/j.agrformet.2013.03.003

Brandon, K. A., Roden-Tice, T. M., & Garver, J. I. (1998). Late Caneozoic exhumation of the cascadia accretionnary wedge in the Olympic montains, northwest Washington State. *Geological Society of America Bulletin*, *110*(8), 985–1009. https://doi.org/10.1130/0016-

7606(1998)110<0985:LCEOTC>2.3.CO;2

Buffington, J., & Woodsmith, R. (2003). Fluvial processes in Puget Sound rivers and the Pacific Northwest. *Book: Restoration of Puget Sound Rivers*.

Childers, D., Kresech, D. L., Gustafson, S. A., Randle, T. J., Melena, J. T., & Cluer, B. (2000). Hydrologic Data Collected During the 1994 Lake Mills Drawdown Experiment, Elwha River, Washington.

Cuo, L., Beyene, T. K., Voisin, N., Su, F., Lettenmaier, D. P., Alberti, M., & Richey, J. E. (2011). Effects of mid-twenty-first century climate and land cover change on the hydrology of the Puget Sound basin, Washington. *Hydrological Processes*, *25*(11), 1729–1753. https://doi.org/10.1002/hyp.7932

Curran, C. A., Konrad, C. P., Higgins, J. L., & Bryant, M. K. (2009). Estimates of Sediment Load Prior to Dam Removal in the Elwha River, Clallam County, Washington. *Scientific Investigations Report 2009-5221*.

Currier, W. R., Thorson, T., & Lundquist, J. D. (2017). Independent Evaluation of Frozen Precipitation from WRF and PRISM in the Olympic Mountains. *Journal of Hydrometeorology*, *18*(10), 2681–2703. https://doi.org/10.1175/JHM-D-17-0026.1

Czuba, J. A., & Foufoula-Georgiou, E. (2014). A network-based framework for identifying potential synchronizations and amplifications of sediment delivery in river basins. *Water Resources Research, 50,* 3826–3851. https://doi.org/10.1002/ 2013WR014227

Daly, C., Neilson, R. P., & Phillips, D. L. (1994). A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology, 33,* 140–158. https://doi.org/10.1175

Duda, J. J., Freilich, J. E., & Schreiner, E. G. (2008). Baseline Studies in the Elwha River Ecosystem Prior to Dam Removal: Introduction to the Special Issue. *Northwest Science*, *82*(sp1), 1–12. https://doi.org/10.3955/0029-344X-82.S.I.1

East, A. E., Pess, G. R., Bountry, J. A., Magirl, C. S., Ritchie, A. C., Logan, J. B., … Shafroth, P. B. (2015). Large-scale dam removal on the Elwha River, Washington, USA: River channel and floodplain geomorphic change. *Geomorphology*, *228*, 765–786. https://doi.org/10.1016/j.geomorph.2014.08.028

Finlayson D.P., Haugerud R.A., Greenberg, H. and Logsdon, M.G. (2000). Puget Sound Digital Elevation Model. University of Washington, (https://www.ocean.washington.edu/data/pugetsound/psdem2000.html).

Frans, C. (2015). Implications of Glacier Recession for Water Resources. *Doctoral Dissertation*. https://doi.org/10.1017/CBO9781107415324.004

Gaeuman, D., Andrews, E. D., Krause, A., & Smith, W. (2009). Predicting fractional bed load transport rates : Application of the Wilcock-Crowe equations to a regulated gravel bed river, *45*(June), 1–15. https://doi.org/10.1029/2008WR007320

Gasparini, N. M., Tucker, G. E., & Bras, R. L. (2004). Network-scale dynamics of grain-size sorting: Implications for downstream fining, stream-profile concavity, and drainage basin morphology. *Earth Surface Processes and Landforms*, *29*(4), 401–421. https://doi.org/10.1002/esp.1031

Gilbert, J. L. (1995). Alluvium distribution in Lake Mills, Glines Canyon Project and Lake Aldwell, Elwha Project, Washington. Elwha Technical Series PN-95-4. U.S. Bureau of Reclamation, Pacific Northwest region, Boise, ID, p. 60.

Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., & Tucker, G. E. (2016). Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. *Earth Surface Dynamics Discussions*, (September), 1–56. https://doi.org/10.5194/esurf-2016-45

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., … Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Knighton, D., 1998: Fluvial Form and Processes, 383 pp., Arnold, London.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., … Lettenmaier, D. P. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *Journal of Climate*, *26*(23), 9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1

Magirl, C., & Olsen, T. (2009). Navigability Potential of Washington Rivers and Streams Determined with Hydraulic Geometry and a Geographic Information System.

Magirl, C. S., Hilldale, R. C., Curran, C. A., Duda, J. J., Straub, T. D., Domanski, M., & Foreman, J. R. (2015). Large-scale dam removal on the Elwha River, Washington, USA: Fluvial sediment load. *Geomorphology*, *246*, 669–686. https://doi.org/10.1016/j.geomorph.2014.12.032

Mauger, G., Lee, S.-Y., Bandaragoda, C., Serra, Y., & Won, J. (2016). Effect of climate change on the hydrology of the Chehalis Basin, *6*(8), 53. https://doi.org/10.2307/3818105

Meyer-Peter, E., & Müller, R. (1948). Formulas for Bed-Load Transport. *Proceedings of the 2nd Meeting of the International Association of Hydraulic Research*, 39–64. https://doi.org/1948-06-07

Minder, J. R., Mote, P. W., & Lundquist, J. D. (2010). Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. *Journal of Geophysical Research Atmospheres*, *115*(14), 1–13. https://doi.org/10.1029/2009JD013493

Montgomery, D. R., & Brandon, M. T. (2002). Topographic controls on erosion rates in tectonically active mountain ranges. *Earth and Planetary Science Letters*, *201*(3–4), 481–489. https://doi.org/10.1016/S0012-821X(02)00725-2

Pacific Northwest National Lab, DHSVM-PNNL, (2017), GitHub repository, https://github.com/pnnl/DHSVM-PNNL/tree/master/CreateStreamNetwork_PythonV.

Parker, G., Klingeman, P. C. and McLean, D. G. (1982). Bed load and size distribution in paved gravel-bed streams, J. Hydraul. Eng., 108(4), 544–571.

Parker, G. (1990). Surface-based bedload transport relation for gravel rivers. *Journal of Hydraulic Research*, *28*(4), 417–436. https://doi.org/10.1080/00221689009499058

Randle, T. J., Bountry, J. A., Ritchie, A., & Wille, K. (2014). Large-scale dam removal on the

Elwha River, Washington, USA: Erosion of reservoir sediment. *Geomorphology*, *246*, 709–728. https://doi.org/10.1016/j.geomorph.2014.12.045

Shen, Z. Y., Gong, Y. W., Li, Y. H., Hong, Q., Xu, L., & Liu, R. M. (2009). A comparison of WEPP and SWAT for modeling soil erosion of the Zhangjiachong Watershed in the Three Gorges Reservoir Area. *Agricultural Water Management*, *96*(10), 1435–1442. https://doi.org/10.1016/j.agwat.2009.04.017

Skamarock, W. C., Klemp, J. B., Dudhi, J., Gill, D. O., Barker, D. M., Duda, M. G., … Powers, J. G. (2008). A Description of the Advanced Research WRF Version 3. *Technical Report*, (June), 113. https://doi.org/10.5065/D6DZ069T

Surian, N. (2002). Downstream variation in grain size along an Alpine river: Analysis of controls and processes. *Geomorphology*, *43*(1–2), 137–149. https://doi.org/10.1016/S0169-555X(01)00127-1

Thornton, P. E., & Running, S. W. (1999). An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agricultural and Forest Meteorology*, *93*, 211–228. https://doi.org/10.1016/S0168-1923(98)00126-9

United States Geological Survey (USGS). (2016, June 15). National Water Information System data available on the World Wide Web. Retrieved from USGS Water Data for the Nation: https://waterdata.usgs.gov/nwis/uv?site_no=12044900

United States Society on Dams. (2015). *Modeling Sediment Movement in Reservoirs*.

Warrick, J. A., Bountry, J. A., East, A. E., Magirl, C. S., Randle, T. J., Gelfenbaum, G., … Duda, J. J. (2015). Large-scale dam removal on the Elwha River, Washington, USA: Source-to-sink sediment budget and synthesis. *Geomorphology*, *246*, 729–750. https://doi.org/10.1016/j.geomorph.2015.01.010

Warrick, J., Draut, A., McHenry, M., & Miller, I. (2011). Geomorphology of the Elwha River and its delta. *Coastal Habitats of the Elwha River—Biological and Physical Patterns and Processes prior to Dam Removal*, 47–73.

Wigmosta, M. S., Vail, L. W., & Lettenmaier, D. P. (1994). A distributed hydrology-vegetation model for complex terrain. *Water Resources Research*. https://doi.org/10.1029/94WR00436

Wilcock, P. R. (1997). The components of fractional transport rate. *Water Resources Research*, *33*(1), 247–258. https://doi.org/10.1029/96WR02666

Wilcock, P. R. (2001). Toward a practical method for estimating sediment-transport rates in gravelbed rivers. *Earth Surface Processes and Landforms*, *26*, 1395–1408.

Wilcock, P. R., & Crowe, J. C. (2003). Surface-based Transport Model for Mixed-Size Sediment. *Journal of Hydraulic Engineering*, *129*(2), 120–128. https://doi.org/10.1061/?ASCE?0733-9429?2003?129:2?120? CE

Wilcock, P. R., & Kenworthy, S. T. (2002). A two-fraction model for the transport of sand/gravel mixtures. *Water Resources Research*, *38*(10), 12-1-12–12. https://doi.org/10.1029/2001WR000684

Appendix

TABLE A1: Historical Data from the Lake Mills Gage Used in Analysis

Parameter Description	Date Range and Frequency of Available Data	Number of samples	Sampling Equipment/Technique	Source	Notes
Streamflow Discharge and Channel Geometry Measurements					
Instantaneous streamflow discharge, channel width, mean channel depth, channel cross-sectional area, mean streamflow velocity	3/25/1994 – 4/27/1994; Intermittent	11	*Velocity/Discharge:* standard Price-AA velocity meter. Measurements made from boat or wading in the water. *Width, depth, area:* Not described.	Childers et al., 1999 (Table 2a)	Also available at: https://waterdata.usgs.gov/wa/nwis/measurements/?site_no=12044900&agency_cd=USGS&amp
Instantaneous streamflow discharge, channel width, mean channel depth, channel cross-sectional area, mean streamflow velocity	5/11/1994 – 10/13/2011; Intermittent	80	*Discharge:* Varies- midsection method, acoustic doppler current profiler or unspecified. Measurements made from boat, causeway or wading in the water. *Velocity:* Varies- standard Price-AA velocity meter, acoustic doupler current profiler or unspecified. *Width, depth, area:* Not described.	U.S. Geological Survey, 2016	Website also includes 7 measurements collected upstream or downstream of the gage as well as 11 measurements from Childers et al., 1999 streamflow data (these 18 samples are excluded from the listed number of samples, 80).
Daily mean streamflow discharge	3/26/1994 – 5/31/1998, 2/18/2004 – 9/30/2011; Daily	3911	*Streamflow Discharge:* USGS streamflow gaging station.	U.S. Geological Survey, 2017	Streamflow discharge collected in 15-minute intervals is also available from the USGS but is not directly used in this analysis.
Bedload Discharge					
Instantaneous bedload discharge, Instantaneous streamflow discharge, Particle size distribution (0.5 mm - 64 mm)	4/8/1994-7/26/1994; Intermittent	19 samples (74 sub-samples)	*Bedload Sampler:* "Elwha sampler" with 10.2 x 20.3 cm intake nozzle, 1.40 expansion ratio and catchment bag with 0.5 mm mesh. Connected to a standard or long wading rod. *Bedload Technique:* MEWI cross-section sampling (Edwards and Glysson, 1999) from the side of a boat. Most samples consisted of four subsamples (range of two to six), each subsample representing an individual transect sample. *Bedload Size Distribution:* Laboratory analysis. *Streamflow Discharge*: standard Price-AA velocity meter with a standard or long wading rod from the side of a boat.	Childers et al., 2000 (Table 2a)	On 6/21/1994, two samples collected within 1 hour of each other were reported with a large difference in bedload discharge. On 7/26/1994, no streamflow discharge was reported. 20 samples were reported however 1 sample (collected on 7/26/1994) was not used in this analysis because it did not include a streamflow discharge measurement.
Instantaneous bedload discharge, Instantaneous streamflow discharge, Particle size distribution (0.5 mm - 64 mm), channel width	11/1/1994-10/15/1997; Intermittent	23 samples (82 sub-samples)	*Bedload Sampler:* "Elwha sampler" with 10.2 x 20.3 cm intake nozzle, 1.40 expansion ratio and catchment bag with 0.5 mm mesh. Connected to a standard or long wading rod. *Bedload Technique:* EWI cross-section sampling (Edwards and Glysson, 1999) from side of a boat or modified hand-held version at low flows when channel could be waded. Most samples consisted of four subsamples (range of one to six), each subsample representing an individual transect sample. Individual transect samples typically consisted of 13 evenly spaced sampling points along cross section. *Bedload Size Distribution:* Laboratory analysis. *Streamflow Discharge:* standard Price-AA velocity meter. Measurements made from boat or wading in the water. *Width:* Not described.	Curran et al., 2009 (Appendix C) and personal correspondence with Chris Curran, USGS	25 samples were reported however 2 samples (collected on 11/4/1994 and 11/30/1995) were not used in this analysis because they did not include particle size distributions for the bedload discharge. Study also uses bedload measurements from Childers et al., 2000.
Channel bed/Gravel Bar/Subsurface Samples					
Wetted channel bed sediments	4/9/1994	1 sample (5 sub-samples)	*Streambed Sampler:* Steel pipe dredge (76 mm diameter, 304.8 mm length) with cap on one end and baler handle on the other, from side of a boat with a rope. *Streambed Technique:* EWI cross-section sampling (Edwards and Glysson, 1999) consisting of 5 evenly spaced sampling points along cross section. *Streambed Size Distribution:* Laboratory analysis. *Streamflow Discharge*: standard Price-AA velocity meter.	Childers et al., 2000 (Table 8)	
Gravel bar surface and subsurface	6/21/1994	2 surface samples, 1 bulk subsurface sample	*Surface:* Particle counts of 200 sampling points. *Subsurface:* Bulk sample collected.	Childers et al., 2000 (Table 9a)	Gravel bar surface particle counts conducted from the upstream and downstream ends of the left bank gravel bar near Lake Mills gage. Bulk sample of subsurface collected with each particle count.

Notes
EDI = equal- discharge increment
EWI= equal- width increment
MEWI= multiple equal- width increment
Additional parameters collected at the Lake Mills gage but not shown in this table included gage height, streamflow temperature, suspended sediment concentration, turbidity and streamflow water quality measurements.