Received: 8 February 2023 | Revised: 31 December 2023 | Accepted: 8 January 2024

DOI: 10.1002/esp.5774

RESEARCH ARTICLE

ESPL WILEY

Strengths and weaknesses of acoustic pipe microphone systems in ephemeral, sandy, gravel-bed rivers

Kyle Stark[1,2] | Daniel Cadol[1] | David Varyu[3] | Eran Halfi[4,5] | Jonathan B. Laronne[4]

[1]New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA

[2]San Francisco Estuary Institute, Richmond, California, USA

[3]US Bureau of Reclamation Technical Service Center, Denver, Colorado, USA

[4]Ben Gurion University of the Negev, Be'er Sheva, Israel

[5]Dead Sea and Arava Science Center, Mount Masada, Israel

Correspondence
Kyle Stark, New Mexico Institute of Mining and Technology, Socorro, NM, USA.
Email: kstark131@gmail.com

Funding information
U.S. Bureau of Reclamation, Grant/Award Number: 9781; National Science Foundation, Grant/Award Number: EAR-1852794; New Mexico Geological Society; New Mexico Water Resources Research Institute

Abstract

We calibrated an acoustic pipe microphone system to monitor bedload flux in a sandy, gravel-bed ephemeral channel. Ours is a first attempt to test the limit of an acoustic surrogate bedload system in a channel with a high content of sand. Calibrations varied in quality; significant data subsetting was required to achieve R^2 values >0.75. Several data quality issues had to be addressed: (1) apparent pulses, which occur when a sensor records an impulse from sediment impacting the surrounding substrate rather than directly impacting the sensor, were frequent, especially at higher signal amplifications. (2) The impact sensors were frequently covered by gravel sheets. This prompted the development of a cover detection protocol that rejected part of the impact sensor record when at least one sensor was partially or fully covered. (3) Because of the lack of sensor sensitivity to impacts of sand-sized particles, which was anticipated, and the considerable sand component of bedload in this channel, a grain size-limited bedload flux was estimated. This was accomplished by sampling the bedload captured by slot samplers and evaluating the variation of grain size with increasing flow strength. This considerably improved the results when compared to attempts at estimating the flux of the entire distribution of grain sizes. This calibration is a successful first attempt, though the impact sensors required several site-specific calibration steps. A universal set of equations using impact sensors to estimate bedload transport of fine-gravel with a large content of sand remains elusive. Notwithstanding, our study demonstrates the utility of impact sensor data, producing relatively low root mean square errors that are independent of measurements of flow strength (i.e. discharge). These tools will be particularly useful in settings that would benefit from new methodologies for estimating bedload transport in sand-rich gravel-bed rivers, such as the American desert Southwest.

KEYWORDS
acoustic impact sensor, bedload, ephemeral channel, gravel-bed river, sand, surrogate

1 | INTRODUCTION

Bedload transport is notoriously difficult to accurately monitor in open channel settings. It is highly variable temporally and spatially due to complex phenomena such as turbulence and bedform migration. These phenomena often cause bedload flux to be nonlinearly related to predictive stream parameters (e.g. shear stress). This causes difficulty in calibrating transport equations (Recking, 2010), though in some ephemeral settings, bedload response to shear is simple (Cohen et al., 2010; Reid & Laronne, 1995). Even so, calibration difficulties are often exacerbated in arid environments, where flow events are rare and flashy in nature, and access to field locations can be challenging.

One solution is to use surrogate instruments to monitor bedload transport. These instruments indirectly monitor bedload by recording the acoustic or seismic vibrations generated due to bedload transport. They have the benefit of being automated and continuously active, with none of the capacity limits that affect direct samplers. Acoustic

monitoring of bedload has been under development since the 1970s (Anderson, 1976; Banziger & Burch, 1990; Richards & Milne, 1979; Thorne, 1986). The development of these instruments has divided along two separate paths: impact-driven systems and passive listening systems. Listening systems utilize hydrophones—microphones placed directly into the water column—to record a wide range of signals, including information related to turbulence and other hydrodynamic processes. These signals can be filtered to isolate acoustic energy related to bedload transport (Geay, Belleudy, Gervaise, et al., 2017). Hydrophone instruments have had demonstrated success in large, wide channels (Geay, Belleudy, Laronne, et al., 2017; Marineau et al., 2016) but limited success in smaller systems.

Impact-driven systems insulate acoustic instruments, typically within a metal pipe or underneath a metal plate, to reduce the signals related to hydrodynamic processes and specifically to target signals arising from bedload transport. This style of instrument typically uses geophones (Hilldale et al., 2014; Rickenmann et al., 2014) or microphones (Halfi et al., 2020; Mizuyama, Laronne, et al., 2010) to monitor the acoustic response of impacts by grains in transport. Some instruments record the full waveform associated with the impact (Nicollier et al., 2020; Wyss et al., 2016b), but this is computationally expensive to record and store. An alternative technique has been refined, such that signal processing procedures produce more manageable datasets that permit quick comparison to bedload data. This technique introduces an impulse-counting concept, reducing the waveform into manageable datasets of recorded pulses. Rickenmann (2017) reviewed the recent efforts to monitor bedload flux through vibration and acoustics using these impact-driven systems, maintaining that the methodology has reached a level of maturity where both total bedload flux and grain size characteristics have been successfully determined in individual rivers. In each case, however, site-specific calibration is still required.

Calibration usually involves collecting samples of bedload for an extended period of time followed by rigorous data analysis. This step of the analysis differs depending on the impact sensor setup but can involve averaging impacts and bedload flux over long periods (Mizuyama, Laronne, et al., 2010), incorporating grain size information in predictive models (Halfi et al., 2020; Mao et al., 2016), incorporating flow velocity into the prediction (Nicollier et al., 2021) or removing erroneous data points due to issues with covered sensors (this study). Common issues reported in calibration include strong effects of grain size on calibration (Mao et al., 2016), differences in acoustic sensor calibration between rising and recession limbs (Halfi et al., 2020), and flow-related limitations (Rickenmann, 2017). In our review of previous studies, many of the challenges for successful sensor calibration were often resolved by limiting or subsetting the datasets used to calibrate the acoustic impact sensor.

To date, these systems have been almost exclusively evaluated in boulder and cobble-bed mountain streams. No efforts have been attempted in sand-rich gravelly environments. Here, we present new calibration data and methods from a sand-rich, gravel-bed ephemeral channel in central New Mexico, USA. Our goals are to:

I. evaluate the effectiveness of impact-driven acoustic surrogate instruments in sand-rich, gravel-bedded channels;
II. use *in situ* field data to address the control of grain size on the surrogate instrument response;
III. compare these results to those using similar experimental setups; and
IV. develop a calibration methodology that may be tested at other equipment installations.

2 | SITE DESCRIPTION AND METHODS

The Arroyo de los Pinos is a 32 km^2 watershed in semi-arid central New Mexico, USA. A detailed description of the site and most of the equipment is available in prior publications (Stark et al., 2021); relevant details and additional equipment descriptions are presented herein. Mean annual rainfall is 240 mm, most of which arrives during high-intensity summer monsoon storms. The Pinos is ephemeral; it is dry more than 99% of the time, flowing three to five times a year. The principal sediment monitoring station resides near the basin outlet (Figures 1 and 2). At the monitoring station, the channel thalweg bed material consists of sand (36%), 2–8 mm gravel (30%) and >8 mm gravel (33%). Silt and clay sized sediment is a minor component, typically comprising 1%–3%. Channel bars are coarser-grained, but only inundated during periods of high flow. The addition of the coarser-grained material from channel bars increases the proportion of bedload coarser >8 mm.

Sediment transported as bedload was physically captured using three Reid-type slot samplers (Reid et al., 1980). In a slot sampler system, bedload falls into a chamber below the bed surface. Wings on the trap entrance prevent lateral movement of sediment into the trap, reducing oversampling from sediment moving laterally. These samplers provide accurate estimates of bedload transport and its grain size distribution (Poreh et al., 1970) but come with a high labour and monetary cost of installation and operation. Slot samplers have capacity limits, such that they typically fill prior to sampling at high water stages. Once the samplers attain ~80% fill, their sampling efficiency rapidly decreases (Habersack et al., 2001). As a sampler's efficiency began to decrease, data from that sampler were excluded from analysis. Bedload flux data were processed following the mass aggregation method developed by Halfi et al. (2020). Data were only assessed once a minimum of 5 kg were collected within the sampler, rather than at regular time intervals (Stark et al., 2021). After flow events, samples were collected from the traps for grain size analysis. A methodological change to the bedload sampling procedure in 2021 limits the ability of data collected from prior years to be used in the calibration process. Prior samples were collected only from the middle of each sampler, but since 2021, separate samples have been taken from both the middle and side of the samplers, as the sediment falling into a sampler develops into a spatio-texturally varying cone. Bedload grain size data were available from the 2018–2020 events, but the information from these events only partially captured the true variance of grain size with changing flow conditions and is not perfectly comparable to the new method. Based on the new methodology and the information gained from the 2021 events, data from previous years were bias-corrected to utilize the entire bedload grain size dataset. The methodological change in 2021 means that the bedload grain size data from prior years have been used to describe trends in grain size, but not to directly calibrate the surrogate instruments. Graphical depictions and extended descriptions of the sampling methods are available in Supporting Information.

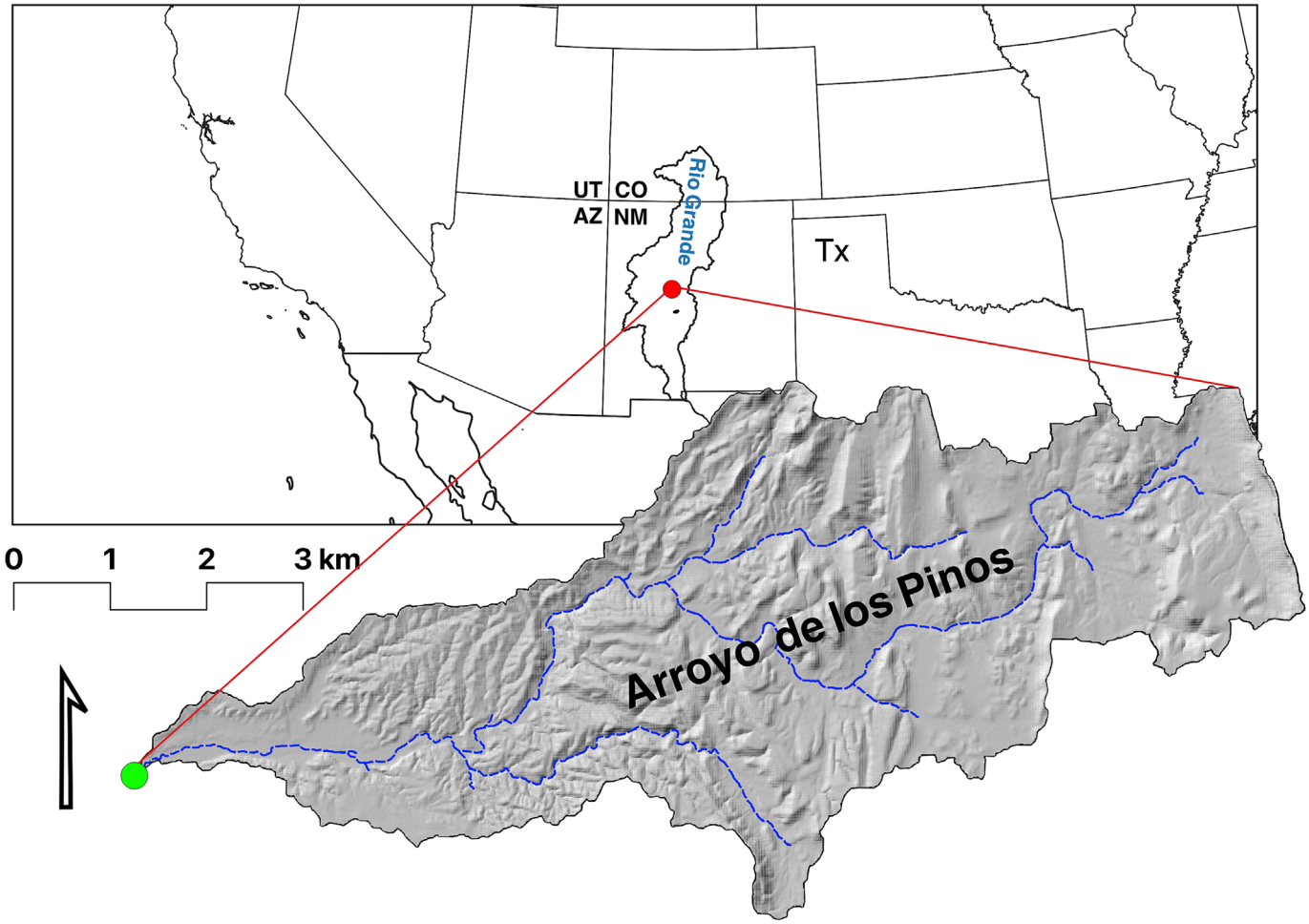
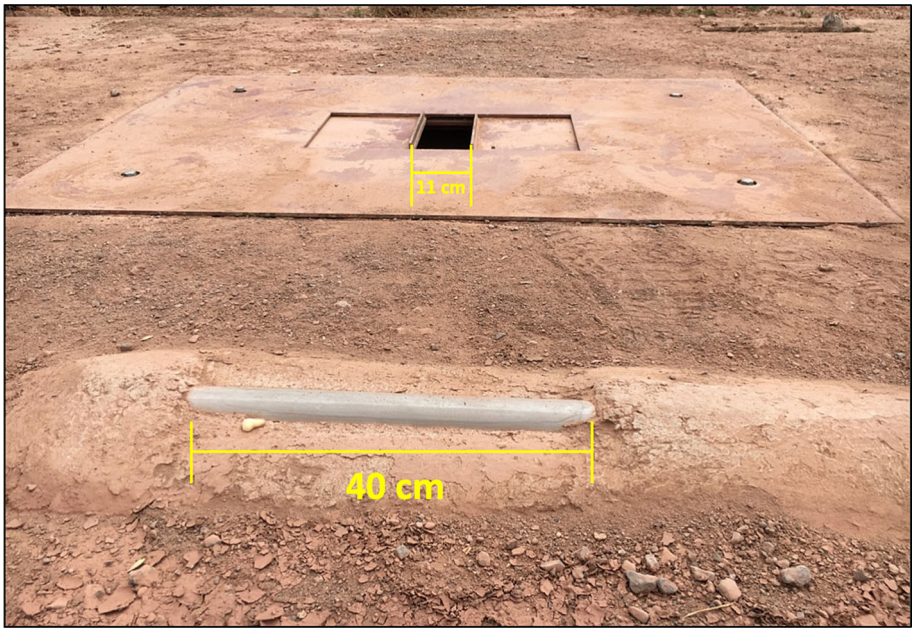


FIGURE 1 Location of the Arroyo de los Pinos within the Rio Grande watershed near Socorro, New Mexico (upper map). Inset (bottom), topography of the Arroyo de los Pinos. Data presented in this manuscript were collected at the basin outlet (green dot).

FIGURE 2 Downstream view of an impact sensor deployment with an attendant bedload slot sampler.



Additional data collected at the site included water depth (monitored by vented pressure transducers), rainfall intensity, and suspended sediment grain size and concentrations obtained via automated pump samplers. Water discharge was estimated using a site-specific rating curve. Average bed shear stress was calculated with the depth-slope product: $\tau = \rho g R s$, where ρ is the density of water, g is the acceleration due to gravity, R is the hydraulic radius, and s is the longitudinal channel slope. This formulation of shear stress does not capture the unsteady nature of water and sediment flux in the Pinos, particularly so for the arrival of flood bores (Thappeta et al., 2023).

These data provide context for the bedload flux and surrogate instrument datasets.

Acoustic impact pipes (often termed Japanese-style impact pipes) were placed in front of the left and right samplers (Figure 2). These instruments were installed and are operated using a methodology developed previously (Mizuyama, Laronne, et al., 2010):

1. A 3-mm-thick, 5-cm-diameter pipe was embedded in the channel, protruding −2.5 cm above the bed with a predetermined exposed length of 40 cm.
2. A microphone sealed inside the pipe recorded the acoustic vibrations resulting from sediment impacting the pipe. These signals were transmitted through a buried cable and processed on-site using a pre-amplifier and digital converter.
3. The digital signal was amplified using a series of 10 amplifications, each twice the gain of the previous (i.e. ×2, ×4, ×8, ×16, ×32, ×64, ×128, ×256, ×512 and ×1024).
4. If the amplified signal exceeded a threshold of 2.5 V, a pulse was counted for that amplification. These pulse counts were the recorded signal output of the pipe microphone system.
5. The microphones were set to sample at a rate of 5 Hz; if the number of counted pulses exceeded this sample rate, the signal was considered saturated. Higher amplifications (×512 and ×1024) saturated often but low amplifications (×2, ×4 and ×8) never saturated.
6. To preserve battery power, the impact sensors were designed to cycle intermittently every 30 s. Data presented in this manuscript were resampled to 1 s, to which a 3-min moving average window was applied. The resulting outputs were reported in pulses s^{-1}.

An acoustic impact plate was also deployed directly in front of the centre sampler. The plate is 25 cm long, 10 cm wide and 5 mm thick. Data from the impact plate were processed identically to those from impact pipes, but the different dimensions of the plate (specifically the thickness) required different data analysis and interpretation.

When evaluating the ability of the impact sensors to predict bedload flux, we chose to utilize a leave-one-out cross-validation strategy (LOOCV; Hastie et al., 2009). This strategy is particularly useful for assessing the extent to which our results generalize to other datasets. Cross-validation is a resampling method that isolates different portions of the entire dataset to test and train a model. Rather than partitioning the dataset into static train-and-test groups, LOOCV isolates a single testing data point (the 'one out'), builds a linear regression with the remaining dataset and predicts the testing point. It iterates through the entire dataset and calculates the statistical performance of each model. Because we evaluate several different model options, we compared the results of LOOCV using the residuals of each model and the R^2 value of the predicted versus observed bedload fluxes for each iteration.

3 | RESULTS

3.1 | Hydrologic and sediment transport data

Construction of the monitoring station was completed in 2018. Since then, 16 flow events have been recorded (Table 1). These events varied in size and hydrograph shape; hydrographs were typically either double or single peaked. The Reid-type slot samplers quickly filled to capacity. On average, only 28% of the flow event duration was monitored using these samplers and generally only the lowest 1/3 of observed flow depths have associated bedload flux data (Figure 3). Quick-rising bore style floods were occasionally observed at the monitoring station. Peak water depth was achieved within minutes, and

TABLE 1 Summary of bedload-transporting events at the Arroyo de los Pinos monitoring station since 2018. Some bedload flux data are missing due to samplers being full from a recent previous event.

Date (YYYY-MM-DD)	Maximum shear stress[a] (N m^{-2})	Peak water discharge[c] (m^3 s^{-1})	Peak channel-average bedload flux (kg s^{-1} m^{-1})
2018-07-16	51.8	10.1	--
2018-07-26	144.0	74.3	11.3[d]
2018-08-09	20.4	2.0	3.8[d]
2018-08-24	34.8	5.0	10.9[d]
2018-09-01	13.9	1.0	1.0
2020–07-23[b]	20.5	2.0	3.8
2020-07-24[b]	33.3	4.6	--
2020-09-01	15.5	1.2	2.0
2021-07-02	35.1	5.0	7.1[d]
2021-07-05	118.0	48.9	11.4[d]
2021-07-06	76.2	20.5	--
2021-07-23	95.2	31.6	6.0[d]
2021-08-12	14.5	1.1	3.2[d]
2021-08-23	21.8	2.2	4.1[d]
2021-08-27	61.2	13.6	8.3[d]
2021-09-28	12.4	0.9	1.4

[a]Calculated using the depth-slope product, detailed in the Section 2.
[b]Right impact pipe data not available.
[c]Calculated using a site-specific rating curve.
[d]Peak bedload flux measured prior to peak water discharge due to limited sampler capacity.

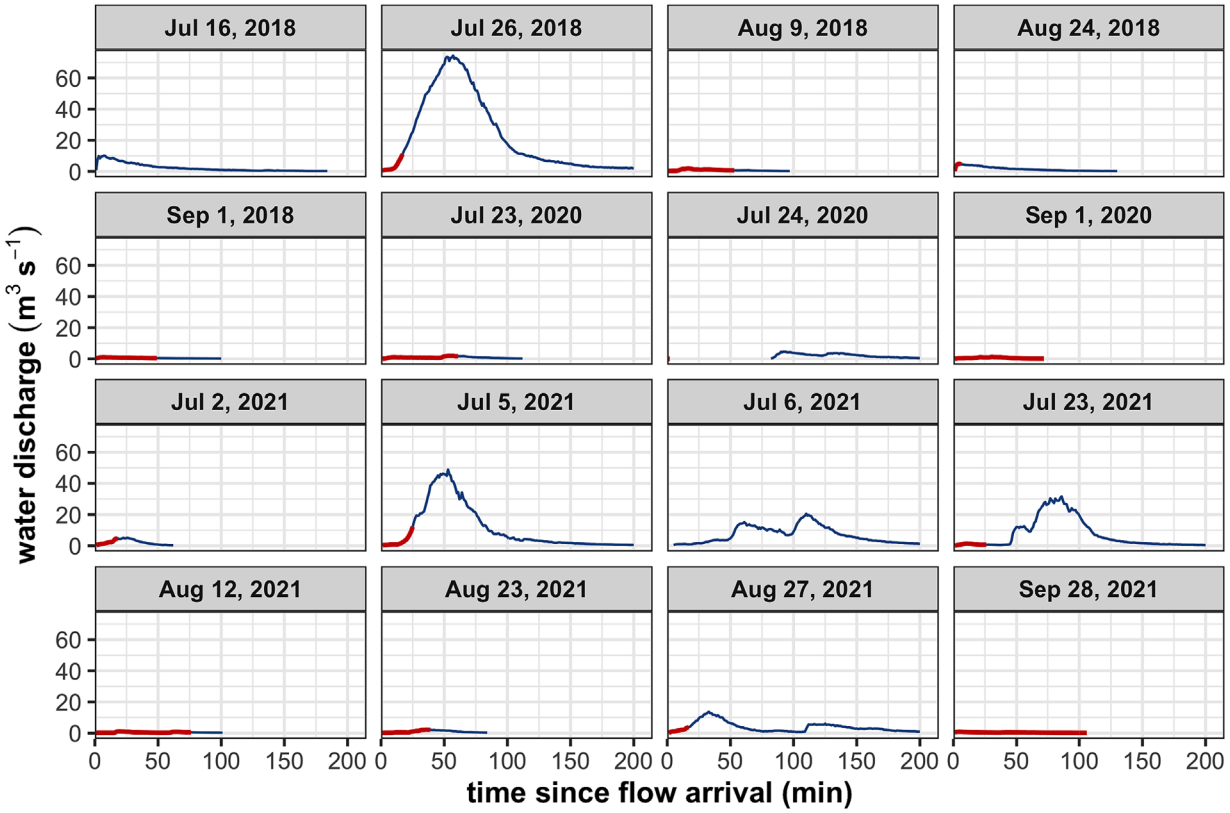


FIGURE 3 Hydrographs of bedload transporting events, 2018–2021. The blue lines are the entire hydrograph; the red portion represents periods when the bedload samplers were active.

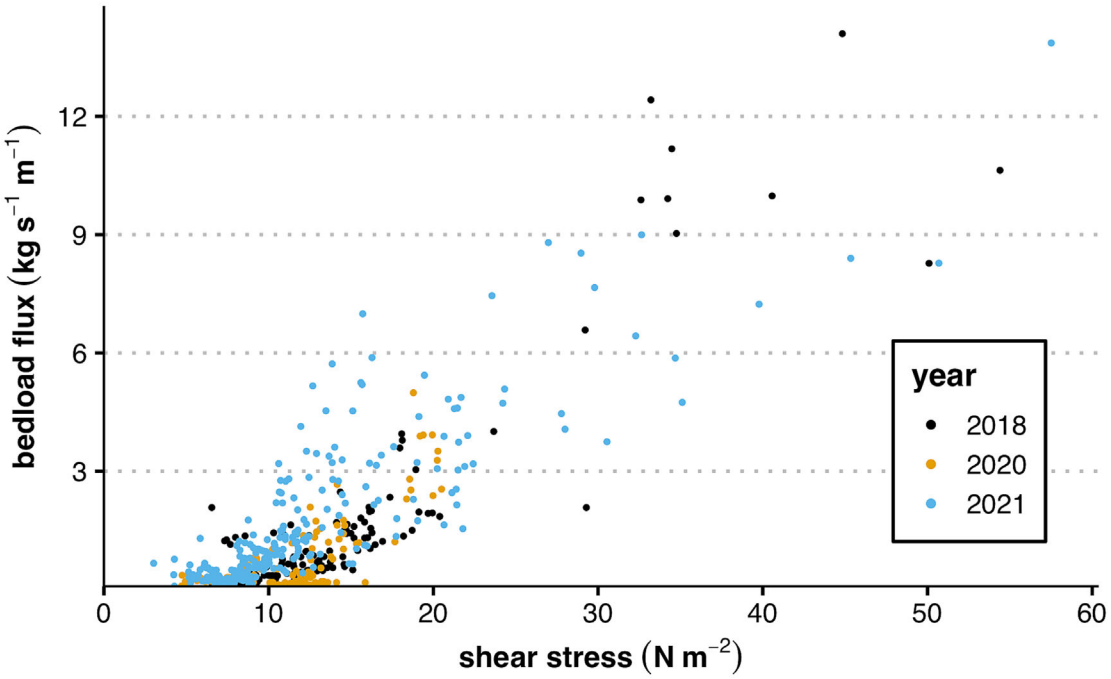


FIGURE 4 Relationship between directly-measured bedload flux and shear stress from the depth-slope product. The shear stresses depicted here represent the lower third of the observed flow strengths due to lack of bedload data for higher water depths.

bedload transport rates were generally higher, producing conditions in which the Reid-type slot samplers reached capacity even faster. Despite the lack of complete event bedload flux datasets, the relationship between channel-averaged bedload flux and shear stress is robust in the range of available data. As shear stress increased, bedload flux increased considerably (Figure 4). Rates of measured bedload transport at the Pinos were high, even for arid environments (Stark et al., 2021). The surrogate impact instruments were deployed with the objective of extending the record of bedload flux beyond the physical measurements.

Bedload grain size distribution changed with increasing flow strength. Steep increases in bedload median grain size, D_{50}, were observed at low shear stresses (Figure 5). The rate of change of bedload grain size promptly decreased as flow strength increased beyond 12 N m^{-2}. We interpret this as the system approaching equal mobility—when tractive forces acting on the bed are sufficiently large to transport bedload in equal proportions to the grain size distribution in the channel bed (Parker & Toro-Escobar, 2002). This is demonstrated by the fact that the D_{50} of bedload material approaches, but never exceeds, that of the bar material (Figure 5). Based on bedload and thalweg bed material grain size data, we interpret this transition to occur between 12 and 20 N m^{-2}. When water depth sufficiently increased to allow the inundation and erosion of bars, bedload grain size should increase slightly and then stabilize at a new equilibrium. Samples from periods of full bar inundation were not available because the slot samplers reached capacity prior to this stage. These samplers have been shown to be less efficient at trapping coarse sediment (Stark et al., 2021) but based on qualitative observations (such as the movement of large boulders between floods), we are confident that the channel bed achieves equal mobility at these low shear stresses. The relative abundance of sand-sized material in the bedload also underwent significant changes at low shear stresses (Figure 5). Fine sands (<0.25 mm) were initially transported as bedload, but as water depth rose, their saltation lengths increased and the transport of this sediment eventually transitioned to suspension (Stark et al., 2021).

3.2 | Impact sensor datasets

The acoustic impact pipe microphones responded instantly to the onset of flowing water and bedload transport (Figure 6). Higher amplifications (e.g. ×512 and ×1024) attained saturation immediately. Lower amplifications, such as ×2, ×4 and ×8, were never saturated but show modest variation in their signal output despite large changes in flow magnitude. Occasionally, the surrogate signal decreased suddenly despite no obvious change in either flow strength or bedload flux (e.g. Figure 6, 23:50–00:00). These changes in signal strength were attributed to partial cover of the sensor. Gravel sheets (Whiting et al., 1988) commonly observed at the Pinos station appear to overrun the instrument and shield it from impacts for a period of time before eventually being transported downstream. These periods of full and partial covering were identified and removed. A method of detecting these periods was implemented using the most sensitive amplifications. Because these amplifications are saturated in nearly all flow conditions during normal instrument operation at the Pinos, tracking periods when they are *not* saturated is an efficient and reproducible manner to identify partial or full covering. At the Pinos, an impact sensor was considered covered if amplification ×256 recorded a pulse rate below 5 pulses s^{-1} while stage was >15 cm. The stage threshold was implemented to avoid removing data before ×256 typically saturated; this was determined based on an analysis of pulse rate vs stage relationships. For the 2021 season, 308 data points were available when all three bedload flux samplers were active, of which 264 values were deemed to be from periods of no covering. In total, 2660 data points were available from all flow events and flow periods (including those when the bedload samplers were full); 1503 of these were deemed to be from periods of no cover.

Correlations between the impact sensor pulse counts and bedload flux were made for those periods of time when the bedload samplers were active and the pipes were not covered by gravel sheets. These correlations form the basis of the instrument calibrations. Correlations between impact sensor pulses from a single pipe compared to bedload flux of the single associated slot sampler show poor predictive ability

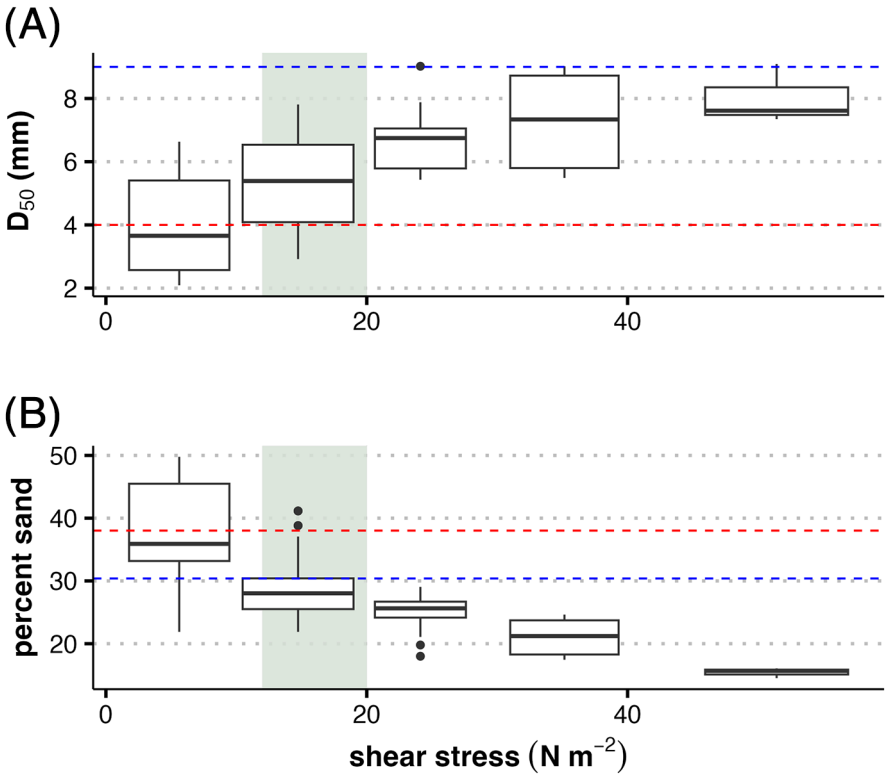


FIGURE 5 Boxplots of bedload D_{50} (a) and percentage sand (b) variation with shear stress. The green shaded area indicates the shear range at which the transition to equal mobility occurs. The dashed red line in each figure represents the D_{50} (a) or percent sand (b) in the channel thalweg bed material, and the blue dashed line is the value in the channel bar sub-surface.

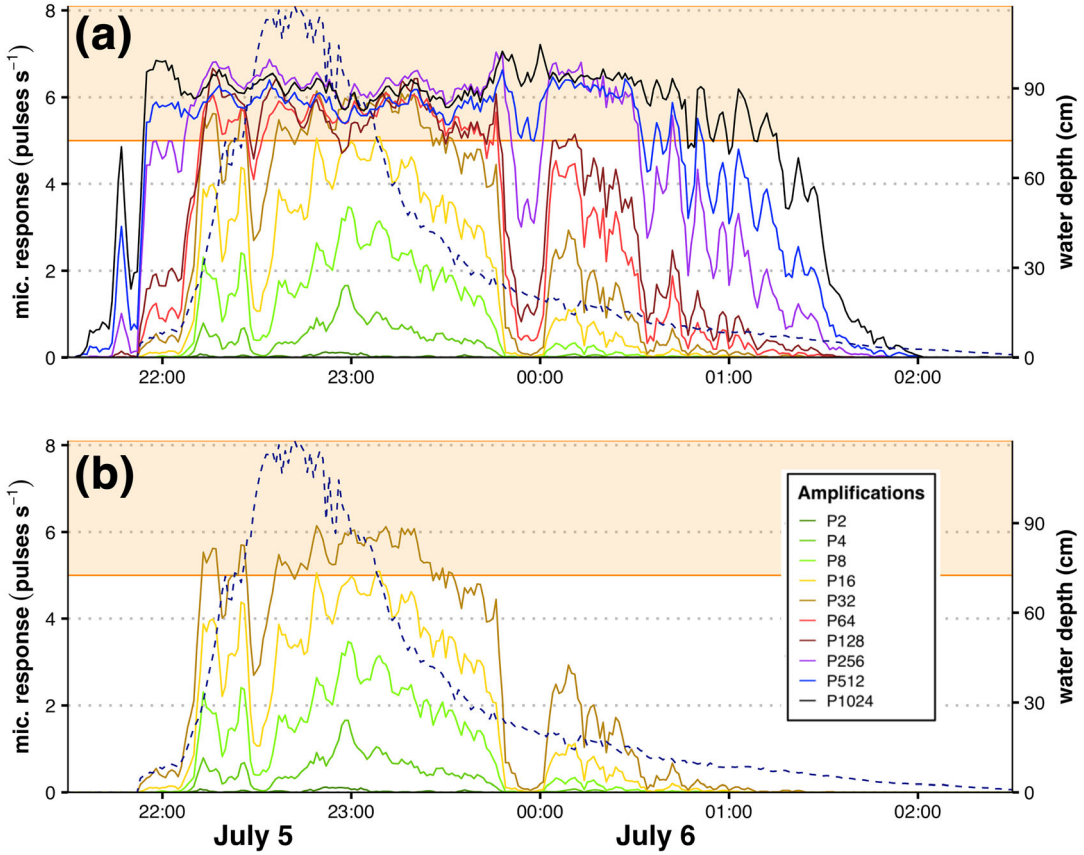


FIGURE 6 An example of an impact sensor time series (left pipe microphone, July 5 and 6, 2021). The solid-coloured lines are the microphone pulse rates for each amplification; the dashed blue line is water depth. Pulse rates in the shaded orange region are considered saturated, above 5 pulses s^{-1}. Panel A includes every amplification; panel B includes amplifications that remain unsaturated for most the event. The pronounced dip between 23:50 and 00:00 is an example of a likely period when the pipe was covered by a sediment sheet.

(LOOCV observed vs. predicted $R^2 = 0.25$). This simple methodology for impact sensor calibration was improved by spatially upscaling the analysis to a channel-averaged approach. Figure 7 depicts the channel-averaged bedload flux from the three slot samplers compared to the average response of the left and right pipe microphones. With this change, predictive ability of a linear model was roughly doubled; the best performing model using the ×32 amplification yielded $R^2 = 0.50$.

The impact plate placed directly in front of the centre sampler performed poorly in comparison to the response of the pipe sensors. Investigation of the acoustic response revealed significant periods when the plate was partially or fully covered. Inspection of the site after flow events reinforced this observation: the plate was often found covered by several centimetres of sediment, while the impact pipes were usually exposed. After removing the data from periods when the plate was covered, the instrument still displayed a poor capacity to estimate bedload flux (maximum $R^2 = 0.31$), possibly due to the thicker metal of the plate.

4 | DISCUSSION AND DATA-INFORMED CALIBRATION

Based on 4 years of data collection and on-site observation, the centre acoustic impact plate was unable to produce satisfactory results for predicting bedload flux in this sand-rich environment. Compared to the pipe instruments, the thicker plate was insufficiently sensitive in detecting the transport of the majority of bedload at the Pinos. Further analysis of the plate-based results is limited to comparisons with other studies of impact plate equipment; the remaining discussion concentrates on improving the quality of the impact pipe calibration.

Bedload flux varied in a fairly predictable fashion with increase in shear stress ($R^2 = 0.76$, Figure 4). Shear stress is a good first approximation for estimating bedload flux (Cohen et al., 2010); indeed, most transport equations use a form of shear stress to predict fluxes (e.g. Gomez & Church, 1989; Reid et al., 1996). Hence, this served as a base case against which to evaluate the calibration of the impact sensors. If a model did not meet the quality of this base case, it was automatically rejected.

4.1 | Investigation of different calibration techniques

Several calibration strategies were considered for the acoustic impact pipes. These included incorporating information related to flow strength, multi-amplification calibrations (where several amplifications were considered jointly in a multiple regression to predict bedload flux), weighted linear regressions, time shifting to better match observed bedload fluxes and transformations of the datasets to improve linearity across the range of bedload fluxes. Table 2 provides a subset of the results from these attempted calibration approaches.

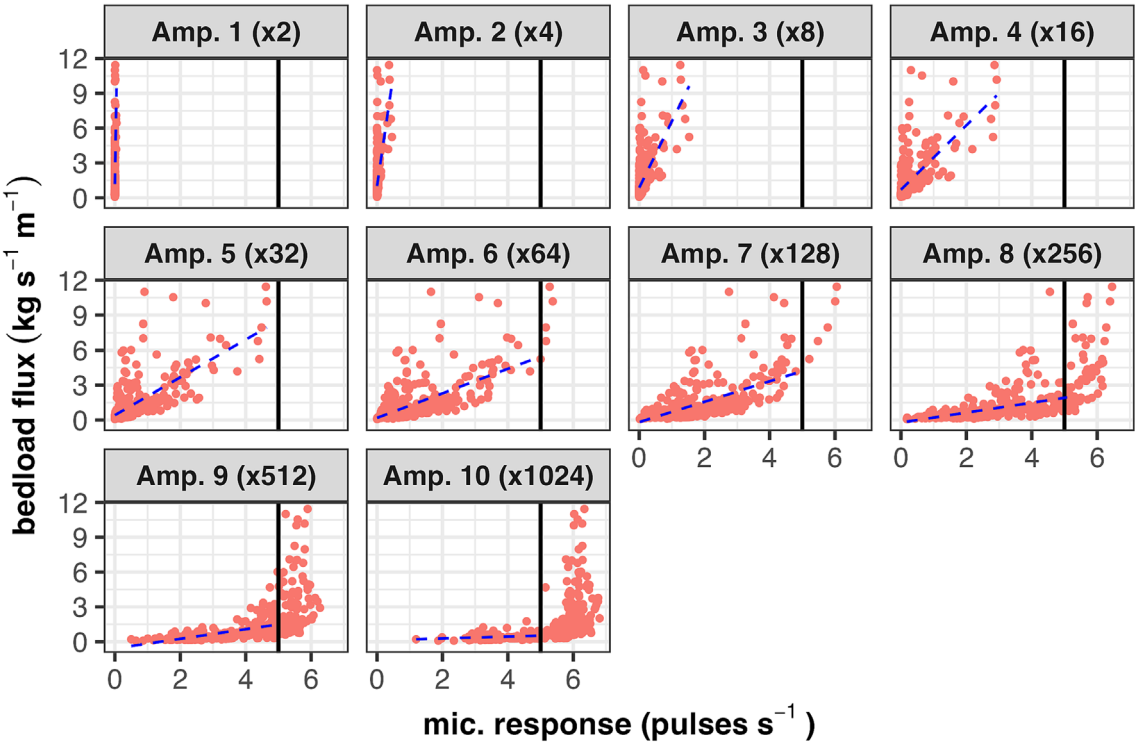


FIGURE 7 Example of simple linear predictions of bedload flux using microphone pulses (dashed blue line). Each facet is a different amplification. The solid black line is the saturation line, five pulses s^{-1}. Data above this line are included for informational purposes but are excluded in the blue dashed lines of the linear models.

TABLE 2 Subset of rejected calibration methods.

Calibration method	Best performing result[a]	Amplification used
Single-amplification model	$R^2 = 0.50$	×32
Log-transformed data	$R^2 = 0.08$	×16
Step-function	$R^2 = 0.52$; flux < 2 kg s^{-1} m^{-1} $R^2 = 0.27$; flux > 2 kg s^{-1} m^{-1}	Multiple
Multiple linear regression models	Adjusted $R^2 = 0.66$	×32 and the D_{10} of bedload
Multiple linear regression + flow strength	Adjusted $R^2 = 0.77$	×16 and shear stress

[a]The R^2 values reported here represent the coefficient of determination between the predicted and observed bedload fluxes using leave-one-out cross-validation (LOOCV).

Ultimately, a simple, single-amplification method was adopted. This method is preferred for several reasons: (1) implementing a data-limited calibration allows other users to utilize calibration coefficients collected at the Pinos for other rivers; (2) reducing the number of predictor variables decreases the risk of overfitting the results with covarying parameters; and (3) using a methodology with a prediction based on single-amplification follows other studies on impact-driven systems (e.g. the studies reviewed by Rickenmann et al., 2014). We chose to exclude the use of flow strength in our calibration (even though it provided better predictability) because bedload fluxes exhibit intrinsic variability in spite of a relatively constant forcing. Bedload is often transported in waves or pulses (Cudden & Hoey, 2003); by detaching our calibrated model from flow strength, we may predict variations of bedload flux that are often independent of flow strength.

When using a single amplification approach, the amplification should remain unsaturated across all observed flow strengths. Amplifications above ×64 were prone to signal saturation in the range of commonly-occurring flow depths, while amplifications ×2 and ×4 displayed very little response to increase in bedload flux. We therefore assessed the quality of calibrations for amplifications ×8, ×16, ×32 and ×64.

4.2 | A grain-size-limited approach

Based on published experience, grain size is a critical parameter for calibrating impact sensors (Wyss et al., 2016a). Laboratory tests have confirmed these sensors are unable to record bedload flux of the finest grain sizes in transport (Mizuyama, Oda, et al., 2010; Wyss et al., 2016b). Investigating when different grain size classes are mobilized determines which amplification is most appropriate for calibration in a sandy, fine-gravel rich channel such as the Pinos. To ensure that our chosen amplifications for calibration were appropriate for the evaluated grain sizes, we calculated the average bedload D_{50} for periods when each amplification began recording pulses. The bulk of

the amplifications began recording pulses when the median grain size in transport was 3–4 mm. Amplifications ×1024 and ×512 began recording immediately with the onset of flow, even when the preponderance of bedload was sand. Amplification ×2 never consistently recorded pulses during the calibration period, indicating grain size must have increased beyond 8 mm before it was activated, while amplification ×4 activated when the median grain size was 7 mm. These values verify that the targeted amplifications (×8–×64) are capable of monitoring most of the grain sizes in transport at the Arroyo de los Pinos, and that we may interpret our results when using these amplifications as the transport of bedload larger than 3–4 mm.

Utilizing the grain size information from sampled bedload, we removed the bedload flux related to the smallest particles and improved the quality of the prediction model. Significant improvement was obtained in the predictability of bedload flux when considering the flux of grains larger than a given threshold (Table 3; Figure 8). These grain-size-limited linear regressions are interpreted as the predictors of bedload flux of a given grain fraction, rather than of the entire distribution of sediment in transport. When the bedload grain size distribution was known (i.e. during the calibration period), these fractional transport rates can be converted back to total bedload flux. The quality of the models peaked when limiting bedload flux to the contribution from 12.5 mm grains and larger. However, choosing the best performing models may not be the best option for estimating bedload flux. In the example of the 12.5 mm limited calibration, only 20% of the mobile grains were included this fraction during the calibration period. Other models using a larger portion of the grain size distribution performed nearly as well. For comparison, on average, 46% of the bedload was coarser-grained than 4.75 mm during calibration. In addition to capturing most of the bedload flux, the possibility of capturing the transport of grains that are ~5 mm in size is important; to date, no other bedload surrogate system has been able to consistently accomplish this.

Given our quality threshold established using the bedload flux-shear stress relationship, none of the ×8 and ×64 amplification models were considered acceptable. Models using the ×16 and ×32 amplifications performed adequately compared to the base case, while also providing a method to estimate bedload flux that is fully independent of a depth-based flow strength parameter. This is crucial for flash flood driven systems such as the Arroyo de los Pinos, where bore-style events are known to be unsteady and cause departures from predictable trends of bedload flux versus shear stress (Halfi et al., 2018; Yager et al., 2018).

Overall, the ×16 and ×32 amplification models perform fairly well; residuals for ×16 and ×32 models have few outliers outside of

TABLE 3 Leave-one-out cross-validation (LOOCV) linear regression coefficients of determination (R^2) derived from isolating the portion of bedload flux greater than a specific grain size and regressing against pulse rates from various amplifications (amp.). Only data from 2021 events were used to generate these linear regressions. Darker shading indicates higher R^2.

Amp.	Total flux	>2 mm	>4.75 mm	>6.3 mm	>8 mm	>12.5 mm	>19 mm	>25 mm
×8	0.42	0.56	0.59	0.61	0.64	0.66	0.71	0.72
×16	0.47	0.62	0.65	0.68	0.71	0.73	0.76	0.77
×32	0.50	0.64	0.67	0.70	0.73	0.75	0.78	0.78
×64	0.41	0.52	0.55	0.58	0.61	0.64	0.66	0.66

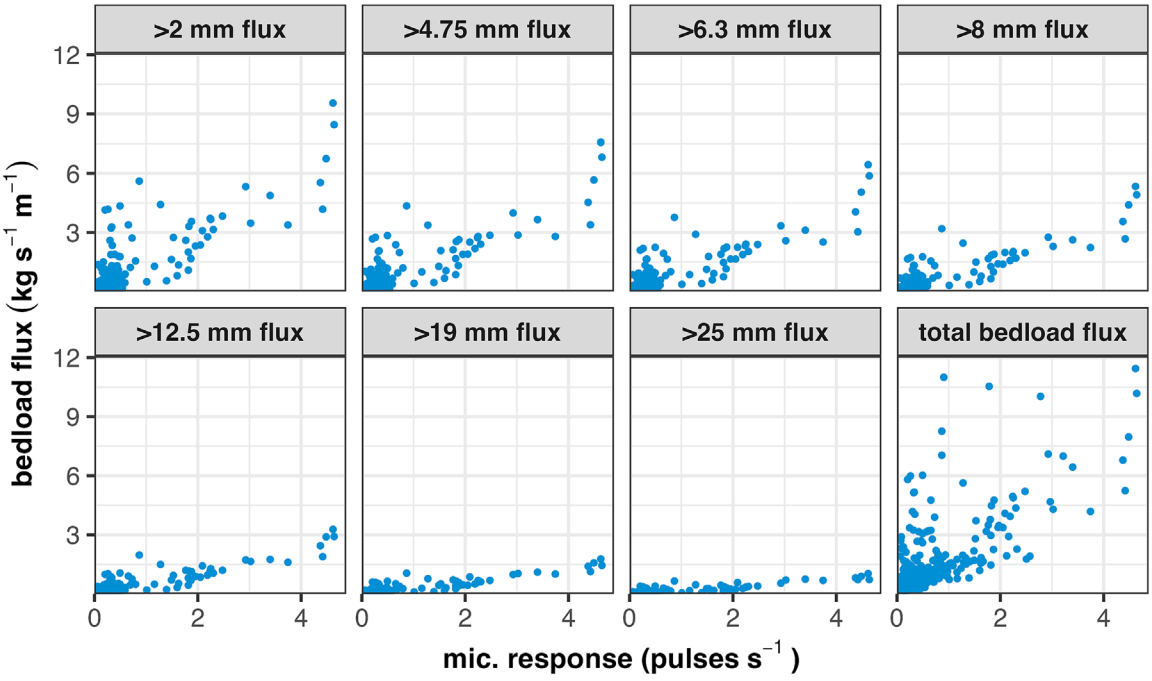


FIGURE 8 Comparison of different grain size-limited bedload fluxes vs. number of pulses (×32 amplification). Limiting bedload flux to various fractions of gravel considerably improves the predictive ability of these amplifications.

the interquartile range (Figure 9). The ×32 models performed slightly better than the ×16 models but occasionally saturated at the highest flow strengthens outside the calibration period. The vast majority of the available data derive from periods of commonly occurring, relatively low bedload fluxes. However, this uneven distribution of data did not affect the overall calibration; the LOOCV approach reduces bias issues in the data distributions and the size of our calibration dataset. Where larger bedload fluxes were monitored, the calibrated models generally all underpredicted bedload flux, regardless of grain size limitation (Figure 9; more positive outliers). The residuals (and root mean square errors) are smaller for models that subset the data to increasingly more limited (coarser) fractions of the bedload flux, in part because the flux values are smaller.

4.3 | Comparison to other impact sensor deployments

Comparing studies using impact sensors is complex, but consideration of differences in equipment, site conditions and data conditioning approaches is necessary to advance this technique beyond site-specific calibrations. Halfi et al. (2020) predicted bedload flux in an ephemeral channel in Israel using a multiple linear regression model. This model used the highest amplification output from an impact plate (equivalent to ×1024 in this study) and the D_{50} measured from the bedload collected from bedload slot samplers. Their study channel is significantly coarser-grained than the Pinos, with sand-sized sediment having been mostly replaced with silt and clays in recent years (Barzilai et al., 2013). This results in a coarser bedload fraction with mostly gravels in transport. Using these two predictors, they achieved a statistically significant regression ($R^2_{adj} = 0.86$) by excluding data collected during the steep hydrograph rises.

Of the available datasets comparing bedload flux to impact acoustic sensors, Halfi et al.'s (2020) study is perhaps the most similar in environmental setting, bedload sampling and observed bedload fluxes. A crucial difference between the sites is grain size and the use of an impact plate as the primary impact sensor system. Using their model (equation 7 in Halfi et al., 2020) and our limited impact plate data, we estimated the bedload flux for the Pinos. The results produced estimates of bedload flux that were less than zero. This is not altogether surprising, given the site-specific calibration of four tuning parameters and suggests that any impact-sensor prediction equation cannot be easily transferred to a new site unless the conditions are near-identical. Their study highlights the likely important differences

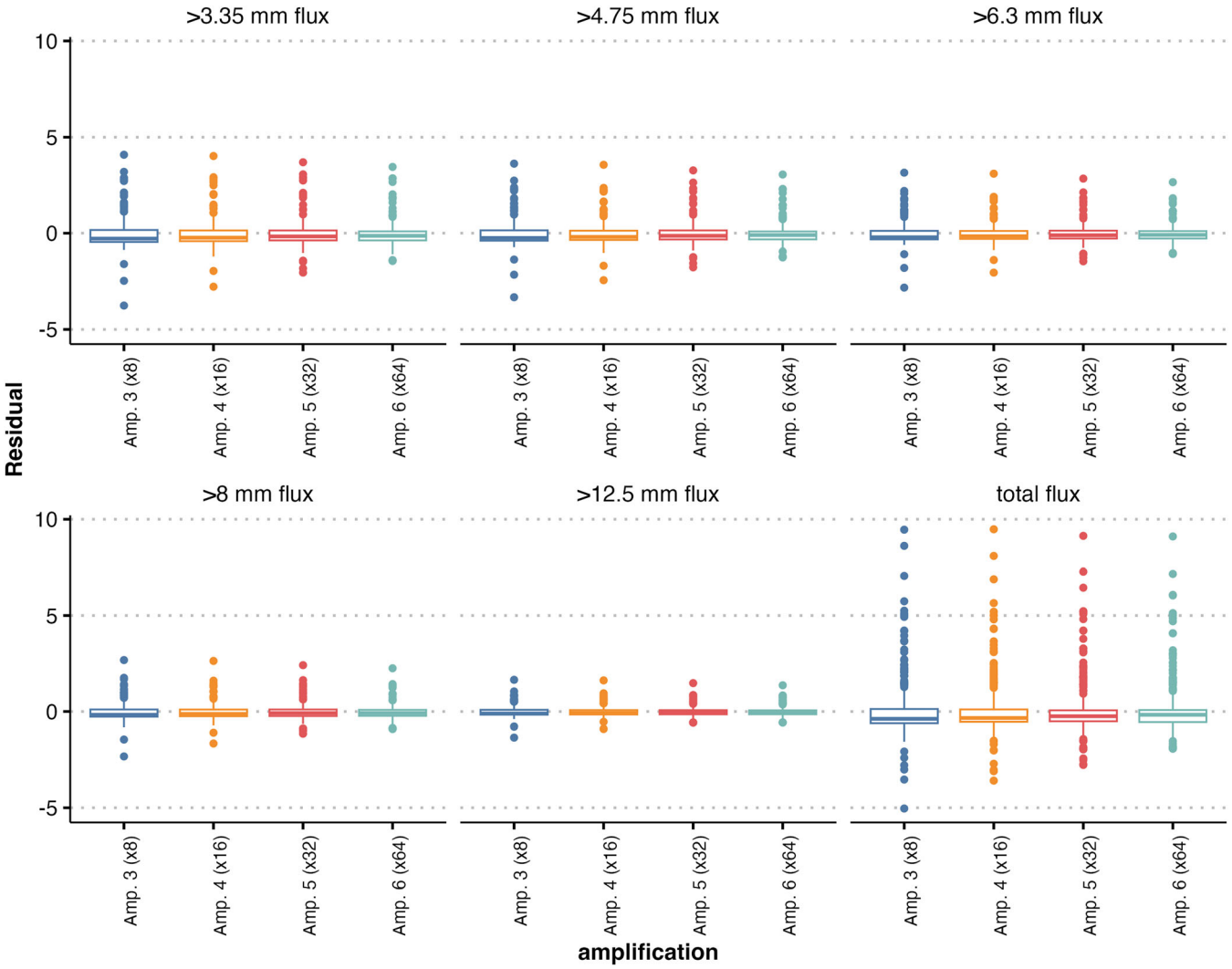


FIGURE 9 Residual (observed bedload flux—model-predicted bedload flux) for a set of selected models. Each model was constructed using leave-one-out cross-validation (LOOCV); the residuals were calculated for each left-out data point.

between hydrograph rise and recession that may exist in a flash-flood driven channel. Although we were unable to observe statistical differences in bedload flux data collected during the hydrograph rise and recession of the calibration dataset, measurable differences in pulse counts between these two segments of the hydrograph can be found in some of the Pinos impact sensor data obtained from larger flows.

Mao et al. (2016) completed a rigorous study of Japanese-style impact pipes using field and flume-based datasets. Their study used very similar equipment and deployment strategies in a starkly different environment than the Pinos. Their research focused on coarser-grained channels with orders of magnitude lower transport rates. Mao et al. (2016) used an approach that relied exclusively on their impact sensor data to predict both grain size and transport across three orders of magnitude. This range of transport rates required elaborate power law regressions using a combination of different amplifications, achieving statistically significant regressions ($R^2 = 0.73$).

Because of deployment similarities, we were able to implement their bedload flux equation (equation 3 in Mao et al., 2016) to predict bedload flux at the Pinos. It underpredicted the observed transport rates by many orders of magnitude; average predicted bedload flux was 0.0003 kg s^{-1} m^{-1} compared to 1.4 kg s^{-1} m^{-1} measured at the Pinos site. Mao et al. (2016) also provided a method to estimate grain size using the impact sensor data (equation 1 in Mao et al., 2016). This had limited success because of the limited impacts counted at the lowest amplifications; when data were available, average estimates of D_{50} in the Pinos using Mao's equation was 17 mm—much higher than the actual average at those times (5.2 mm). Finally, we attempted to estimate bedload flux in the Estero Morales (the channel evaluated in the field component of Mao et al., 2016). Although we did not have access to attendant bedload flux data, our equations appeared to perform poorly for these data. Average bedload fluxes were at least one order of magnitude larger than observations reported in Mao et al., 2016. Comparisons like these highlight why universal calibrations for acoustic impact sensors have hitherto proved to be elusive. Individual site conditions or equipment setups may be similar, but calibrations remain distinctly site-specific. Several factors beyond grain size differences contribute to the site-specific nature of these calibrations, including bed slope, differences in flow strength and flow velocity (Gray et al., 2010; Nicollier et al., 2021). Data from a range of river systems are required to determine the fundamental controls of impact sensor calibration.

4.4 | Overall quality of calibration

In the original design of the Pinos monitoring station, attendant pairs of samplers and sensors were intended to be independently calibrated. This calibration proved to be inferior to the upscaled version, likely because of the phenomenon of apparent pulses (Nicollier et al., 2021), which occur when an impact sensor records a pulse without receiving a direct impact by transported sediment. When the sediment station was constructed in 2018, it was assumed that the acoustic sensors were sufficiently insulated to record only direct impacts of sediment. These sensors often record indirect impacts onto constructed materials (Nicollier et al., 2021), such as the concrete sill at the Pinos. To determine the degree to which these apparent pulses are measurable, we performed several tests simulating transport on different areas of the managed cross-section. Some impacts from larger grains were recorded across the entire channel, confirming that at least some of the recorded pulses are apparent, rather than only those due to direct impacts. The rate of these apparent pulses is unknown and their effect on different amplifications complicates the matter further—higher amplifications are prone to record more apparent pulses. The result is that individual impact sensors cannot be expected to match well with any one bedload sampler, but instead, a channel-average approach is required. This explains why our calibration efforts improved when using a spatially-averaged calibration.

Sensor cover also impaired the ability of these sensors to predict bedload flux at the Pinos. Cover issues have been observed in other studies of acoustic impact sensors, but not nearly as frequently as was found in this study. Nearly 45% of the entire impact sensor record was rejected because one or both pipes were deemed to be partially or fully covered. Viewing the sensors independently, only 19% of the left sensor record and 28% of the right sensor record were deemed to be from covered periods. The two sensors were rarely covered at the same time, suggesting migration of the thalweg or zone of most active transport. Covered periods also varied widely between events, ranging from 8% to as much as 88% of one flow event. The interaction of pipe cover issues and the apparent pulse phenomena is poorly understood in our dataset. If a pipe is covered but the other half of the concrete sill is not, apparent pulses could still be detected by the covered sensor, masking the low pulse rate as an indicator of burial. This is expected to be most pronounced for higher amplifications, providing additional motivation to use the lowest possible amplification for cover detection. Instrument cover is a clear disadvantage of surrogate instruments being directly deployed on the channel bed. Recent advances using other instruments, such as seismometers, to monitor bedload transport indirectly, are also in development at the Pinos sediment monitoring station. These instruments have the distinct advantage of being placed outside of the stream, allowing continuous monitoring of bedload transport without the effect of instrument cover (Bakker et al., 2020; McLaughlin et al., 2021; Roth et al., 2014).

The significant fraction of sand and fine gravel <8 mm in the Pinos bed material means that acoustic impact sensors are inevitably unable to capture the entire bedload transport rate. While our preferred calibration method predicts bedload flux only for larger grain sizes, the contribution from this size class will increase as flow strength increases beyond our calibration data. Sand will continue to move into suspension while larger grains will be incorporated as more bar material is activated (Figure 5). Choosing a model that only considers sediment coarser than 4.75 mm results in predicting >70% of the total bedload mass for the largest fluxes available in the calibration period. Using information related to grain size, we may scale the grain-limited bedload flux back to total bedload flux. Scaling to total flux beyond the calibration periods adds uncertainty, given the lack of bedload grain size distribution data at larger flow strengths. However, as flow strength increases and larger grains are mobilized and smaller particles transition to transport via suspension, the scaling adjustment for total flux estimation is expected to decrease.

Predicting the transport rates of the largest grains is also most relevant when considering channel-forming discharges. Only when channel bars are fully activated and bedload grain sizes are large do significant changes occur within the channel. Impact sensors provide additional information related to characteristics of bedload material,

such as grain size, though site-specific calibrations appear to be required, as demonstrated by our effort to use an equation from Mao et al. (2016) to estimate D_{50}. In light of this, calibration of these instruments towards movement of the largest grains is most useful in understanding flows that change channel morphology, but less so for the commonly occurring smaller flows. Users of these instruments should consider their need and deploy these instruments accordingly.

Considering the challenges identified with implementing acoustic impact sensors in a sand-rich environment, we consider this calibration effort an important step towards consistent bedload transport measurements using surrogate instruments. With a single impact sensor amplification, we were able to match the prediction ability of shear stress in estimating bedload flux (Figure 10) and have produced results comparable to others attempting similar calibrations (Coviello et al., 2022; Dell'Agnese et al., 2014; Mao et al., 2016; Mizuyama, Laronne, et al., 2010). For a given bedload flux and amplification (e.g. Figure 10), as much as 85% of the data were within a factor of five. This compares favourably with others who have evaluated these types of acoustic instruments, particularly when accounting for the significant amount of sand and fine gravel in the Pinos system. Choosing which model methodology to use was non-trivial, and ultimately, after having explored complex options, we opted for simplicity over complexity.

4.5 | Recommendations for future calibration attempts

Acoustic impact sensor systems provide a cost-effective alternative to direct sampling of bedload transport. Implementation of these systems is somewhat straightforward in pre-existing-built environments such as bridge piers and dams. But before widespread adoption of these impact sensors to monitor bedload, we recommend continued fundamental research. Future improvements to the Pinos system of equipment and calibration techniques are being undertaken, including replacing the centre impact plate with a pipe hydrophone, obtaining physical measurements of bedload flux and grain size during higher water discharges, and evaluating the effect that water velocity has on instrument calibration. We also offer the following considerations and suggestions for others attempting to use similar instruments or deployment techniques:

1. A wide range of flow conditions need to be considered during calibration. This is more important than a high number of individual data points. Our dataset was fairly large (264 data points used in calibration) but lacked crucial information at higher flow rates.
2. Some information about grain size is necessary. Ideally, the size of grains in transport should be measured but grain size information from the channel bed can be sufficient in interpreting the signals from these acoustic systems.
3. When possible, collecting acoustic datasets from different locations across a channel cross-section is important. We observed some differences between our acoustic impact sensors spread across the channel. Those differences helped in our interpretation and analysis during calibration.
4. Site-specific issues will arise when deploying these sensors in-channel. At the Pinos, sensor covering and apparent pulses were our biggest issues, but others may occur. It is important to adapt your system or analyses to these issues.

In the future, more sophisticated calibration and analysis techniques could be employed to better predict bedload flux using acoustic impact sensors. These could include machine learning techniques (where many data streams are used to build predictive models), reviewing and evolving the pulse counting method (such as the methods highlighted by Choi et al., 2020), or utilizing the full acoustic waveform. The techniques and calibration described here were selected because of the ease-of-use in other channels or systems. With advancement in computing and the continued data collected from an ever-growing set of channels and flow regimes, these more sophisticated analyses could be used in channels such as the Pinos.

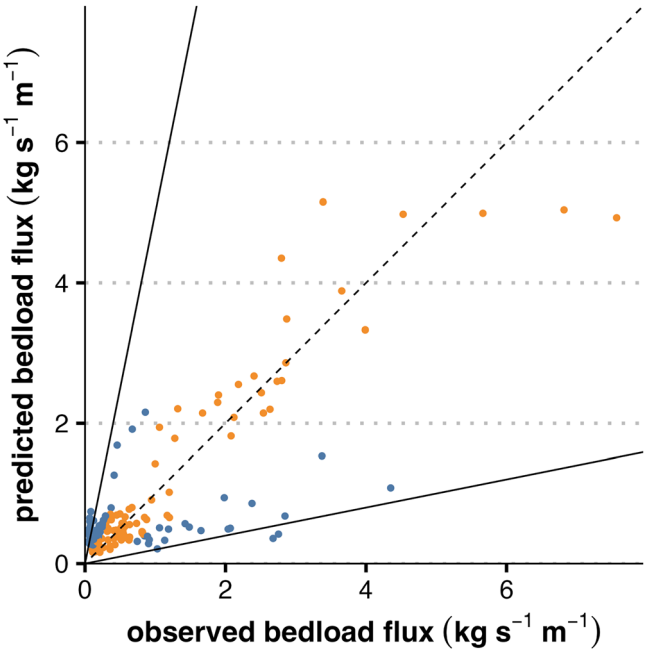


FIGURE 10 Comparison of predicted bedload flux and observed bedload flux for the 4.75 mm, 32× amplification model. Orange data are within a factor of two (blue more than a factor of two) while the black solid lines delimit data within a factor of five.

5 | CONCLUSIONS

We have deployed an acoustic impact sensor array, capable of estimating bedload flux with a similar predictive power to shear-stress-based predictions, but independent of hydraulic measurements and capable of observing hysteresis and other complex bedload dynamics. When the sensors are active and functioning properly, the gravel fraction of bedload flux may be estimated with a high degree of certainty. Several calibration strategies were evaluated before settling on a method that prioritizes simplicity and reproducibility while maintaining effectiveness. We chose a single-amplification approach and calibrated it against the range of grain sizes that it could detect, rather than estimating the total bedload flux. Amplifications with gains of ×16 and ×32 (the two most sensitive channels that do not have their signals consistently saturated during high discharges) provided the best results.

Calibrating a pipe microphone system in an environment with a considerable sand and fine gravel fraction presented many unique challenges. Estimating the total bedload flux was only in part successful, because the instruments could not consistently detect the movement of sediment smaller than 4 mm. Sediment finer than 4 mm regularly constitutes <30% of the transported bedload material during larger events. Issues of instrument cover were common: as much as 47% of the entire impact sensor record derived from periods when the instruments were fully or partially covered by immobile sediment sheets. Apparent pulse detection, resulting from indirect impacts on surrounding concrete surfaces, was observed but is difficult to quantify or to filter. Despite these challenges, we consider this calibration attempt an important step towards surrogate-based bedload transport monitoring. Any method that provides an estimation of bedload transport independent of shear stress is valuable, especially one that is automated and unlimited in capacity.

Future research should focus on extending the calibration record to larger, less common flows, while also considering the effects of flow velocity. Our vision for the Arroyo de los Pinos monitoring station is to rigorously test different indirect sediment monitoring techniques—especially of bedload—against the highest quality direct, physical measurements. By selecting the Pinos as our observatory, we aim to advance sediment monitoring in rivers with a wide range of transported grain size.

AUTHOR CONTRIBUTIONS

Kyle Stark: Conceptualization; methodology (including methodological development); investigation (e.g. data collection); resources (provision of data etc.); software (its provision and development); writing—initial draft; writing—reviewing and editing. **Daniel Cadol**: Conceptualization; funding acquisition; investigation (e.g. data collection); resources (provision of data etc.); supervision; writing—reviewing and editing. **David Varyu**: Conceptualization; funding acquisition; supervision; writing—reviewing and editing. **Eran Halfi**: Methodology (including methodological development); resources (provision of data etc.); writing—reviewing and editing. **Jonathan B. Laronne**: Conceptualization; funding acquisition; methodology (including methodological development); investigation (e.g. data collection); resources (provision of data etc.); supervision; writing—reviewing and editing.

ACKNOWLEDGEMENTS

Several funding sources have made this research possible. Principally, the U.S. Bureau of Reclamation (project 9781) and the National Science Foundation (award EAR-1852794) provided student research funding to Stark. Additional support was provided through grant-in-aid programmes by the New Mexico Geological Society and the New Mexico Water Resources Research Institute.

The authors thank tireless efforts of the Pinos team: Sandy, Zach, Mitchell, Loc and Rebecca. Their contributions to this study made this manuscript possible. The authors also thank Yaniv Munwes for his efforts in designing the systems and troubleshooting through several years of project iteration. His patience and availability allowed us to troubleshoot issues in a timely manner.

DATA AVAILABILITY STATEMENT

Bedload transport data are archived alongside other bedload transport datasets in a publicly-accessible database: Stark, et al., 2022. Other Data related to the Arroyo de los Pinos project are publicly available on the U.S. Reclamation Information Sharing Environment (RISE): https://data.usbr.gov/. The RISE environment provides an interactive platform for accessing data collected during Reclamation sponsored projects. Pinos data can be accessed by searching 'Arroyo de los Pinos'.

ORCID

Kyle Stark https://orcid.org/0000-0001-9768-4880
Daniel Cadol https://orcid.org/0000-0002-4408-8548
Eran Halfi https://orcid.org/0000-0002-4045-2824
Jonathan B. Laronne https://orcid.org/0000-0002-2889-9316

REFERENCES

Anderson, M.G. (1976) An inexpensive circuit design for the acoustic detection of oscillations in bedload transport in natural streams. *Earth Surface Processes*, 1(3), 213–217. Available from: https://doi.org/10.1002/esp.3290010303

Bakker, M., Gimbert, F., Geay, T., Misset, C., Zanker, S. & Recking, A. (2020) Field application and validation of a seismic bedload transport model. *Journal of Geophysical Research - Earth Surface*, 125(5), e2019JF005416. Available from: https://doi.org/10.1029/2019JF005416

Banziger, R. & Burch, H. (1990) Acoustic sensors (hydrophones) as indicators for bed load transport in a mountain torrent. *Int'l Assoc. Hydrol. Sci. Hydrology in Mountainous Regions I. Hydrological Measurements*, 193, 207–214.

Barzilai, R., Laronne, J.B. & Reid, I. (2013) Effect of changes in fine-grained matrix on bedload sediment transport in a gravel-bed river. *Earth Surface Processes and Landforms*, 38(5), 441–448. Available from: https://doi.org/10.1002/esp.3288

Choi, J.-H., Jun, K.-W. & Jang, C.-D. (2020) Bed-load collision sound filtering through separation of pipe hydrophone frequency bands. *Water*, 12(7), 1875. Available from: https://doi.org/10.3390/w12071875

Cohen, H., Laronne, J.B. & Reid, I. (2010) Simplicity and complexity of bed load response during flash floods in a gravel bed ephemeral river: a 10 year field study. *Water Resources Research*, 46(11), W11542 Available from: https://doi.org/10.1029/2010WR009160

Coviello, V., Vignoli, G., Simoni, S., Bertoldi, W., Engel, M., Buter, A., et al. (2022) Bedload fluxes in a glacier-fed river at multiple temporal scales. *Water Resources Research*, 58(10), e2021WR031873. Available from: https://doi.org/10.1029/2021WR031873

Cudden, J.R. & Hoey, T.B. (2003) The causes of bedload pulses in a gravel channel: the implications of bedload grain-size distributions. *Earth Surface Processes and Landforms*, 28(13), 1411–1428. Available from: https://doi.org/10.1002/ESP.521

Dell'Agnese, A., Mao, L. & Comiti, F. (2014) Calibration of an acoustic pipe sensor through bedload traps in a glacierized basin. *Catena*, 121, 222–231. Available from: https://doi.org/10.1016/j.catena.2014.05.021

Geay, T., Belleudy, P., Gervaise, C., Habersack, H., Aigner, J., Kreisler, A., et al. (2017) Passive acoustic monitoring of bed load discharge in a large gravel bed river. *Journal of Geophysical Research - Earth Surface*, 122(2), 528–545. Available from: https://doi.org/10.1002/2016JF004112

Geay, T., Belleudy, P., Laronne, J.B., Camenen, B. & Gervaise, C. (2017) Spectral variations of underwater river sounds. *Earth Surface Processes and Landforms*, 42(14), 2447–2456. Available from: https://doi.org/10.1002/ESP.4208

Gomez, B. & Church, M. (1989) An assessment of bed load sediment transport formulae for gravel bed rivers. *Water Resources Research*, 25(6), 1161–1186. Available from: https://doi.org/10.1029/WR025i006p01161

Gray, J. R., Laronne, J. B., & Marr, J. (2010). *Bedload-surrogate monitoring technologies*. S. Geological Survey. Reston, VA. 2010–5091, 37 p. Retrieved from https://pubs.usgs.gov/sir/2010/5091/

Habersack, H.M., Nachtnebel, H.P. & Laronne, J.B. (2001) The continuous measurement of bedload discharge in a large alpine gravel bed river. *Journal of Hydraulic Research*, 39(2), 125–133. Available from: https://doi.org/10.1080/00221680109499813

Halfi, E., Deshpande, V., Johnson, J.P.L., Katoshevski, D., Reid, I., Storz-Peretz, Y., et al. (2018) Characterization of bedload discharge in bores and very unsteady flows in an ephemeral channel. *E3S Web of Conferences*, 40, 02036. Available from: https://doi.org/10.1051/e3sconf/20184002036

Halfi, E., Paz, D., Stark, K., Yogev, U., Reid, I., Dorman, M., et al. (2020) Novel mass-aggregation-based calibration of an acoustic method of monitoring bedload flux by infrequent desert flash floods. *Earth Surface Processes and Landforms*, 45(14), 3510–3524. Available from: https://doi.org/10.1002/esp.4988

Hastie, T., Tibshirani, R., Friedman, J.H. & Friedman, J.H. (2009) *The elements of statistical learning: data mining, inference, and prediction*, Vol. 2. New York: springer, pp. 1–758.

Hilldale, R.C., Carpenter, W.O., Goodwiller, B., Chambers, J.P., Randle, T.J., Wre, D., et al. (2014) Installation of impact plates to continuously measure bed load: Elwha River, Washington, USA. *Journal of Hydraulic Engineering*, 141(3), 06014023. Available from: https://doi.org/10.1061/(ASCE)HY.1943-7900.0000975

Mao, L., Carrillo, R., Escauriaza, C. & Iroume, A. (2016) Flume and field-based calibration of surrogate sensors for monitoring bedload transport. *Geomorphology*, 253, 10–21. Available from: https://doi.org/10.1016/J.GEOMORPH.2015.10.002

Marineau, M.D., Wright, S.A. & Gaeuman, D. (2016) Calibration of sediment-generated noise measured using hydrophones to bedload transport in the Trinity River, California, USA. *Proceeding of River Flow*, 1, 1519–1526.

McLaughlin, J.M., Bilek, S., Cadol, D., Laronne, J., Stark, K. & Glasgo, S. (2021) Seismic monitoring of bedload in monsoon floods in gravel bed arroyos in central New Mexico. In: *AGU fall meeting*, Vol. 2021. American Geophysical Union, New Orleans: Retrieved from https://ui.adsabs.harvard.edu/abs/2021AGUFM.S55A0127M/abstract

Mizuyama, T., Laronne, J.B., Nonaka, M., Sawada, T., Satofuka, Y., Matsuoka, M., et al. (2010) Calibration of a passive acoustic bedload monitoring system in Japanese mountain rivers. *US Geological Survey Scientific Investigations Report*, 5091, 296–318. Retrieved from https://pubs.usgs.gov/sir/2010/5091/

Mizuyama, T., Oda, A., Laronne, J.B., Nonaka, M. & Matsuoka, M. (2010) Laboratory tests of a Japanese pipe geophone for continuous acoustic monitoring of coarse bedload. *US Geological Survey Scientific Investigations Report*, 5091, 319–335. Retrieved from https://pubs.usgs.gov/sir/2010/5091/

Nicollier, T., Rickenmann, D., Boss, S., Travaglini, E. & Hartlieb, A. (2020) Calibration of the Swiss plate geophone system at the Zinal field site with direct bedload samples and results from controlled flume experiments. In: *River flow 2020—proceedings of the 10th conference on fluvial hydraulics*, pp. 901–909 https://doi.org/10.1201/B22619-127/. CRC Press, Boca Raton, FL.

Nicollier, T., Rickenmann, D. & Hartlieb, A. (2021) Field and flume measurements with the impact plate: effect of bedload grain-size distribution on signal response. *Earth Surface Processes and Landforms*, 46(8), 1504–1520. Available from: https://doi.org/10.1002/ESP.5117

Parker, G. & Toro-Escobar, C.M. (2002) Equal mobility of gravel in streams: the remains of the day. *Water Resources Research*, 38(11), 461–468. Available from: https://doi.org/10.1029/2001WR000669

Poreh, M., Sagiv, A. & Seginer, I. (1970) Sediment sampling efficiency of slots. *Journal of the Hydraulics Division*, 96(10), 2065–2078. Available from: https://doi.org/10.1061/JYCEAJ.0002729

Recking, A. (2010) A comparison between flume and field bed load transport data and consequences for surface-based bed load transport prediction. *Water Resources Research*, 46(3), W03518. Available from: https://doi.org/10.1029/2009WR008007

Reid & Laronne, J.B. (1995) Bed load sediment transport in an ephemeral stream and a comparison with seasonal and perennial counterparts. *Water Resources Research*, 31(3), 773–781. Available from: https://doi.org/10.1029/94WR02233

Reid, I., Layman, J.T. & Frostick, L.E. (1980) The continuous measurement of bedload discharge. *Journal of Hydraulic Research*, 18(3), 243–249. Available from: https://doi.org/10.1080/00221688009499550

Reid, I., Powell, D.M. & Laronne, J.B. (1996) Prediction of bed-load transport by desert flash floods. *Journal of Hydraulic Engineering*, 122(3), 170–173. Available from: https://doi.org/10.1061/(ASCE)0733-9429(1996)122:3(170)

Richards, K.S. & Milne, L.M. (1979) Problems in the calibration of an acoustic device for the observation of bedload transport. *Earth Surface Processes*, 4(4), 335–346. Available from: https://doi.org/10.1002/esp.3290040404

Rickenmann, D. (2017) Bed-load transport measurements with geophones and other passive acoustic methods. *Journal of Hydraulic Engineering*, 143(6), 03117004. Available from: https://doi.org/10.1061/(ASCE)HY.1943-7900.0001300

Rickenmann, D., Turowski, J.M., Fritschi, B., Wyss, C., Laronne, J.B., Barzilai, R., et al. (2014) Bedload transport measurements with impact plate geophones: comparison of sensor calibration in different gravel-bed streams. *Earth Surface Processes and Landforms*, 39(7), 928–942. Available from: https://doi.org/10.1002/esp.3499

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. Available from: https://doi.org/10.1016/J.EPSL.2014.07.019

Stark, K., Cadol, D., Varyu, D. & Laronne, J.B. (2021) Direct, continuous measurements of ultra-high sediment fluxes in a sandy gravel-bed ephemeral river. *Geomorphology*, 382, 107682. Available from: https://doi.org/10.1016/j.geomorph.2021.107682

Stark, K., Laronne, J. B., Cadol, D., Reid, I., Powell, M., Billi, P., Cohen, H., Lucía, A., Liébault, F., & Zapico, I. (2022) bedload from ephemeral channels. Harvard Dataverse.

Thappeta, S.K., Johnson, J.P., Halfi, E., Storz-Peretz, Y. & Laronne, J.B. (2023) Bed shear stress in experimental flash flood bores over dry beds and over flowing water: a comparison of methods. *Journal of Hydraulic Engineering*, 149(4), 04023001. Available from: https://doi.org/10.1061/JHEND8.HYENG-13029

Thorne, P.D. (1986) Laboratory and marine measurements on the acoustic detection of sediment transport. *Journal of the Acoustical Society of America*, 80(3), 899–910. Available from: https://doi.org/10.1121/1.393913

Whiting, P.J., Dietrich, W.E., Leopold, L.B., Drake, T.G. & Shreve, R.L. (1988) Bedload sheets in heterogeneous sediment. *Geology*, 16(2), 105–108. Available from: https://doi.org/10.1130/0091-7613(1988)016<0105:BSIHS>2.3.CO;2

Wyss, C.R., Rickenmann, D., Fritschi, B., Turowski, J.M., Weitbrecht, V. & Boes, R.M. (2016a) Measuring bed load transport rates by grain-size fraction using the swiss plate geophone signal at the Erlenbach. *Journal of Hydraulic Engineering*, 142(5), 04016003. Available from: https://doi.org/10.1061/(ASCE)HY.1943-7900.0001090

Wyss, C.R., Rickenmann, D., Fritschi, B., Turowski, J.M., Weitbrecht, V., Travaglini, E., et al. (2016b) Laboratory flume experiments with the Swiss plate geophone bed load monitoring system: 2. Application to field sites with direct bed load samples. *Water Resources Research*, 52(10), 7760–7778. Available from: https://doi.org/10.1002/2016WR019283

Yager, E.M., Venditti, J.G., Smith, H.J. & Schmeeckle, M.W. (2018) The trouble with shear stress. *Geomorphology*, 323, 41–50. Available from: https://doi.org/10.1016/j.geomorph.2018.09.008

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Stark, K., Cadol, D., Varyu, D., Halfi, E. & Laronne, J.B. (2024) Strengths and weaknesses of acoustic pipe microphone systems in ephemeral, sandy, gravel-bed rivers. *Earth Surface Processes and Landforms*, 1–14. Available from: https://doi.org/10.1002/esp.5774