Smartrock transport in a mountain stream: bedload hysteresis and changing thresholds of motion

Kealie L. G. Pretzlav[1,*], Joel P. L. Johnson [1,†] , D. Nathan Bradley[2]

[1] Dept. of Geological Sciences, The University of Texas at Austin, Austin, TX, USA

[2] U.S. Bureau of Reclamation, Denver, CO, USA

*Now at Balance Hydrologics, Berkeley, CA, USA

†Corresponding author, joelj@jsg.utexas.edu, ORCID 0000-0001-6286-9949

Abstract

Bedload movement is fundamentally probabilistic. Our quantitative understanding of gravel transport is particularly limited when flow conditions just exceed thresholds of motion, in part because of difficulties in measuring transport statistics during floods. We used accelerometer-embedded tracer clasts to precisely measure the timing of grain motions and rests during snowmelt floods in Halfmoon Creek, a gravel-bed mountain stream in Colorado, USA. These new data let us explore how probabilities of tracer movement vary with snowmelt discharge. Bedload hysteresis occurred over both daily and seasonal timescales, and included clockwise, counter-clockwise, and figure-eight patterns. We quantitatively explain these observations in terms of how thresholds of motion progressively evolved over 22 days during a seasonal snowmelt flood. Our results suggest that thresholds of motion are functions of both (a) cumulative shear stress and (b) temporal changes in shear stress during floods.

Plain Language Summary

Predicting the effects of floods on mountain river channels remains difficult but important because floods affect people, communities and ecosystems. Our research paper shows that the amount and timing of gravel that moves downstream depends not only on how much water is flowing in the channel at a given time, but also on how much flow and sediment movement has occurred previously during a flood or previous recent floods. We developed "smartrocks" that each hold sensors and batteries to measure the exact timing of these artificial tracer gravels. We collected field data during a month-long flood in a stream in the Rocky Mountains near Leadville, Colorado, USA. By measuring exactly when grains move during floods we can better understand how to predict when channels will be stable or change during future floods of different sizes, and how much change is likely occur.

1. Introduction

Interactions between society and the natural environmental have motivated our understanding of sediment transport since the early days of its study. Gilbert (1914, 1917) quantified movement in order to evaluate how hydraulic mining and agriculture affected channel aggradation and flooding. Supplies of water and sediment to river channels continue to be perturbed by human land use, and also by climate-related changes in flood frequency and magnitude. Changes in sediment transport dynamics can also critically affect natural ecosystem and habitat health. Predicting bedload transport and corresponding erosion and deposition during high-flow events of all different sizes is therefore important for the effective management of the interface between natural and engineered systems.

Bedload transport in gravel-bed rivers is controlled not only by spatial and temporal variations in flow, but also by thresholds of motion (e.g., Buffington and Montgomery, 1997; Bunte et al., 2013; Church et al., 1998; Yager et al., 2018). Even at a given discharge, gravel transport rates in individual rivers can span orders of magnitude (e.g., Lenzi et al., 2004; Rickenmann, 2001; Turowski et al., 2011). Many of the simplest yet arguably most widely-used bedload transport models have a typical form of $q_s \propto (\tau - \tau_{cr})^{3/2}$, empirically assuming that transport rate (q_s) is a power-law function of just two variables: shear stress (τ, a function of discharge) and a threshold for grain motion (τ_{cr}) (e.g., Meyer-Peter and Müller, 1948; Wong and Parker, 2006). Although threshold data usually exhibit a great deal of scatter, τ_{cr} values are generally treated as temporally constant (e.g., Buffington and Montgomery, 1997; Lamb et al., 2008). However, factors related to grain interactions such as clast clustering, sheltering and protrusion, overlapping and interlocking, packing density, surface roughness, force chain development, sand and gravel supply, and local erosion and deposition have been shown to influence thresholds of motion, and can also evolve through time in response to flow history (e.g., Hassan et al., 2020; Kirchner et al., 1990; Marquis and Roy, 2012; Masteller and Finnegan, 2017; Ockelford and Haynes, 2013; Recking, 2012; Sanguinito and

Johnson, 2012; Wilcock and Crowe, 2003; Yager et al., 2012). Using field data, Turowski et al. (2011) found systematic differences in the discharges at which bedload transport started and ended and started again from one flood to the next, suggesting systematic changes in thresholds. Building on these results, Masteller et al. (2019) demonstrated that τ_{cr} tended to progressively increase over seasonal timescales in response to small to intermediate flood discharges. In contrast, the largest floods caused thresholds to decrease. Johnson (2016) developed a equations to describe the temporal evolution of τ_{cr} as a function of sediment supply and local erosion and deposition, and compared them to laboratory experiments. Yager et al. (2018) combined field, laboratory and numerical model constraints to argue that friction and interlocking between grains is a key control on thresholds of motion.

Bedload hysteresis is a specific example of discharge-dependent transport variability that is almost always observed in both field and laboratory settings (e.g., Alexandrov et al., 2007; Mao et al., 2014; Meirovich et al., 1998; Moog and Whiting, 1998; Olinde and Johnson, 2015). Current bedload transport models have difficulty predicting hysteresis. Clockwise hysteresis (higher transport rates on rising limbs of hydrographs) is sometimes attributed to gradual decreases in sediment availability, or to progressive increases in bed surface stability through the evolution of structures such as coarse grains clustering or the degree of surface armoring (e.g., Mao, 2012; Roth et al., 2014). Counter-clockwise hysteresis (higher transport rates on falling limbs) can be caused by temporal lags as bedforms adjust to changing discharge (Bombar et al., 2011;

Martin and Jerolmack, 2013), or to the destabilization of surface structures during hydrograph rising limbs, increasing falling limb transport rates (Kuhnle, 1992). Many of these mechanisms proposed to explain hysteresis have also been shown to cause changes in thresholds of motion. Bedload equations such as $q_s \propto (\tau - \tau_{cr})^{3/2}$ could predict transport hysteresis if τ_{cr} evolved systematically through time.

The overall goal of the present research is to better understand how and why bedload transport probabilities and corresponding thresholds of motion change in response to river discharge during floods. Over what timescales do thresholds evolve? How well (or poorly) can we predict timeseries of threshold changes from timeseries of discharge? What physical mechanisms for threshold evolution are consistent with our unique smartrock-based constraints on transport probabilities? We found that thresholds changed between daily hydrographs, and even between rising and falling limbs of hydrographs, which could explain hysteresis. Thresholds evolved systematically with both cumulative shear stress and with the change in shear stress from one flood to the next. Finally, we interpret that changes in grain interlocking probably provide the most plausible physical mechanism for rapid changes in thresholds.

Instrumented tracer particles offer great potential for improving our statistical understanding of coarse particle transport in Earth surface processes, but their use also poses many challenges (e.g., Gimbert et al., 2019; Gronz et al., 2016; Maniatis et al., 2017). Because another goal of this work has been to improve the design of instrumented tracer “smartrocks”, we next describe the sensors, their new housings, and equipment limitations. Methods developed for data analysis include an algorithm to infer rest and hop durations from our time series of particle acceleration, validated using flume experiments. We then use our field data to calculate how transport probabilities changed through time.

2. Methods

2.1 *Tracer Design and Motion Sensor Technology*

Olinde and Johnson (2015) used concrete-encased accelerometer tracers to measure the timing of bedload motion during snowmelt floods in Reynolds Creek, Idaho, USA. They used Onset HOBO Pendant G data loggers which have ample battery life but limited data storage. Sensors were sampled once every 10 minutes, which allowed them to determine if a given particle had moved in the last 10 minutes. However, the duration and number of particle hops were unknown over shorter timescales.

The "smartrock" tracers we developed for this study sample nearly 4 orders of magnitude faster, letting us measure the precise timing and duration of motions and rests. We chose an off-the-shelf motion sensor from Gulf Coast Data Concepts which used an InvenSense 9150 9-axis inertial measurement unit (IMU) to measure acceleration, rotation rate, and compass direction with a 3-axis $\pm$16g accelerometer, 3-axis $\pm 2000\ °/s$ gyroscope, and 3-axis $\pm\ 1200\ \mu T$ magnetometer, respectively (Figure 1). The IMU can sample at up to 100 Hz, but because faster sampling consumes more power, a slower rate of 10 Hz was chosen to balance duration of data collection with data resolution. Each sensor was powered using three 3.6V, 2.6 Ah non-rechargeable lithium batteries connected in parallel, which together could power the device for as long as 40 days. Each battery had the same dimensions as a common 1.5V AA battery. Each sensor recorded data on a micro SD card. Battery life, rather than data storage, limited data collection. Each tracer clast also held a backup HOBO Pendant G logger, which sampled 3 orthogonal axes of acceleration once every 10 minutes. These allowed data to be logged for approximately five months, ensuring that we would have some constraint on motions that occurred after the sampling span of the other sensors.

Each motion sensor was enclosed in a custom manufactured case which we designed (Figure 1). The case dimensions were chosen to make the tracer as small as possible, but with enough room to hold the motion sensors, batteries, and a circular 30 mm RFID tag. We chose an ellipsoid-like shape for the case with major, intermediate, and minor axes diameters of 12.0, 7.2, and 6.4 cm, respectively. The case was injection-molded using a highly-durable thermoplastic mixed with a copper powder to increase the density to 3.3 g/cm^3. Accounting for void space, batteries and sensor components the bulk tracer density was 2.65 g/cm^3. Two identical halves were held together in four places with bolts and nuts resistant to loosening. An o-ring helped prevent water from entering the cavity (Figure 1).

The high-density plastic was originally chosen instead of metal with the hope of using RFID-technology for tracer recovery, as metal interferes with radio frequencies. Although preliminary testing suggested that tags would be readable through this plastic, after production we found that the dispersed copper powder, batteries, and sensors were unfortunately sufficient to block the RFID signal. Therefore, the passive RFID tag was used for identification purposes when cases were open, but tracers were found visually on the stream bed and by using a metal detector when buried.

2.2 *Study site: Halfmoon Creek, Colorado, USA*

Previous bedload transport studies conducted in Halfmoon Creek, a gravel-bed stream that drains Mount Elbert and Mount Massive in Colorado, USA (Figure 2a), include Torizzo and Pitlick (2004), Mueller and Pitlick (2005), Bradley and Tucker (2012), and Bradley (2017). The drainage area at the study site is approximately 61.5 km^2 and the elevation is approximately 3015 m. There are no significant tributaries between our study site and USGS gage 07083000 located 1.5 km downstream. The gage has operated continuously since August 1946. Discharge is dominated by spring snowmelt and produces an annual flood that typically lasts from mid-May to Mid-July. The spring 2015 flood peaked at 11.5 m^3/s on June 17th (Figure 2b). Based on a 69-year record, this discharge had a 10-year recurrence interval.

The study reach is alluvial, with bed-surface grain sizes ranging from fine gravel to meter-scale boulders. For this study, we measured median surface grain sizes (D_{50}) of 6.4 and 12.9 cm based on two Wolman-type point counts *(N=400)* in two short reaches (Figure 2a). D_{84} for both locations was about 29 cm. Our 7.2 cm intermediate axis tracers correspond to the 51st and 40th percentiles of surface grain sizes in the upstream and downstream locations, respectively. Bradley and Tucker (2012) reported D_{50}=5.5 cm measured over a somewhat longer reach which includes our study area. Mueller and Pitlick (2005) report surface D_{50} between 5.0 and 7.2 cm for six reaches within ~2 km upstream and downstream of the gaging station. Considering the reach-scale variability in grain sizes, we estimate that our tracer with intermediate axis of 7.2 cm is a reasonable approximation of the reach-averaged D_{50}.

The thalweg of the channel is approximately 1 m below the banks and the channel is approximately 10 m wide. The slope of the thalweg of the longer study reach from Bradley and Tucker (2012) is approximately 1%, while Mueller and Pitlick (2005) report slopes over similar reaches of 0.84-0.86%. There are several low-angle alternating bars with one large bar in the inside of a sharp bend approximately 200 m downstream of the deployment location (Figure 2a).

This field site was chosen for several reasons. First, the timing of snowmelt floods is predictable, usually peaking in late May or early June. Due to limited battery life, predictability in the timing of flow above transport thresholds was important. Second, Bradley and Tucker (2012) conducted a multi-year passive tracer campaign in this reach which provides important context for this study. Finally, the stream is wadeable at low flow, allowing tracer recovery necessary to retrieve the motion data.

2.3 *Field Methods*

We deployed 33 motion tracers on May 13th, 2015 in Halfmoon Creek, Colorado, in a similar location to the RFID-embedded tracers deployed by Bradley and Tucker

(2012) (Figure 2a). Tracers were positioned across the width of the portion of the channel that was subaqueous at the time of deployment. Following the methodology of Bradley and Tucker (2012), tracers were placed on the streambed inside the pocket made by gently removing a similarly sized grain, with the goal of minimizing enhanced mobility during the first few motions. Deployment occurred when channel discharge was approximately 0.8 m^3/s, well below the threshold of motion for the tracers. Pressure transducers (HOBO depth loggers) were installed in two locations in our study reach near the channel bank, and recorded the water depth at each location once every five minutes (Figure 2a).

Tracers were recovered in October 2015, when the stream discharge was approximately 0.3 m^3/s and easily wadeable. We were able to recover 27 of the 33 deployed tracers, an 82% recovery rate. Search efforts extended approximately 400 meters downstream beyond the farthest recovered tracer. Most of the recovered tracers were on the bed surface and were found by eye. Four recovered tracers were buried below the surface of the large aggrading bar, and were located with a metal detector. Because the remaining six unrecovered tracers were likely buried, our data may have a bias toward surface grains. Deployment positions were surveyed using a total station with sub-centimeter resolution. Recovery positions were measured with a Trimble XT GPS giving ± 1 m accuracy after post-processing. Of the 27 recovered, five tracer housings leaked due to subtle unrecognized warping of the housings during manufacture, and those IMUs did not record any motions before logging failed. One IMU logger malfunctioned despite remaining dry. The following analysis uses the remaining 21 tracers. Total logging times ranged from 24 to 40 days, with most lasting at least 30 days. The HOBOs recorded data once every 10 minutes until recovery in October.

2.4 Algorithm to Identify Motions and Rests

The accelerometer and gyroscope record the near-instantaneous acceleration and rate of rotation along three axes (x, y, and z) at 10 Hz. We use these data to detect the timing of particle entrainment and disentrainment. In practice, raw sensor data are noisy, and the motion sensors record all grain movements including wobbling of grains in place. We therefore developed a simple empirical algorithm using acceleration, rotation, and duration thresholds to identify motions which likely correspond to downstream translation of the particle. Controlled laboratory experiments were used to validate the algorithm and calibrate its parameters.

When a particle is at rest, the gyroscope records a rotation rate of zero for all three axes. When at rest, the accelerometer feels gravity and should record a vector sum of acceleration ($\sqrt{A_x^2 + A_y^2 + A_z^2}$) equal to 1 g (where g is gravitational acceleration, 9.81 m/s^2, and A_x, A_y, A_z are accelerations measured along each axis). In practice, noise on the ±16g accelerometers produces a vector sum of 1±0.1g at rest. Changes in acceleration on different axes as well as non-zero gyroscope readings should indicate particle motion. To remove acceleration noise during rests while preserving acceleration changes indicating motion, we applied a two-second moving window median filter to accelerations along all three axes. Cobble motion is generally initiated as a grain rotation out of a bed pocket.

Significant acceleration changes may not be detected on all three axes because the change in acceleration of a given axis can vary from 0 to 1 g depending on the particle orientation relative to the axis of rotation. Therefore, entrainment was detected when the value of the filtered accelerometer data of at least one axis changed by 0.1 g/s. We found that the gyroscope data were most effective at determining when a particle movement ended. A tracer particle was considered at rest when the gyroscope reading falls below an empirically derived threshold (0.3 rad/s) for any of the three axes. Motions and rests are only detected if they persist for two or more samples (0.2 s).

Flume experiments were used to evaluate and calibrate the algorithm, and suggest that it accurately identifies movements > 0.5 s in duration. We video-recorded a sample tracer in a 0.5 m wide laboratory flume with a mobile gravel bed, and compared manually detected motions from the video to the motion detection algorithm (Figure 3). To make sure the timing of entrainments and disentrainments was clearly observable in the video, we set the flume discharge to be large enough to maintain motion if the particle was already in motion, but not too large so that the particle would instantly begin moving once placed on the bed. Throughout the test, the particle was placed on the bed surface in the upstream portion of the video frame, and then pushed slightly by hand to initiate motion. Once the particle reached the edge of the video frame and stopped moving, we repositioned the tracer to the upstream portion of the video frame. The threshold values (0.1 g/s, 0.3 rad/s) were determined in order to allow all observed displacements to be identified correctly. Two instances of particle wobble (500 and 740 seconds) and three impacts by another larger cobble (540, 550, and 690 seconds) were correctly not identified as motions.

While the algorithm reasonably identifies particle entrainment and disentrainment, the flume test revealed two limitations. First, in two instances a single motion was incorrectly identified as two motions separated by a brief 0.2 second rest when the particle stopped rotating and momentarily slid across the pea gravel surface. The algorithm only detects particle rotations as a motion indicating displacement downstream, and not pure sliding with no rotation. However, the coarse and rough bed surface in Halfmoon Creek means that our ellipsoid-like tracer particles are not likely to be able to slide across the bed surface without rotation very often. Second, two brief motions were identified when the particle was not actually displaced, approximately 580 and 810 seconds into the test. In both cases, the particle was artificially jostled by a hand resulting in a permanent rotation but not displacing it downstream (arrows in Figure 3). This would most likely occur in the field when a particle partially rotates up from the bed but does not fully exit its pocket. The results suggest that identified motions less than about 0.5 s may be less reliably detected than longer-duration motions. We assume that these uncertainties in detecting movements are acceptable for our analyses.

2.5 *Hydraulic Forcing and Bedload Transport Probabilities*

To frame results in terms of hydraulic forcing, we calculate bed shear stress τ using the depth-slope product, $\tau = \rho g h S$, where ρ is water density (1000 kg/m^3) and h is water depth. For reach slope S we use the average water surface slope between the two pressure transducers, 0.5%. Unfortunately, temporal changes in water surface slope were not resolved with sufficient accuracy relative to noise in the pressure transducer data, and so for simplicity we assume that the water surface slope remained at 0.5% during both rising and falling limbs of the floods. We also confirmed that the reach bed slope was 0.5% using the surveyed recovery positions of the tracers found in the channel thalweg. The time-dependent record of water depth is derived from the two pressure transducer records (Figure 2a). An offset measured in the field was used to infer water depths from the stage records. The two depth records were averaged so the time series of shear stress best represented reach-averaged conditions. Finally, we calculate dimensionless shear stress (Shields stress) as

$$\tau^* = \frac{\tau}{(\rho_s - \rho)gD}\,, \tag{1}$$

where ρ_s is sediment density (2650 kg/m^3) and D is intermediate grain diameter (0.072 m).

From the time series of tracer motions and rests we calculate the probability of transport, P_q , as:

$$P_q = {}^{n_m}/_{n_s}\,, \tag{2}$$

where n_m is the number of measurements that indicate a particle is in motion, and n_s is the total number of measurements in that sampling interval, calculated for all of the tracers recording data over a given time interval. For temporal calculations of n_s we used

10 minute intervals. For example, 10 tracers recording data over 10 minutes at a sampling frequency of 10 Hz would correspond to 60,000 total records, so n_s=60,000. If, during the same interval and for the same tracers, we detected that 120 of these measurements (0.1 s each sample) indicated motion, then n_m=120 and P_q=120/60,000.

3. Results

Beginning June 3rd, 2015 (21 days after deployment), there were 32 consecutive days with tracer transport. Because only the first 22 days had a substantial number of tracers recording data, the following analysis of hysteresis and thresholds of motion is limited to the first 22 diurnal flood events (Figure 4). Discharge increased over the first 15 days and then decreased (Figure 4a). Superimposed diurnal floods are defined from the flow minimum of one day to the next. We began with a population of 21 functional tracers. Different sensors stopped working at different times (Figure 4b), and our calculations of P_q account for decreasing numbers of functional tracers. Olinde and Johnson (2015) calculated P_q in the same way but since their motion sensors sampled once every 10 minutes, values of n_s represent fewer samples collected in a period of time. As a result, our P_q values are much lower than those presented in Olinde and Johnson (2015).

We first explore how transport probabilities varied with Shields stress (Figure 5). Rather than binning in time, samples (n_s) and motions (n_m) were binned into τ^* increments of 0.0004. P_q was calculated for each bin using equation (2). From these data we visually determine an overall threshold of motion of τ^*_{cr} =0.0387, which corresponds to a probability of transport of about 10^{-4} (Figure 5). The corresponding threshold stage and discharge are 0.92 m and 3.5 m^3/s, respectively. Several short-duration motions less than half a second were identified during lower flows, but with exceedingly small transport probabilities.

A logistic function fits the relationship between P_q and τ^* well (Figure 5; $R^2 = 0.96$):

$$P_q = \frac{1}{1 + e^{-429.5(\tau^* - 0.0625)}} \text{ for } \tau^* \geq 0.0387. \tag{3}$$

A power law can also be fit with $R^2 = 0.95$:

$$P_q = 10^{24.2}\, \tau^{*20.4} \text{ for } \tau^* \geq 0.0387. \tag{4}$$

We use the logistic function for most of our analyses below because it asymptotes towards the physical limit of P_q =1 for higher τ^*. For example, equation (3) predicts P_q ≈0.97 for τ^* =0.071. In contrast, the power-law fit predicts mathematically possible but unphysical transport probabilities of P_q >1 for τ^* ≥0.065. Nonetheless, neither equation is expected to be accurate outside of the range of the fitted data ($0.0387 < \tau^* < 0.05$; Figure 5).

3.1 Bedload Hysteresis

The time-averaged analysis in Figure 5 effectively treats transport hysteresis as noise, which it is not. Figure 6 compares temporal relationships between transport probability P_q and hydraulic forcing characterized by τ^*. Over the 22 days of snowmelt flood used in our analysis, average discharge increased over 15 days and then decreased, with superimposed diurnal floods (Figure 4a). Figure 6a plots P_q and τ^* calculated every hour, but averaged over a 24-hour moving window to smooth away the diurnal fluctuations. We find overall clockwise hysteresis with significantly higher transport probabilities on the overall rising limb (events 1-15) compared with the falling limb. A decrease in discharge corresponding to events 10 – 12 (Figure 4a) produced the smaller clockwise loop superimposed in the rising limb (Figure 6a).

A similar procedure is applied to each of the 22 diurnal flood events. Figure 6b-f shows events 8, 9, 13, 14, and 15; all 22 events are plotted in the supplementary material. Probability of transport is calculated over 15 minute intervals with data smoothed over a 2-hour moving window to reduce variability. Over the 22 diurnal flood events, hysteresis patterns are highly variable. For convenience we categorize them into four groups: Clockwise hysteresis (events 7, 8, 12, 13, 15, 16, and 18), counter-clockwise hysteresis (events 1 and 14), figure-eight hysteresis with higher transport at different times on both rising and falling limbs (events 10, 17, and 19-22), and “low-transport” with both minimal hysteresis and low transport rates throughout (events 2-6 and 11). Some events could be classified in two ways. For example, event 9 has figure-eight hysteresis but also higher average transport probabilities on the falling limb indicating net counter-clockwise hysteresis (Figure 6c).

3.2 *Thresholds of Motion*

Next, we determine how thresholds of motion that are a function of time, notated as $\tau_{cr}^*(t)$, would have to change to explain the observed transport hysteresis. The flow-based τ^* timeseries (Figure 4a) and population-averaged τ_{cr}^* =0.0387 are used to calculate what the dimensionless transport rate (q^*) would be following the modified Meyer-Peter and Müller bedload formulation of Wong and Parker (2006):

$$q^* = 4.93(\tau^* - \tau_{cr}^*)^{1.6} . \tag{5}$$

We also calculate a transport rate based instead on transport probability (P_q). To do this, we first rearrange equation (3) and change notation by substituting $\tau_{P_q}^*$ for τ^*:

$$\tau_{P_q}^* = \frac{-1}{429.5} \, ln\left(\frac{1}{P_q} - 1\right) + 0.0625. \tag{6}$$

$\tau_{P_q}^*$ represents a time-dependent Shields stress, calculated from the time-dependent probability of transport. Next, we modify equation (5) in two ways, by first setting $\tau_{cr}^* = 0.0387$ and second substituting in $\tau_{P_q}^*$ for τ^* using equation (6), to develop an equation for $q_{P_q}^*$, a non-dimensional transport rate estimate based on P_q:

$$q^*_{P_q} = 4.93\left[\frac{-1}{429.5}\ ln\left(\frac{1}{P_q} - 1\right) + 0.0625\ -\ 0.0387\right]^{1.6}. \qquad (7)$$

We then assume that $q^* = q^*_{P_q}$, and that temporal discrepancies between shear stress-based q^* and motion tracer-based $q^*_{P_q}$ are caused by temporal changes in τ^*_{cr}. We equate equations (5) and (7) and solve for time-dependent $\tau^*_{cr}(t)$ to give:

$$\tau^*_{cr}(t) = \ \tau^* + \frac{1}{429.5} ln\left(\frac{1}{P_q} - 1\right) - 0.0238\,, \qquad (8)$$

where $\tau^*_{cr}(t)$, τ^*, and P_q all vary through time. Figure 7a shows the evolution of $\tau^*_{cr}(t)$ calculated from our data, both averaged over each diurnal flood event, and also averaged separately over each falling and rising hydrograph limb. Lower thresholds on rising limbs than falling limbs correspond to clockwise hysteresis, and vice versa (Figure 5, S1).

The diurnal-averaged $\tau^*_{cr}(t)$ gradually increases during the 22 days of flood. At the same time, $\tau^*_{cr}(t)$ tends to decrease when discharge increases from one day to the next. To quantitatively evaluate correlations between diurnal flood-averaged thresholds of motion and flow, we conducted an ordinary least squares multi-parameter linear regression analysis (MLR) using five hydraulic variables: (a) flood-averaged shields stress ($\overline{\tau^*}$), (b) peak shields stress ($\widehat{\tau^*}$), (c) difference between average rising and average falling shields stress within each diurnal flood (${\tau^*}_{rf}$), (d) cumulative shields stress (${\tau^*}_{+}$), and (e) the change in flood-averaged shields stress ($\Delta\tau^*$). The parameter ${\tau^*}_{+}$ is a cumulative sum of Shields stress over the analyzed 21 events. Positive values of $\Delta\tau^*$

indicate an increase in Shields stress from one diurnal flood event to the next. For example, $\Delta\tau^*$ for flood event 2 is equal to $\overline{\tau^*}$ for flood event 2 minus $\overline{\tau^*}$ for flood event 1. Flood event 1 is used to calculate change in Shields stress ($\Delta\tau^*$), but is excluded from the other regression analyses because its low critical Shields stress indicates that these motions were likely influence by the initial placement of the tracers on the channel bed.

Table 1 shows regression analysis results for each variable individually and considered together. Figure 7b-7d shows correlations between select variables. Single variable linear regressions indicate that $\Delta\tau^*$ is best correlated with $\tau^*_{cr}(t)$ (R^2=0.496), while ${\tau^*}_+$ is only slightly lower (R^2=0.453). Inclusion of all variables in the MLR resulted in $R^2 = 0.77$, but with coefficient-specific t-test p-values > 0.05 for all variables except $\Delta\tau^*$. Because these variables have some degree of inter-dependency, various groupings were calculated using MLR to infer the most relevant variables without over-fitting the data. Every MLR that did not include $\Delta\tau^*$ had a considerably lower R^2 (Table 1). No combination of three parameters produced a MLR in which all parameters had statistically significant t-test p-values for the slope coefficient. In a two-parameter MLR, use of $\Delta\tau^*$ with either $\overline{\tau^*}$, ${\tau^*}_+$, or $\widehat{\tau^*}$ resulted in statistically significant coefficients and with R^2 ranging from 0.726 to 0.748. The MLR using ${\tau^*}_+$ and $\Delta\tau^*$ predicts $\tau^*_{cr}(t)$ with $R^2 = 0.726$ (Figure 7d):

$$\tau^*_{cr}(t) = 0.002{\tau^*}_+ - 0.897\Delta\tau^* + 0.0389 \tag{9}$$

Equation (9) shows that, in our data, thresholds of motion tend to (a) increase with cumulative discharge, and (b) decrease when discharge increases from one diurnal flood to the next. Although both of the other two-parameter MLR produce similar results, we selected ${\tau^*}_+$ as the second parameter as it explains a higher proportion of the data variability alone than the other two variable choices (i.e. $R^2 = 0.453$).

4. Discussion

Because smartrock-based transport probabilities are a novel but untested method for quantifying bedload transport variables, we first demonstrate that our calculated transport capacities (τ^*/τ^*_{cr}) are reasonable for gravel-bed rivers. Similarly, the logistic and power-law fits to our temporally-averaged transport data (equations (3) and (4); Figure 5) provide new insights while being consistent with previous work. We then interpret that the evolution of grain interlocking is probably the most plausible mechanism to explain transport hysteresis and how motion thresholds changed quickly in our data, although coarse grain clustering may also adjust rapidly.

4.1 *Threshold channels*

Threshold compilations suggest that τ^*_{cr} can easily vary between perhaps 0.02 and 0.1 for gravel-bed rivers with slopes comparable to Half Moon Creek (e.g., Buffington and Montgomery, 1997; Mueller et al., 2005; Lamb et al., 2008). Mueller and Pitlick (2005) present a relation suggesting $\tau^*_{cr} \approx 0.039$ for a Halfmoon Creek reach close to ours (their equation 6 for their reach 3), comparable to the average τ^*_{cr} we found. While transport thresholds vary systematically in our data, the range of $\tau^*_{cr}(t)$ is small, from

≈0.038 to 0.043 when thresholds are averaged over each diurnal hydrograph (Figure 7a). However, threshold values only control transport in relation to shear stresses. The range of daily-averaged Shields stress during the 22 days of monitored flooding is ≈0.040-0.048 (e.g. Figure 6a). Thus, even though nondimensional threshold stresses only varied over a small range of values (≈0.05), they span over half of the range of Shields stresses (≈0.08) that occurred while flow was above threshold conditions during this 10-year flood.

In addition, the average transport capacity for this flood was $\tau^*/\tau^*_{cr} \approx 0.044/0.04 \approx 1.1$. Previous empirical and theoretical work suggests that bedload transport during bankfull floods usually occurs close to thresholds of motion for gravel bed rivers (e.g., Parker, 1978; Mueller et al., 2005; Phillips and Jerolmack, 2016, 2019), although sediment supply may also influence τ^*/τ^*_{cr} at bankfull (Pfeiffer and Finnegan, 2018). Quantitatively, our $\tau^*/\tau^*_{cr} \approx 1.1$ value is similar to the “closure” condition of $\tau^*/\tau^*_{cr} \approx 1.2$ for the middle of channels proposed by Parker (1978) for bankfull flow in gravel-bed rivers. Dunne and Jerolmack (2018) suggest that alluvial river banks adjust (through widening or narrowing) to have $\tau^*/\tau^*_{cr} \approx 1$. Thus $\tau^*/\tau^*_{cr} \approx 1.1$ is expected for threshold channels, suggesting that our probability-based smartrock threshold calculations are reasonable. Statistics calculated from the relatively small number of smartrocks that successfully recorded data (Figure 4b) appear to be sufficient for calculating bulk transport characteristics for gravels, probably because each tracer recorded large numbers of individual movements and rests.

4.2 *Power law logistics*

Interestingly, equation (4)—which shows that transport probabilities scale highly nonlinearly as ${\tau^*}^{20.4}$ for Shields stresses between 0.0387 and 0.05—is arguably consistent with previous work showing that transport is strongly nonlinear at low Shields stresses. Parker (1990) suggested that gravel movement should scale with $\tau^{15.7}$ for "very low sediment transport rates". Paintal (1971) empirically found that transport rates scaled as $q_s^* \propto \tau^{16}$ for $0.01 < \tau^* < 0.050$, and then transitioned to much less nonlinear transport with $q_s^* \propto \tau^{2.5}$ for $\tau^* > 0.05$. While our best-fit exponent is 20.4, Figure 5 shows that imposing a τ^* exponent of 16 and regressing to find the scaling factor alone also gives a strong fit of $P_q = 10^{18.4}\,{\tau^*}^{16}$, with R^2=0.93, p<0.0001.

Mathematically equation (4) predicts that transport probability $P_q = 1$ at $\tau^*/\tau^*_{cr} \approx \frac{0.065}{0.0387} \approx 1.8$. This transport capacity of 1.8 is high but not unphysical for gravel-bed rivers during floods. Flume experiments clearly demonstrate that gravel transport rates do not saturate near these flow conditions (e.g., Wilcock and Crowe, 2003). Therefore, P_q must be much less than 1 at $\tau^*/\tau^*_{cr} \approx 1.8$. The extremely nonlinear ${\tau^*}^{20.4}$ or ${\tau^*}^{16}$ scaling exponents must decrease systematically at Shields stresses higher than the range of our data (Figure 5), and would be different when measured over different data ranges.

In addition to fitting our data slightly better than the power law, the key benefit of our proposed logistic function (equation 3) is that it asymptotes to the physical transport limit of $P_q = 1$. However, our particular empirical fit asymptotes at a transport capacity of ≈ 2, which is undoubtedly too low. Equation (3) describes a symmetric sigmoid. An asymmetric logistic function, with one or more additional fitting parameters, would likely be required to also fit higher τ^*/τ^*_{cr} data. Powell et al. (2001) present bedload data collected up to $\tau/\tau_{cr} \approx 8.5$, and found that a gradual transition from size-selective to size-independent gravel transport occurred at $\tau/\tau_{cr} \approx 4.5$. Wilcock and Crowe (2003) present experimental bedload data up to $\tau/\tau_{cr} \approx 10$, which is broadly where grains gradually transition from energetic saltation to suspension (e.g. Sklar and Dietrich, 2004).

4.3 *Threshold evolution and correlations with flow*

The pervasiveness of near-threshold conditions ($\tau^*/\tau^*_{cr} \approx 1.1$) in gravel-bed channels highlights the importance of understanding subtle changes in $\tau^*_{cr}(t)$ in order to accurately predict bedload transport rates and corresponding channel changes during floods. Our data suggest that temporal changes in thresholds of motion ($\tau^*_{cr}(t)$) can be driven by the history of hydraulic forcing, in particular changes in diurnal Shields stress ($\Delta\tau^*$) and cumulative above-threshold Shields stress (τ^*_+) (Figure 7; Table 1). Together, a multiple linear regression with both variables gives a higher $R^2 = 0.726$ (equation 9; Figure 7d) than either variable alone. Furthermore, $\Delta\tau^*$ and τ^*_+ are not significantly correlated in our data ($R^2 = 0.096, p = 0.073$). We interpret that cumulative Shields stress and the change in Shields stress both are sufficiently independent to influence $\tau^*_{cr}(t)$ evolution in different ways.

When shear stress increases from one event to the next ($\Delta\tau^* > 0$), the threshold of motion tends to decrease (Figure 7b). This is most noticeable for events 4, 7, 13, and 14, which have relatively large increases in stress compared to the other events (Figure 7a). Much smaller increases in the diurnal flood event peak stress do not seem to produce the same decrease in $\tau^*_{cr}(t)$ (event 6, 9, 18). This is consistent with equation (9), because cumulative shear stress also increases with each event. Small $\tau_{cr}^*(t)$ decreases predicted by the $\Delta\tau^*$ term in equation (9) may be offset by small $\tau_{cr}^*(t)$ increases predicted from the cumulative τ_+^* term.

4.4 *Possible mechanisms for threshold evolution over short timescales*

Our analysis suggests that gravel thresholds of motion can evolve surprisingly quickly, e.g. within rising and falling limbs of individual daily floods (Figure 7a). Our field data provide a unique window into cobble transport statistics, but we do not have simultaneous observations of the bed surface or spatial interactions with other grains that could prove which flow-dependent mechanisms caused thresholds to evolve.

In the absence of independent constraints, we highlight different flow and transport conditions and corresponding mechanisms that previous work suggests might cause rapid threshold evolution. We then hypothesize which mechanisms may best explain our data. First, cumulative discharge both somewhat below and somewhat above "threshold" flow conditions tends to increase bed stabilization through a variety of recognized mechanisms including bed compaction, changes in bed surface roughness, and decreasing protrusion (e.g., Marquis and Roy, 2012; Masteller and Finnegan, 2017; Ockleford and Haynes, 2013; Paphitis and Collins, 2005). Second, changes in sediment supply from upstream can influence local bed mobility and thresholds of motion through grain size size distribution changes, grain impacts, and other possible mechanisms (e.g., Johnson, 2016; Pfeiffer and Finnegan, 2018; Recking, 2012). For example, grains smaller than the bed surface average (including sand sizes) can preferentially fill topographic lows and smooth the bed, in turn influencing near-bed shear stresses (e.g., Wilcock and Crowe, 2003; Venditti et al., 2010).

Third, sorting during transport can spatially organize surface grains into coarse grain clusters and other stabilizing structures, which in turn influence drag and bedload transport. Most although not all studies have found that increased clustering tends to enhance the overall stability of the bed surface, decreasing transport rates (e.g., Church et al., 1998; Hassan and Church, 2000; Hassan and Reid, 1990; Johnson, 2017; Piedra et al., 2012; Strom et al., 2004). Using flume experiments, Hassan et al. (2020) showed that clusters can dynamically expand, contract, and change through particle exchange between the cluster and transported grains, and interpret that clusters may buffer the bed by rapidly changing in response to short-term supply perturbations.

Fourth, forces between surrounding grains—due to interlocking, intergranular friction, and overlapping—can evolve over short timescales and may be dominant and underlying controls on thresholds of motion. In this granular physics view, force chains and particle contacts dictate mobility, though are difficult to view directly. Intergranular friction is a distinct mechanism from coarse grain clustering in that it that does not require grains becoming spatially reorganized by moving past other grains. Yager et al. (2018) combined field measurements of dislodgment forces with discrete element modeling of interacting spheres to support their model in which interparticle friction and grain protrusion relative to the surrounding bed are key variables controlling grain threshold distributions. In their model and data, grain resistance to motion can be 3-10 times larger than grain weight for particles with low protrusion. Earlier work has explored related factors that enhance thresholds of motion such as friction angles due to pocket geometry and resisting forces from grain overlap (e.g., Kirchner et al., 1990; Sanguinito and Johnson, 2012).

Overall, we hypothesize that the mechanisms described by Yager et al. (2018)—intergranular friction, protrusion, and overlap—are the primary drivers of $\tau_{cr}^{*}(t)$ evolution in our dataset. These factors can evolve quickly, explaining the observed transport hysteresis in our data, and describing $\tau_{cr}^{*}(t)$ changes over timescales significantly shorter than individual floods. The first hydrograph had an anomalously low $\tau_{cr}^{*}(t)$, which we associate with initial tracer positions being less stable (Figure 7) because the grains were not interlocked with surrounding grains. However, starting as soon as event 2 it appears that grains had attained more stable positions. We observe that when hydrograph-averaged shear stress was stable or slightly decreasing from one hydrograph to the next, $\tau_{cr}^{*}(t)$ tended to gradually increase. We interpret that this was caused by grains being gradually jostled in place (building force chains and increasing intergranular friction while also compacting the bed over time) and/or transported to adjacent positions that were more stable (Masteller and Finnegan, 2017). Statistically, some grains will be transported to less stable adjacent positions as well, but it is less probable that those grains remain there, as continued flow and turbulence will progressively move grains until they find more stable and interlocked positions. Thus, cumulative flow slightly above threshold conditions tends to increase thresholds of motion.

In general, given constant or perhaps very gradual increases in discharge, we interpret that intergranular friction remains constant and/or increases to balance the applied shear stress. When discharge drops and τ^* subsequently decreases, the higher intergranular friction remains, resulting in modest macroscopic increases in $\tau_{cr}^*(t)$ that depend on the previous level of τ^* (e.g. events 10, 11, 16). A grain will tend to be stable at shear stresses less than or equal to the stress that initially transported the grain to a given position. However, we interpret that moderate increases in shear stress from one hydrograph to the next (i.e., the $\Delta\tau^*$ term in equation 9) can break up force chains and overwhelm the intergranular friction developed at lower τ^*, "releasing" grains. This transport may further disrupt the bed through particle impacts and/or changes in local bed geometry. The net result is an increase in transport rates and decrease in thresholds of motion.

In addition to changes in intergranular friction among adjacent grains, it is also possible that our data reflect coarse grain clusters or other surface structures that developed through transport and were broken up over short timescales, enhancing stability as they expanded by adding grains, and enhancing transport by releasing sediment when they shrank or disintegrated (Hassan et al., 2020). Strom et al. (2004) found, using experiments with spherical grains over an immobile bed, that clustering generally increased over a range of τ^*/τ_{cr}^* ≈1.25 to 2, acting as a net sink for moving grains. Clusters broke up and then decreased due to increasingly energetic transport for $\tau^*/\tau_{cr}^* > 2.25$, acting as a net source. However, our stresses are much lower than these values. This may suggest that intergranular friction/grain interlocking are more important than cluster changes in our data, assuming that these transport capacity ranges are appropriate for natural grains and mobile beds.

We cannot discount the possibility that the overall trend of increasing $\tau^*_{cr}(t)$ with time and cumulative shear stress primarily reflects tracers being progressively worked into the bed or scour and fill effects (e.g., Haschenburger, 2011), and that the gradual threshold increase is an artifact of unstable initial conditions. However, nearly all of our particles were found on the bed surface, which suggests that progressive burial or scour and fill effects are unlikely to explain the gradual threshold increase found in our dataset. In addition, previous work shows that thresholds increase with cumulative flow without relying on tracer data (e.g., Paphitis and Collins, 2005). Using 19 years of monitoring data from the Erlenbach Torrent, Switzerland, Masteller et al. (2019) found that gravel thresholds tended to progressively increase due to the cumulative effects of small floods and below-threshold flows between floods, but that larger floods caused thresholds to decrease. Earlier monitoring work showed similar trends over multiple years of floods (Lenzi et al., 2004). Masteller et al. (2019) interpret that intense sediment transport in sufficiently large events disrupts bed surface grains enough to reset the "memory" of past flow conditions that led to shear stress increases.

While these researchers found increasing and decreasing transport thresholds from before and after relatively large floods over multi-year timescales, our results expand the parameter space of our understanding by documenting systematic threshold evolution over shorter timescales (changes within individual diurnal floods) and smaller changes in discharge. We interpret that threshold changes need not reflect complete surface destabilization or a significant reduction in the availability of mobile sediment, but can also reflect subtle changes in grain interlocking.

We also cannot discount the possibility that the threshold evolution we observed was caused by changes in surface grain size. Perhaps thresholds increased because the bed progressively coarsened overall over the 22 day period, while sand or finer gravel pulses also moved through the reach during times of highest discharge, temporarily decreasing thresholds, but were then transported out of the reach. While we did not did repeat surface GSD measurements directly before or during the data collection period, the point counts done after smartrock recovery (fall 2015) are consistent with previous measurements in Halfmoon creek (Bradley and Tucker, 2012). We thus feel like it is unlikely that overall surface coarsening with punctuating fining within our 22-day study period are responsible for observed trends. Masteller et al. (2019) similarly see no evidence that their threshold trends were controlled by systematic seasonal coarsening.

Hysteresis in bedload transport can likely be caused by a variety of mechanisms which are not mutually exclusive. Mao et al. (2014) observed day-to-day changes between clockwise and counterclockwise hysteresis, and suggested that a combination of migrating sediment waves and seasonal sediment supply changes from melting banks might explain daily to seasonal hysteresis trends. Roth et al. (2014) observed bedload hysteresis using near-stream seismic signals in a channel with high gravel sediment supply, and explore plausible mechanisms for gravel-bed systems including time lags between discharge and bedform adjustment or surface roughness changes (Mao, 2012; Martin and Jerolmack, 2013), bedload wave migration, and surface sorting/grain size changes (Humphries et al., 2012). Our data are similarly insufficient to confirm any particular mechanism responsible for the hysteresis we observe. Nonetheless, we suggest that grain interlocking and intergranular friction (Yager et al., 2018) may cause surface thresholds to evolve over timescales of individual diurnal floods, explaining hysteresis. Existing bedload transport models could predict transport hysteresis if threshold parameters evolve over time as functions of discharge as well as sediment supply (Johnson, 2016).

4.5 *Implications and Applications*

Our results may help improve predictions of bedload transport and bed stability, particularly in managed gravel-bed rivers. Channel reaches downstream of large dams tend to develop static and tightly interlocked armor layers, both as a function of reduced sediment supply, and decreased transport capacity as a result of reservoir-attenuated flood peaks (e.g., Viparelli et al., 2011). An increasing number of dam managers are considering downstream impacts to habitat (e.g. salmonid spawning) and including a 'naturalized' flow hydrograph with flow sufficient to mobilize the bed and reduce embeddedness of gravels. Understanding evolving thresholds of motion, and in particular how larger controlled floods might cause thresholds to decrease, could improve estimates and uncertainties of bed mobilization, as well as guide monitoring plans that could be implemented during managed floods and used for real-time decision making and hydrograph adjustment. For example, knowing when beds first destabilize during floods could be used to minimize water volumes released while still attaining bed mobilization goals. Conversely, overly mobile transport could potentially destabilize salmonid redds or impact other aquatic habitat. Future work could also explore how rates of hydrograph rise and fall influence bed mobilization.

Erosion and deposition can lead to channel avulsions and bank failures which can negatively impact life and infrastructure. In gravel-bed rivers, thresholds of motion are critical for predicting stability and transport across the range of flood magnitudes, including destructive floods that cause the most rapid channel evolution. Understanding bedload dynamics in flood events will become critically important in predicting river behavior as climate change and human land use change continue to impact the natural environment. Increases in the frequency and magnitude of floods are expected in many locations for decades or centuries to come (e.g. Milly et al., 2002), combined with climatic changes to hillslope hydrology and vegetation that will influence sediment supply to river networks. Our data suggest that how mountain channels respond to these environmental perturbations will be influenced by history-dependent thresholds of motion.

5. Conclusion

In 2015 we measured accelerations of Smart Rock tracer particles in Halfmoon Creek, Colorado during a seasonal snowmelt flood with a 10-year recurrence interval. Transport data was collected during 22 daily hydrographs which had flow above threshold transport conditions. We used tracer particle accelerations to infer the precise timing of motion and rest using an empirical algorithm, which was tested and calibrated in a controlled laboratory setting.

Our results suggest that the critical thresholds of motion for populations of particles evolved systematically over time with changes in discharge. In particular, increases in average shear stress from one day to the next correlate with decreases in thresholds of motion. Conversely, thresholds of motion increase as cumulative shear stress increases over the duration of the entire flood. Together, these two factors can explain ≈73% of the variability we observe in transport thresholds (Figure 7b; equation 9). Mechanistically, a variety of processes that influence bed stability could potentially explain our results, including changes in surface armoring, grain size changes and sediment supply pulses, clustering, and interlocking of grains. Given that we observe rapid changes in thresholds of motion between rising and falling limbs of daily hydrographs, we interpret that changes in intergranular friction and evolving force chains between grains are most likely the explanation, as these mechanisms could evolve rapidly and sensitively in response to local shear stresses.

Evolving threshold of motion are also illustrated by hysteresis in transport rates, which occurs in clockwise, counter-clockwise, and figure eight patterns. Progressive stabilization of grains and increasing thresholds of motion are supported by overall clockwise hysteresis over the 22-day above-threshold measurement period, and increasing entrainment thresholds after successive diurnal flood events with similar Shields stress. Counter-clockwise hysteresis after increases in Shields stress from one daily flood to the next suggests that thresholds decrease, potentially due to changes in grain interlocking or clustering. Our data provide a unique look into the dynamics of coarse sediment transport in the field under rapidly changing hydraulic forcing.

Acknowledgements

The work was funded by the National Science Foundation (EAR 1053508 to JPLJ) and by The University of Texas at Austin Jackson School of Geosciences. All authors contributed to project design and to field work. KLGP developed the methodologies used and conducted most of the data analysis. KLGP and JPLJ wrote the manuscript, with input from DNB. We thank Lindsay Olinde for informative discussions. We will also thank the journal editors and reviewers. Data will be available in a data archive that complies with WRR guidelines.

Figures and tables

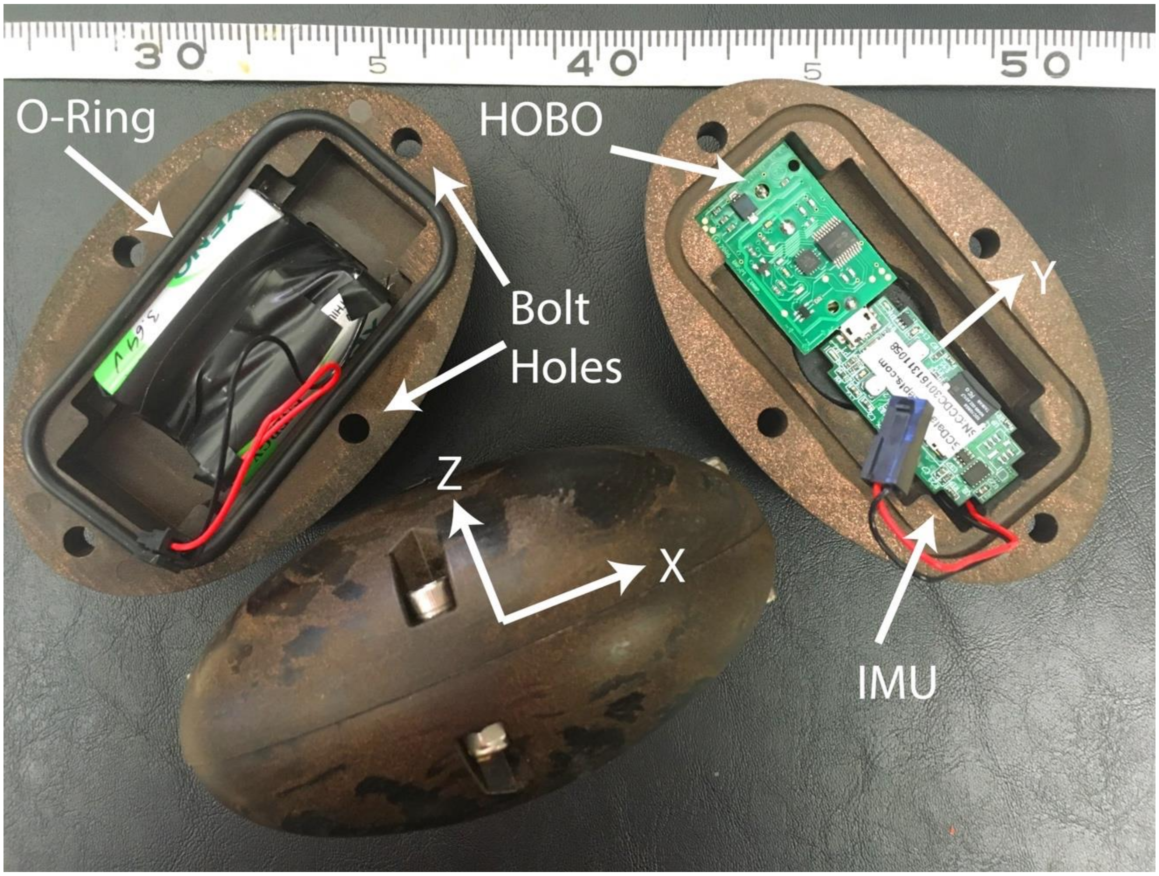

Figure 1. Tracer high-density plastic case with HOBO, IMU, and battery pack (left, mostly covered in black electrical tape) visible. The tape measure is in cm. Major, intermediate, and minor axis diameters are 12.0, 7.2, and 6.4 cm, respectively. X, Y, and Z axes for the IMU are indicated.

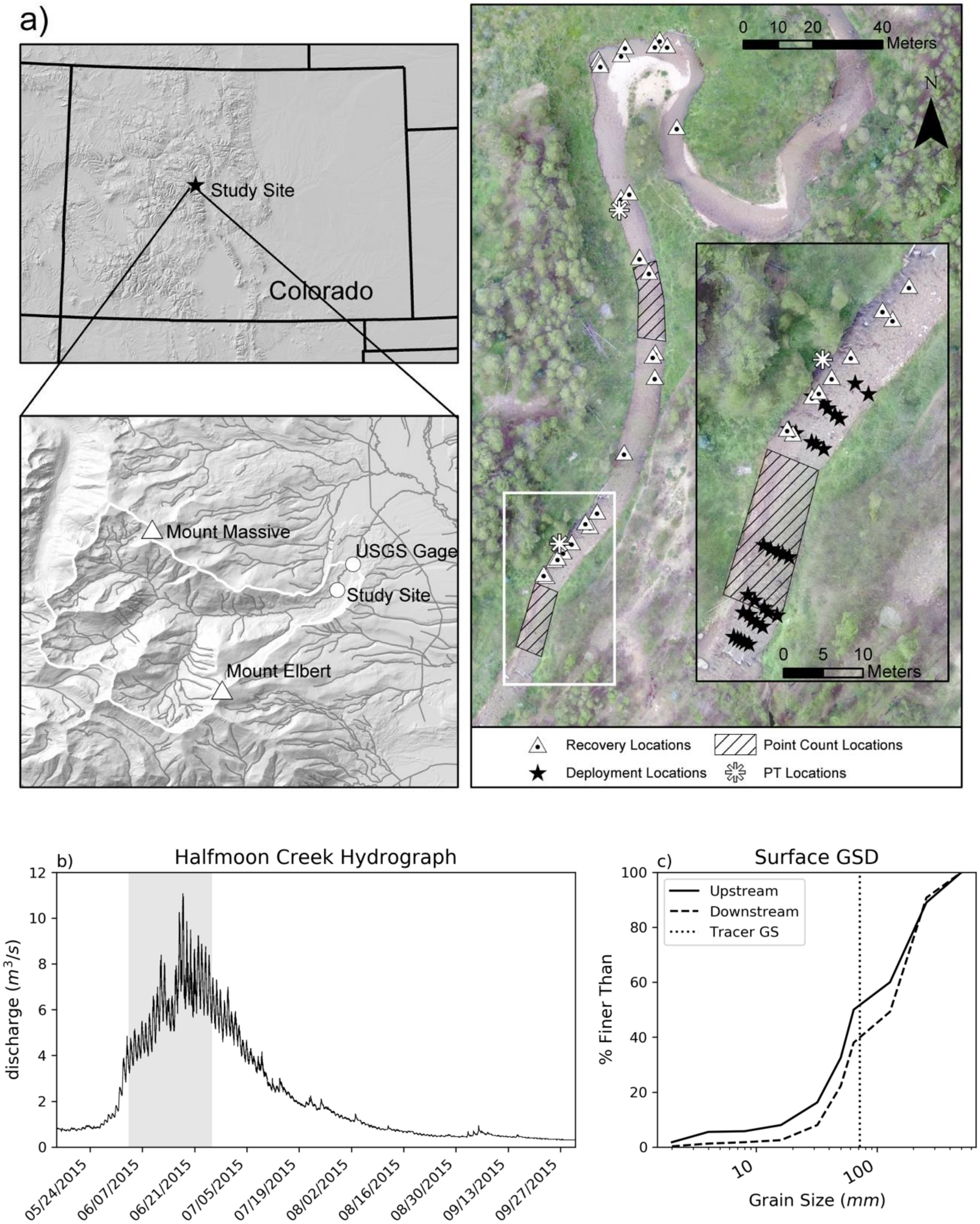

Figure 2. (a) Halfmoon Creek drainage area and project site, including installation and recovery locations of tracers. "PT" means pressure transducers, used to monitor reach

flow depth. (b) Summer 2015 Halfmoon Creek Hydrograph from USGS gage 07083000. Shaded grey region indicates 22 diurnal events used in the threshold of motion analysis. c. Bed surface GSD from two Wolman-type point counts (N=400). Dotted line is intermediate diameter of tracers.

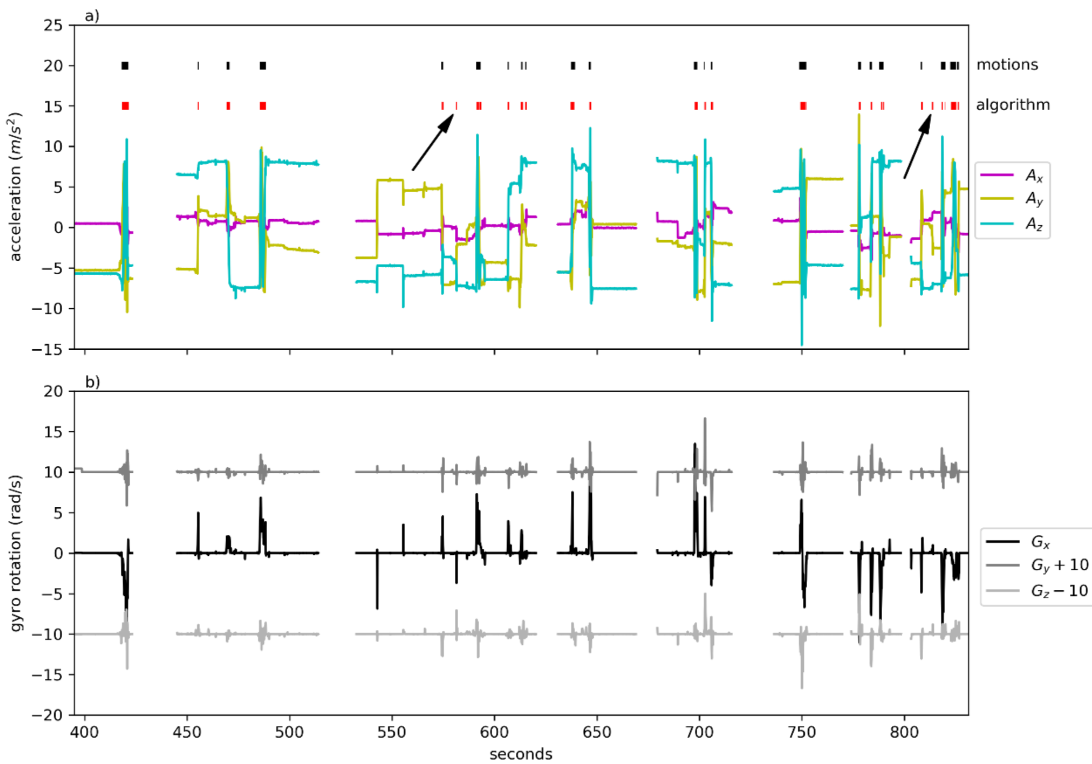

Figure 3. (a) Acceleration and (b) gyroscope data from experimental motion test. Black bars indicate when particle displacement actually occurred. Red bars indicate motions identified by the algorithm. Data not plotted when test particle was repositioned inside video frame. Arrows indicate two times identified by algorithm as motion but were not a displacement.

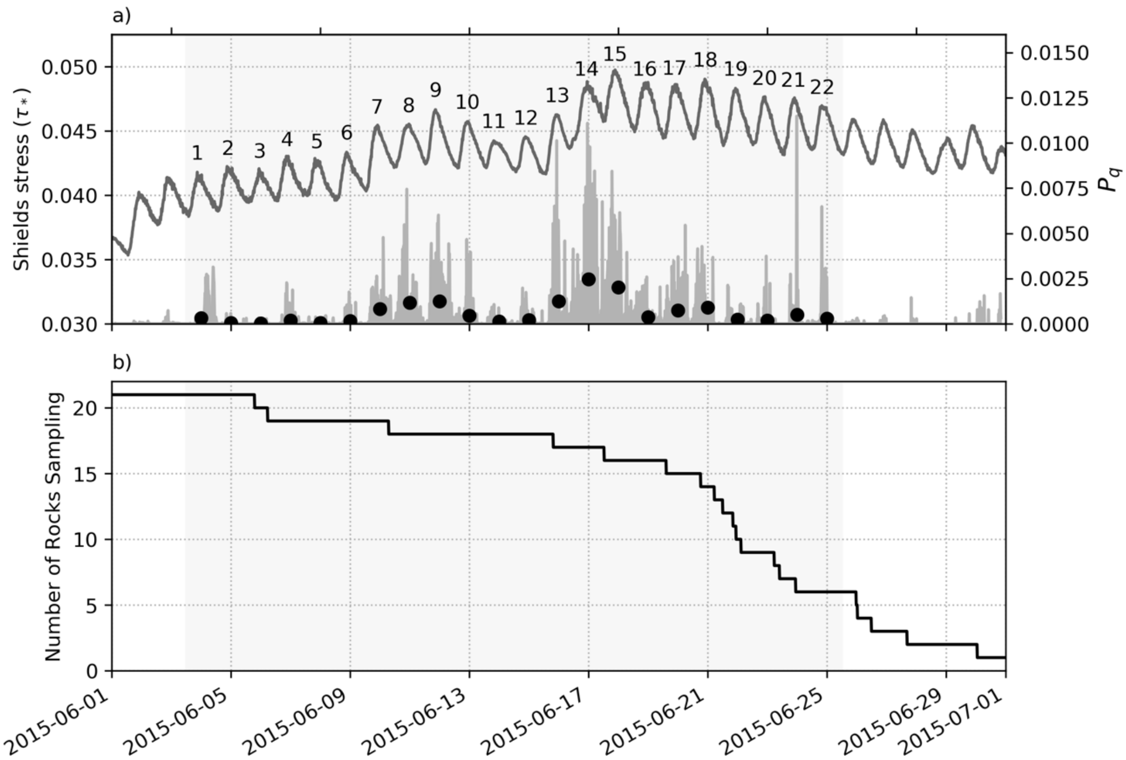

Figure 4. (a) Shields stress (dark gray lines) and 15-minute P_q (light grey) verses time. The shaded region covers daily flood events 1-22 that are used in the threshold of motion analysis. Black points are event-averaged values of P_q. (b) Number of rocks sampling over the 22-day analysis window.

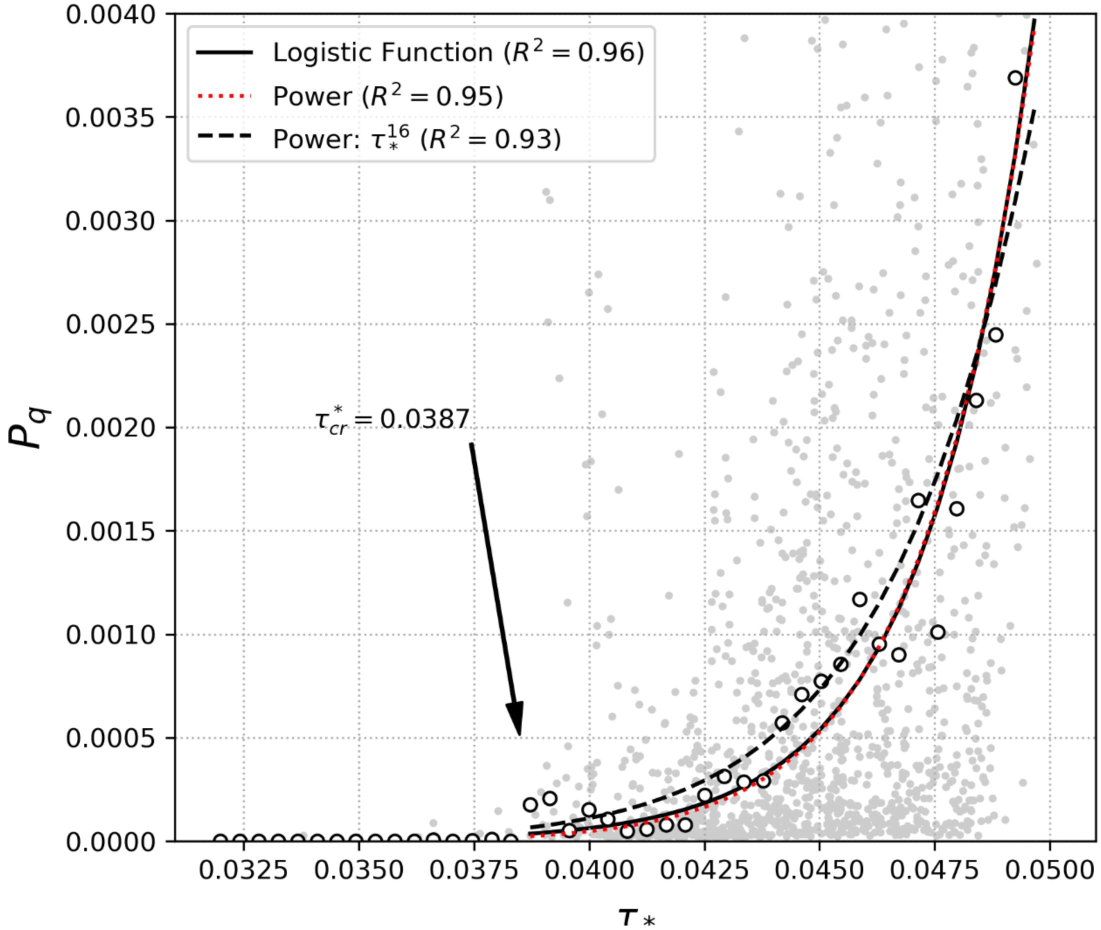

Figure 5. Gray dots represent unsmoothed 15 minute data. The open circles are bin-averaged data, in increments of τ^*=0.0004. The "Logistic Function" and "Power" refer to equations (3) and (4) respectively. A small number of unsmoothed 15 minute data fall above P_q=0.004.

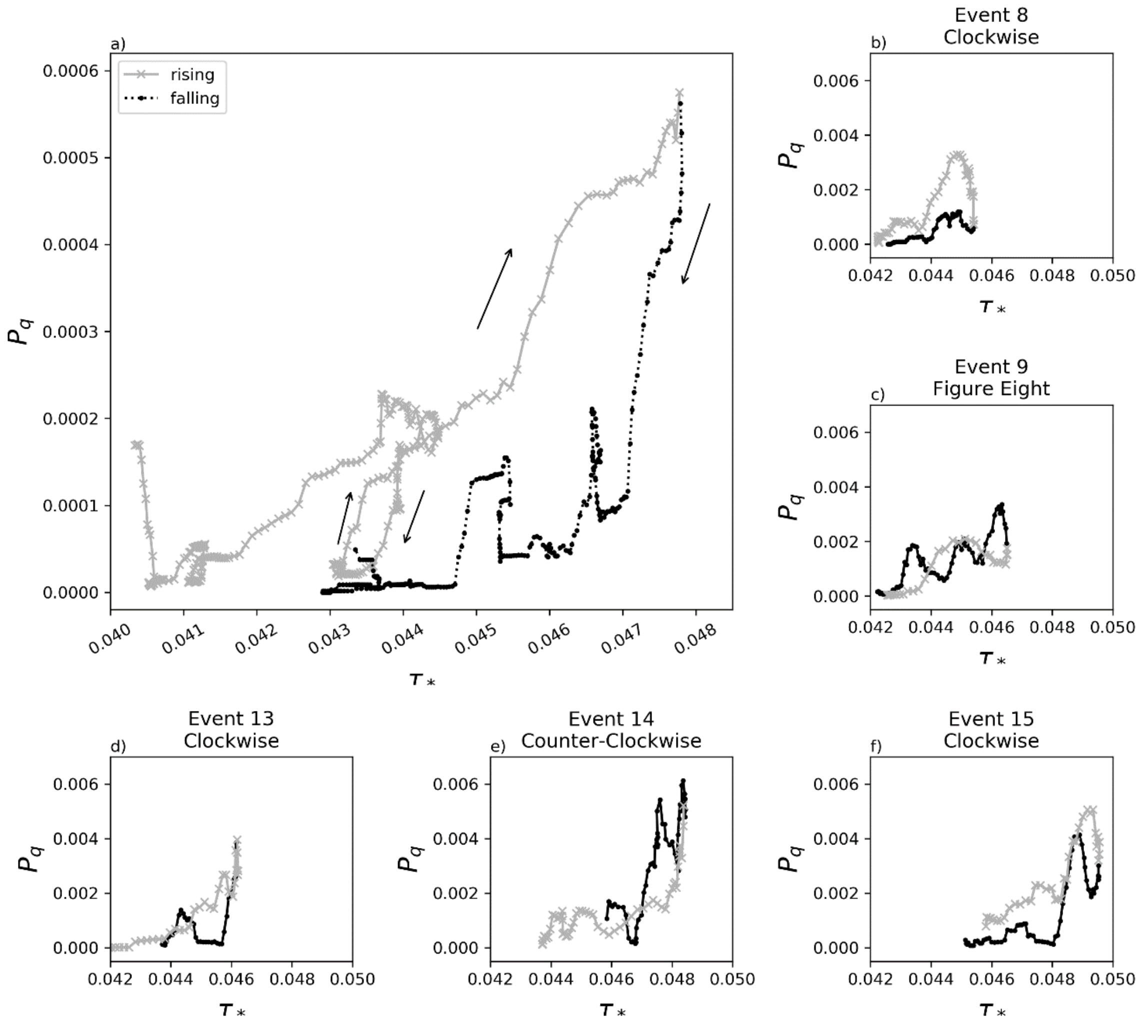

Figure 6. (a) Hysteresis curve for 22 days of recorded data. P_q and τ_* were smoothed using a 24-hour median moving window, plotted every hour. Calculated over 22 flood events, transition from rising to falling coincides with peak flow in event 15. (b-f) Examples of clockwise (b, d, f), figure-eight (c), and counter-clockwise (e) hysteresis patterns found in daily flood events, calculated using a 2-hour median moving window plotted every 15 minutes. Hysteresis plots for all 22 hydrographs are in the supplementary material.

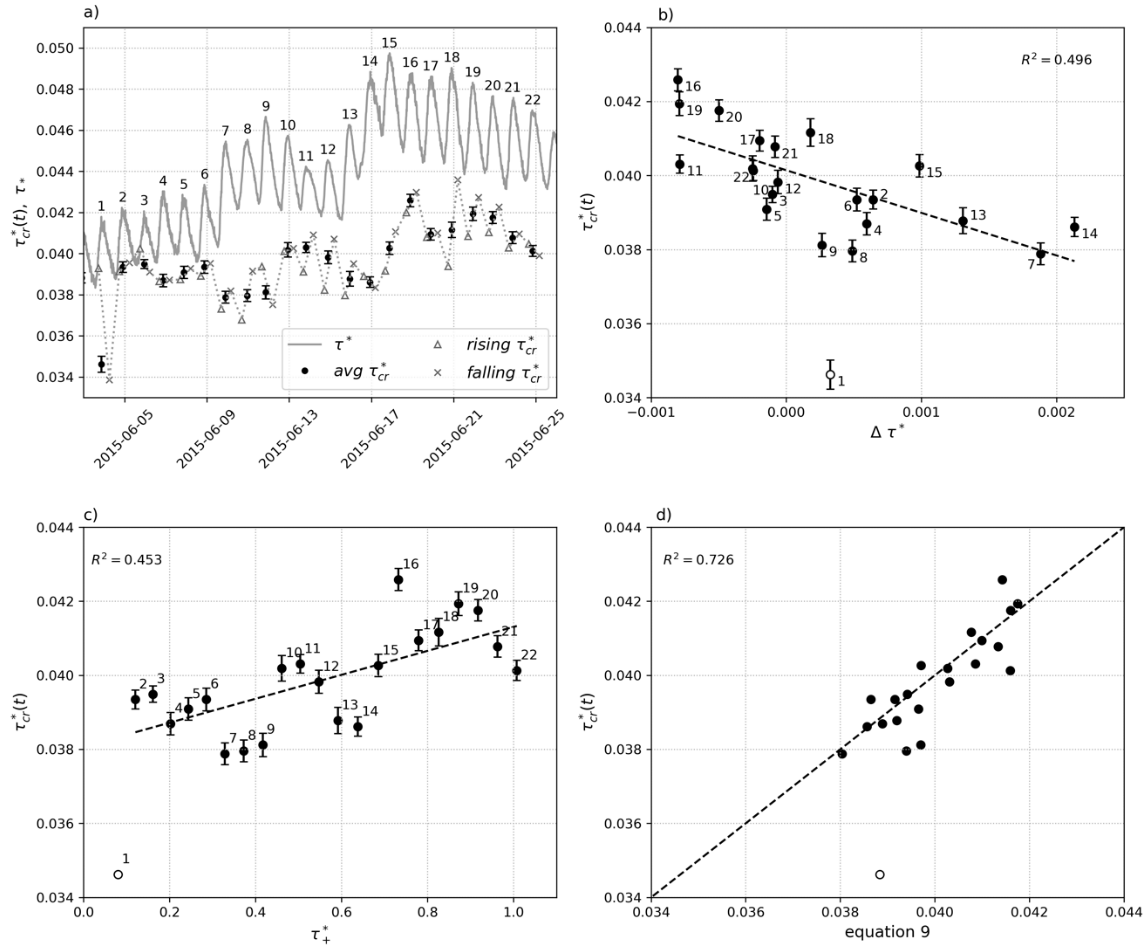

Figure 7 (a) Daily threshold of motion, averaged over the duration of each daily flood event (avg $\tau^*_{cr}(t)$), and also averaged separately over each rising and falling limb. Error bars are standard deviation of threshold of motion for 15-minute data; the grey triangles and circles represent the critical threshold of motion for each rising and falling limb, respectively. (b) Linear regression between $\Delta\tau^*$ and $\tau^*_{cr}(t)$, (c) Linear regression between $\tau^*{}_+$ and $\tau^*_{cr}(t)$, and (d) equation (9) vs $\tau^*_{cr}(t)$.

Table 1. Regression results

Single Linear Regression Results

Parameter	R^2	*p-value*	*Equation*
$\tau^*{}_+$	0.453	0.0008	$\tau^*_{cr}(t) = 0.003\tau^*{}_+ + 0.0383$
$\Delta\tau^*$	0.496	0.0004	$\tau^*_{cr}(t) = -1.151\Delta\tau^* + 0.0401$
$\overline{\tau^*}$	0.265	0.0170	$\tau^*_{cr}(t) = 0.314\overline{\tau^*} + 0.0260$
$\widehat{\tau^*}$	0.248	0.0217	$\tau^*_{cr}(t) = 0.272\widehat{\tau^*} + 0.0027$
$\tau^*{}_{rf}$	0.107	0.1475	$\tau^*_{cr}(t) = -0.321\tau^*{}_{rf} + 0.0396$

Two-parameter Regression Results

Parameters	R^2	*p-value*	*Equation*
$\Delta\tau^*, \tau^*{}_+$	0.726	<0.0001	$\tau^*_{cr}(t) = 0.002\tau^*{}_+ - 0.897\Delta\tau^* + 0.0389$ (Equation (9))
$\Delta\tau^*, \overline{\tau^*}$	0.747	<0.0001	$\tau^*_{cr}(t) = 0.305\overline{\tau^*} - 1.135\Delta\tau^* + 0.0275$
$\Delta\tau^*, \widehat{\tau^*}$	0.748	<0.0001	$\tau^*_{cr}(t) = 0.271\widehat{\tau^*} - 1.156\Delta\tau^* + 0.0266$

Multiple Linear Regression Results

Parameters	R^2	*p-value*	*Equation*
	0.077	0.0002	$\tau^*_{cr}(t) = 0.0008\tau^*{}_+ - 1.061\Delta\tau^* - 0.244\overline{\tau^*} + 0.393\widehat{\tau^*} - 0.144\tau^*{}_{rf} + 0.0323$

	R^2 *without parameter*	*Paired t-test p-value of coefficient*
$\tau^*{}_+$	0.453	0.5368
$\Delta\tau^*$	0.496	0.0006
$\overline{\tau^*}$	0.265	0.7493
$\widehat{\tau^*}$	0.248	0.5638
$\tau^*{}_{rf}$	0.107	0.3071

6. References

Alexandrov, Y., Laronne, J.B., and Reid, I. (2007). Intra-event and inter-seasonal behaviour of suspended sediment in flash floods of the semi-arid northern Negev, Israel: Geomorphology, v. 85, no. 1-2, p. 85–97, doi: 10.1016/j.geomorph.2006.03.013 .

Bombar, G., Elçi, Ş., Tayfur, G., Güney, M., and Bor, A. (2011). Experimental and Numerical Investigation of Bed-Load Transport under Unsteady Flows: J. Hydraul. Eng, v. 137, no. October, p. 1276–1282, doi: 10.1061/(ASCE)HY.1943-7900.0000412.

Bradley, D.N, and Tucker, G.E., 2012, Measuring gravel transport and dispersion in a mountain river using passive radio tracers: Earth Surface Processes and Landforms, v. 37, no. 10, p. 1034–1045, doi: 10.1002/esp.3223.

Bradley, D. N. (2017). Direct Observation of Heavy-Tailed Storage Times of Bed Load Tracer Particles Causing Anomalous Superdiffusion. Geophysical Research Letters, 44(24), 12227-12235. doi:10.1002/2017gl075045

Buffington, J.M., and Montgomery, D.R., 1997, A systematic analysis of eight decades of incipient motion studies, with special reference to gravel-bedded rivers: Water Resources Research, v. 33, no. 8, p. 1993–2029, doi: 10.1029/96WR03190.

Bunte, K., Abt, S. R., Swingle, K. W., Cenderelli, D. A., & Schneider, J. M. (2013). Critical Shields values in coarse-bedded steep streams. Water Resources Research, 49(11), 7427-7447. doi:10.1002/2012wr012672

Church, M., Hassan, M. a., and Wolcott, J.F., 1998, Stabilizing self-organized structures in gravel-bed stream channels: Field and experimental observations: Water Resources Research, v. 34, no. 11, p. 3169–3179, doi: 10.1029/98WR00484 .

Dunne, K. B. J. and Jerolmack, D. J. (2018). Evidence of, and a proposed explanation for, bimodal transport states in alluvial rivers, Earth Surf. Dynam., 6, 583–594, https://doi.org/10.5194/esurf-6-583-2018 .

Gilbert, G.K. (1914). Transport of debris by running water. - U.S. Geol. Sur. Prof. Paper 86. US Government Printing Office.

Gilbert, G.K. (1917). Hydraulic-Mining debris in the Sierra Nevada. - U.S. Geol. Sur. Prof. Paper 105. US Government Printing Office.

Gimbert, F., Fuller, B. M., Lamb, M. P., Tsai, V. C., & Johnson, J. P. L. (2019). Particle transport mechanics and induced seismic noise in steep flume experiments with accelerometer-embedded tracers. Earth Surface Processes and Landforms, 44(1), 219-241. doi:10.1002/esp.4495

Gronz, O., Hiller, P. H., Wirtz, S., Becker, K., Iserloh, T., Seeger, M., . . . Ries, J. B. (2016). Smartstones: A small 9-axis sensor implanted in stones to track their movements. Catena, 142, 245-251. doi:https://doi.org/10.1016/j.catena.2016.03.030

Haschenburger, J. K. (2011). The rate of fluvial gravel dispersion. Geophysical Research Letters, 38(24). doi:10.1029/2011gl049928

Hassan, M.A., and Church, M., 2000, Experiments on surface structure and partial sediment transport on a gravel bed: Water Resources Research, v. 36, no. 7, p. 1885, doi: 10.1029/2000WR900055.

Hassan, M. a., and Reid, I., 1990, The influence of microform bed roughness elements on flow and sediment transport in gravel bed rivers: Earth Surface Processes and Landforms, v. 15, no. 8, p. 739–750, doi: 10.1002/esp.3290150807.

Hassan, M. A., Saletti, M., Zhang, C., Ferrer-Boix, C., Johnson, J. P. L., Müller, T., and von Flotow, C. (2020) Co-evolution of coarse grain structuring and bed roughness in response to episodic sediment supply in an experimental aggrading channel. *Earth Surf. Process. Landforms*, 45: 948– 961. https://doi.org/10.1002/esp.4788.

Humphries, R., Venditti, J. G., Sklar, L. S., & Wooster, J. K. (2012). Experimental evidence for the effect of hydrographs on sediment pulse dynamics in gravel-bedded rivers. Water Resources Research, 48(1). doi:10.1029/2011wr010419

Johnson, J. P. L. (2016). Gravel threshold of motion: a state function of sediment transport disequilibrium? Earth Surface Dynamics, 4(3), 685-703. doi:10.5194/esurf-4-685-2016

Johnson, J. P. L. (2017). Clustering statistics, roughness feedbacks, and randomness in experimental step-pool morphodynamics. Geophysical Research Letters, 44(8), 3653-3662. doi:10.1002/2016gl072246

Kirchner, J.W., Dietrich, W.E., Iseya, F., and Ikeda, H. (1990). The variability of critical shear stress, friction angle, and grain protrusion in waterworked sediments.: Sedimentology, v. 37, no. 4, p. 647, doi: Article.

Kuhnle, R.A., 1992, Bed load transport during rising and falling stages on two small streams: Earth Surface Processes and Landforms, v. 17, no. 2, p. 191–197, doi: 10.1002/esp.3290170206.

Lamb, M. P., Dietrich, W. E., & Venditti, J. G. (2008). Is the critical Shields stress for incipient sediment motion dependent on channel-bed slope? Journal of Geophysical Research, 113(F2), F02008. doi:10.1029/2007jf000831

Lenzi, M. A., Mao, L., & Comiti, F. (2004). Magnitude-frequency analysis of bed load data in an Alpine boulder bed stream. Water Resources Research, 40. doi:10.1029/2003WR002961

Maniatis, G., Hoey, T. B., Hassan, M. A., Sventek, J., Hodge, R., Drysdale, T., & Valyrakis, M. (2017). Calculating the Explicit Probability of Entrainment Based on Inertial Acceleration Measurements. Journal of Hydraulic Engineering, 143(4), 04016097. doi:doi:10.1061/(ASCE)HY.1943-7900.0001262

Mao, L., Dell'Agnese, A., Huincache, C., Penna, D., Engel, M., Niedrist, G., & Comiti, F. (2014). Bedload hysteresis in a glacier-fed mountain river. Earth Surface Processes and Landforms, 39(7), 964-976. doi:10.1002/esp.3563

Mao, L. C. (2012). The effect of hydrographs on bed load transport and bed sediment spatial arrangement. Journal of Geophysical Research-Earth Surface, 117. doi:10.1029/2012jf002428

Marquis, G. A., and A. G. Roy (2012), Using multiple bed load measurements: Toward the identification of bed dilation and contraction in gravel-bed rivers, J. Geophys. Res., 117, F01014, doi:10.1029/2011JF002120.

Martin, R.L., and Jerolmack, D.J. (2013). Origin of hysteresis in bed form response to unsteady flows: Water Resources Research, v. 49, no. 3, p. 1314–1333, doi: 10.1002/wrcr.20093.

Masteller, C. C., & Finnegan, N. J. (2017). Interplay between grain protrusion and sediment entrainment in an experimental flume. Journal of Geophysical Research-Earth Surface, 122(1), 274-289. doi:10.1002/2016jf003943

Masteller, C. C., Finnegan, N. J., Turowski, J. M., Yager, E. M., & Rickenmann, D. (2019). History-Dependent Threshold for Motion Revealed by Continuous Bedload Transport Measurements in a Steep Mountain Stream. Geophysical Research Letters, 46(5), 2583-2591. doi:10.1029/2018gl081325

Meirovich, L., Laronne, J. B., & Reid, I. (1998). The variation of water-surface slope and its significance for bedload transport during floods in gravel-bed streams. Journal of Hydraulic Research, 36(2), 147-157. doi:10.1080/00221689809498630

Meyer-Peter, E., and Müller, R. (1948). Formulas for bed-load transport: *Proc., 2nd Meeting*, IAHR, Stockholm, Sweden, 39–64.

Milly, P.C.D., Wetherald, R.T., Dunne, K. a, and Delworth, T.L., 2002, Increasing risk of great floods in a changing climate: Nature, v. 415, no. 6871, p. 514–517, doi: 10.1038/415514a.

Moog, D. B., & Whiting, P. J. (1998). Annual hysteresis in bed load rating curves. Water Resources Research, 34(9), 2393-2399. doi:10.1029/98wr01658

Mueller, E. R., and Pitlick, J. (2005), Morphologically based model of bed load transport capacity in a headwater stream, *J. Geophys. Res.*, 110, F02016, doi:10.1029/2003JF000117.

Mueller, E. R., Pitlick, J., and Nelson, J. M. (2005), Variation in the reference Shields stress for bed load transport in gravel-bed streams and rivers, *Water Resour. Res.*, 41, W04006, doi:10.1029/2004WR003692.

Nathan Bradley, D., and Tucker, G.E., 2012, Measuring gravel transport and dispersion in a mountain river using passive radio tracers: Earth Surface Processes and Landforms, v. 37, no. 10, p. 1034–1045, doi: 10.1002/esp.3223.

Ockelford, A. M., & Haynes, H. (2013). The impact of stress history on bed structure. Earth Surface Processes and Landforms, 38(7), 717-727. doi:10.1002/esp.3348

Olinde, L., and Johnson, J.P.L. (2015). Using RFID and accelerometer-embedded tracers to measure probabilities of bed load transport, step lengths, and rest times in a mountain stream: Water Resources Research, v. 51, no. 9, p. 7572–7589, doi: 10.1002/2014WR016120.

Paphitis, D., & Collins, M. B. (2005). Sand grain threshold, in relation to bed 'stress history': an experimental study. Sedimentology, 52(4), 827-838. doi:10.1111/j.1365-3091.2005.00710.x

Parker, G. (1978). Self-formed straight rivers with equilibrium banks and mobile bed. Part II. The gravel river. Journal of Fluid Mechanics, 89(1), 127-148. doi: https://doi.org/10.1017/S0022112078002505

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

Paintal, A.S. (1971). Concept Of Critical Shear Stress In Loose Boundary Open Channels, Journal of Hydraulic Research, 9:1, 91-113, DOI: 10.1080/00221687109500339

Pfeiffer, A. M., & Finnegan, N. J. (2018). Regional Variation in Gravel Riverbed Mobility, Controlled by Hydrologic Regime and Sediment Supply. Geophysical Research Letters, 45(7), 3097-3106. doi:10.1002/2017gl076747

Phillips, C. B., & Jerolmack, D. J. (2016). Self-organization of river channels as a critical filter on climate signals. Science, 352(6286), 694-697. doi:10.1126/science.aad3348

Phillips, C. B., & Jerolmack, D. J. (2019). Bankfull Transport Capacity and the Threshold of Motion in Coarse-Grained Rivers. Water Resources Research, 55(12), 11316-11330. doi:10.1029/2019wr025455

Piedra, M.M., Haynes, H., and Hoey, T.B., 2012, The spatial distribution of coarse surface grains and the stability of gravel river beds: Sedimentology, v. 59, no. 3, p. 1014–1029, doi: 10.1111/j.1365-3091.2011.01290.x.

Powell, D. M., Reid, I., & Laronne, J. B. (2001). Evolution of bed load grain size distribution with increasing flow strength and the effect of flow duration on the caliber of bed load sediment yield in ephemeral gravel bed rivers. Water Resources Research, 37(5), 1463-1474. doi:10.1029/2000wr900342

Recking, A. (2012). Influence of sediment supply on mountain streams bedload transport. Geomorphology, 175, 139-150. doi:10.1016/j.geomorph.2012.07.005

Rickenmann, D., 2001, Comparison of bed load transport in torrents and gravel bed streams: Water Resource Research, v. 37, no. 12, p. 3295–3305, doi: 10.1029/2001WR000319.

Roth, D. L., Finnegan, N. J., Brodsky, E. E., Cook, K. L., Stark, C. P., & Wang, H. W. (2014). Migration of a coarse fluvial sediment pulse detected by hysteresis in bedload generated seismic waves. Earth and Planetary Science Letters, 404, 144-153. doi:10.1016/j.epsl.2014.07.019

Sanguinito, S., & Johnson, J. (2012). Quantifying gravel overlap and dislodgement forces on natural river bars: implications for particle entrainment. Earth Surface Processes and Landforms, 37(1), 134-141. doi:10.1002/esp.2237

Sklar, L. S., & Dietrich, W. E. (2004). A mechanistic model for river incision into bedrock by saltating bed load. Water Resources Research, 40(6). doi:10.1029/2003wr002496

Strom, K., Papanicolaou, A. N., Evangelopoulos, N., & Odeh, M. (2004). Microforms in gravel bed rivers: Formation, disintegration, and effects on bedload transport. Journal of Hydraulic Engineering-Asce, 130(6), 554-567. doi:10.1061/(asce)0733-9429(2004)130:6(544)

Torizzo, M., & Pitlick, J. (2004). Magnitude-frequency of bed load transport in mountain streams in Colorado. Journal of Hydrology, 290(1-2), 137-151. doi:doi:10.1016/j.jhydrol.2003.12.001

Turowski, J. M., Badoux, A., & Rickenmann, D. (2011). Start and end of bedload transport in gravel-bed streams. Geophysical Research Letters, 38(4), L04401. doi:10.1029/2010gl046558

Venditti, J. G., Dietrich, W. E., Nelson, P. A., Wydzga, M. A., Fadde, J., & Sklar, L. (2010). Mobilization of coarse surface layers in gravel-bedded rivers by finer gravel bed load. Water Resources Research, 46(7). doi:10.1029/2009wr008329

Viparelli, E., Gaeuman, D., Wilcock, P., & Parker, G. (2011). A model to predict the evolution of a gravel bed river under an imposed cyclic hydrograph and its application to the Trinity River. Water Resources Research, 47(2). doi:10.1029/2010WR009164

Wilcock, P. R., & Crowe, J. C. (2003). Surface-based transport model for mixed-size sediment. Journal of Hydraulic Engineering-Asce, 129(2), 120-128. doi:10.1061/(asce)0733-9429(2003)129:2(120)

Wong, M., & Parker, G. (2006). Reanalysis and Correction of Bed-Load Relation of Meyer-Peter and Müller Using Their Own Database. Journal of Hydraulic Engineering, 132(11), 1159-1168. doi:10.1061/(ASCE)0733-9429(2006)132:11(1159)

Yager, E. M., Turowski, J. M., Rickenmann, D., & McArdell, B. W. (2012). Sediment supply, grain protrusion, and bedload transport in mountain streams. Geophysical Research Letters, 39. doi:10.1029/2012gl051654

Yager, E. M., Schmeeckle, M. W., & Badoux, A. (2018). Resistance is not futile: Grain resistance controls on observed critical Shields stress variations. Journal of

Geophysical Research: Earth Surface, 123, 3308–3322. https://doi.org/10.1029/2018JF004817