Estimating Sand Bedload in Rivers by Tracking Dunes: a comparison of methods based on bed elevation time-series

Kate C.P. Leary[1,2] and Daniel Buscombe[3]

[1]Department of Geography, University of California Santa Barbara, Santa Barbara, CA, USA.
[2]Department of Earth and Environment Sciences, New Mexico Tech, Socorro, NM, USA.
[3]Sch. Earth & Sustainability, Northern Arizona University, Flagstaff, USA, AZ.

Correspondence: Kate Leary (learykcp@ucsb.edu)

Abstract. Quantifying bedload transport is paramount to the effective management of rivers with sand or gravel-dominated bed material. However, a practical and scalable field methodology for reliably estimating bedload remains elusive. A popular approach involves calculating transport from the geometry and celerity of migrating bedforms, extracted from time-series of bed elevation profiles acquired using echosounders. Various echosounder sampling methodologies of how to extract bed ~~elevations~~ elevation profiles exist. Using two sets of repeat multibeam sonar surveys with ~~large~~ high spatio-temporal resolution and coverage, we compute bedload using three field techniques (one actual and two simulated) for acquiring bed elevation profiles: repeat multi-, single-, and multiple single-beam sonar. Significant differences in flux arise between repeat multibeam and single beam sonar. Mulitbeam and multiple single beam sonar systems can potentially yield comparable results, but the latter relies on knowledge of bedform geometries and flow that collectively inform optimal beam spacing and sampling rate. These results serve ~~to guide~~ as a guide for design of optimal sampling ~~,~~ and for comparing transport estimates from different sonar configurations.

1 Introduction

Bedload is usually a significant proportion of total load in rivers with sand and/or gravel-dominated bed material, and the relative importance of suspended load and bedload often changes with flow and the location within the channel ~~(e.g. ?)~~ (e.g. Gomez, 1991). Whereas instrumentation and protocols for sampling suspended sediment loads are relatively well established (e.g. Nolan et al., 2005; Wren et al., 2000; Edwards and Glysson, 1999), reliable estimates of bedload are more difficult to obtain because bedload in transport is difficult to sample directly (e.g. Emmett, 1980; Gomez, 1991), define (e.g. Church, 2006; Yang, 1986), or estimate with empirical formulas (e.g. Van Rijn, 1984; Martin and Church, 2000). Therefore, the effectiveness of sediment management in river systems is often predicated on the accuracy and representativeness of available bedload measurements.

Reliable estimates of bedload transport have been shown to result from application of the Exner equation (Simons et al., 1965; Engel and Lau, 1980) to time-series of bed elevation profiles (Simons et al., 1965; Van Den Berg, 1987; Dinehart, 2002; Villard and Church, 2003; Wilbers and Ten Brinke, 2003; Claude et al., 2012; Guala et al., 2014) acquired with an echosounder. Simons et al. (1965) ~~first~~ showed that bedload flux can be estimated by tracking the average celerity, V_c, of the downstream migration of dunes with a known average height, H, and average length, λ. These variables are averaged over a field of dunes to satisfy the necessary assumptions that suspended sediment load, q_s, is in equilibrium ($d_{qs}/dx = 0$), and with continuity of mass ($d_{qb}/dx + d\eta/dt = 0$), where x and η are downstream distance and bed elevation, respectively (Simons et al., 1965). The Simons et al. (1965) approach therefore quantifies only the first-order bedload flux due to dune translation, not accounting for any exchanges in bed material load between suspended and bedload fractions that deform the dune and may contribute to net transport (McElroy and Mohrig, 2009).

Complicating matters, however, is the inherent variability of bedform size ~~a~~ and shape in natural river systems (**?**). Sediment and water discharge conditions ~~,~~ vary continuously in natural rivers causing bed morphology to be out of equilibrium with prevailing flow conditions. Numerous field studies suggest that bedform disequilibrium is likely the norm rather than an exception in natural river systems ~~(e. g., Frings and Kleinhans, 2008; Julien et al., 2002; Ten Brinke et al., 2009; Wilbers and Ten Brinke, 2003; Fielding et al., 2009).~~ (e.g. **?**Julien et al., 2002; Wilbers and Ten Brinke, 2003; Ten Brinke et al., 1999). Even with these complications, bedload flux estimated from translating dunes remains one of the most accurate bedload estimate techniques for sand-bedded rivers (Nittrouer et al., 2008; Wilbers and Ten Brinke, 2003; **?**) and as such bed elevation measurements are of great importance.

In practice, bed elevation profiles might be acquired in three ways using echosounders: repeat multibeam, single-beam, and multiple single-beam sonar (Fig. 1A). Repeat multibeam ~~measures~~ surveys measure a spatially extensive three-dimensional bed, $\eta(x, y, t)$, from a moving vessel, from which it is possible to independently and simultaneously estimate V_c, H, and λ. Single beam systems measure a one-dimensional bed, $\eta(t)$, at a single (x, y) location using a stationary (fixed reference frame) sonar. Single-beam ~~echosounder~~ echosounders have also been used to acquire data from a moving vessel (specifically for longitudinal profiles in shallow environments). The study presented herein does not address bed elevation profiles and resulting flux estimates calculated from moving vessel single beam methods. Lastly, multiple single-beam ~~measures~~ surveys measure a spatially limited three-dimensional bed, $\eta(x, y, t)$, at a few (x, y) locations using stationary sonar. All three sampling methodologies have been and are currently employed to collect bed elevation data in experimental (Blom et al., 2003; Van der Mark et al., 2008; Guala et al., 2014; Curran et al., 2015), natural fluvial (Simons et al., 1965; Ten Brinke et al., 1999; Ashworth et al., 2000; Julien et al., 2002; Parsons et al., 2005; Gaeuman and Jacobson, 2007; Nittrouer et al., 2008; Claude et al., 2012; Shugar et al., 2010; Rodrigues et al., 2015; Wintenberger et al., 2015; Huizinga, 2016; Kaplinski et al., 2017; Buscombe et al., 2017; Hackney et al., 2018), and estuarine/coastal environments (Villard and Church, 2003; Hoekstra et al., 2004; Schmitt and Mitchell, 2014).

There are practical benefits and drawbacks to each data collection method. Repeat multibeam is spatially extensive, but relatively expensive, only practical in relatively deep, safely navigable rivers, and limited in temporal coverage. Therefore, the use of in situ stationary echosounders ~~(also called altimeters, here referred to as single beam and multiple single beam systems)~~

is becoming an increasingly popular alternative (Gray et al., 2010). These systems are especially useful in shallow water, are less expensive than multibeam systems, and ~~generating~~ generate longer time-series (e.g. Moulton et al., 2014). However, only measuring the bed elevation at a single location means it is not possible to resolve V_c, H, and λ simultaneously ~~.~~ and, since space and time are not linearly substitutable (Guala et al., 2014), not trivial to estimate λ from V_c or vice versa. This has implications for bedload estimates that are explored in this paper.

Since different methodologies may be employed to collect bed elevation profiles, and thus calculate bedload transport rates, it is important that resulting bedload flux estimates are compared to test for consistency. Do differences in sampling methodology cause systematic differences in bed elevation profiles and bedload flux estimates? Are different sampling methods equally able to measure/account for ~~bedform disequilibrium dynamics? How does bedform disequilibrium~~ changes in bedform size and shape with changing flow conditions? If not, how do changes in bedform size affect the bed elevation measurements made by different types of echosounders?

Presently, it is unclear how differences in bed elevation data acquired with different methodologies translate to the fidelity with which dune migration is captured, and finally to bedload transport estimates. Furthermore, multibeam, single beam, and multiple single beam datasets do not generally exist in the same location and time, making direct comparison of these datasets difficult. In order to examine these issues, we establish a virtual experiment using a repeat multibeam dataset to directly compare bedload transport estimates calculated from bed elevation profiles extracted to mimic the three different field survey methodologies outlined above.

2 Methods

2.1 Study Area and Survey Data

We use an extensive repeat multibeam dataset consisting of bed elevation from a large area of migrating dunes at high spatio-temporal resolution. Data come from a approx. 300 m long by 40 m wide reach upstream of the Diamond Creek USGS gage site on the Colorado River in Grand Canyon National Park (Fig. 1B, 1C), where flows are regulated by releases through Glen Canyon Dam 385 km upstream. Flow depths at the downstream extent ranged from 6.8m on the lowest flow of 310 m^3/s, and 7.9m on the highest flow of 590 m^3/s, although there are deeper holes within the surveyed reach. We simulate data from simultaneous single beam and multiple single beam deployments by extracting time-series of bed elevations from the repeat multibeam datasets (Fig. 1D). This 'virtual echosounder' experiment allows us to directly compare flux estimates from all three methodologies. We assess the relative accuracy of the single beam and multiple single beam techniques at estimating bedload transport compared to repeat multibeam-derived bedload, and suggest practical guidelines for developing sampling and processing protocols that maximize accuracy.

Repeat multibeam surveys were collected at two different discharges (Fig. 2A) from just upstream of the Diamond Creek USGS gage site (Fig. 1C). All bathymetric data were collected using a Teledyne-Reson 7125 multibeam echosounder, with sensor attitudes provided by a vessel-mounted inertial navigation system, and positions telemetered to the survey vessel at 20 Hz using a robotic total station situated onshore on monumented control. Data were collected with a 50% overlap between adjacent sweeps, providing up to 1000 individual soundings per meter-squared. Each sounding was edited manually. Further details of this system, survey, and processing methods are given by Buscombe et al. (2014; 2017) and Kaplinski et al. (2017). The channel bed was entirely composed of fine to medium sand with no gravel patches (Buscombe et al., 2017). At each discharge, data were collected every 6-10 minutes for 12 hours. A digital elevation model of the riverbed was produced for each survey, using coincident 0.25x0.25 m grids. The March 2015 (around 283 m^3s^{-1}) and July 2015 (around 566 m^3s^{-1})repeat multibeam surveys occured during mostly increasing and decreasing hydrographs, respectively(Fig. 2A). The precision of the repeat surveys was high, with mean cell elevation standard deviation of 0.012m computed over rocks known to be immobile (Kaplinski et al., 2017).

2.2 Extraction of Bed Elevation Profiles

A 35 x 30 m subsection in approximately the middle of the area surveyed by the repeat multibeam was selected for detailed bedload analyses using repeat multibeam, single beam, and multiple single beam bed elevation profiles (Fig. 1C). This subsection was then divided into 40 different repeat multibeam bed elevation profiles (8.67 m in length, 3.67 m lateral spacing from one another; Fig. 1D) for March and 20 different repeat multibeam bed elevation profiles for July (17.34 m in length, 3.67 m lateral spacing from one another; Fig. 1D). The length of the bed elevation profiles was determined by considering the maximum dune wavelength. All repeat multibeam bed elevation profiles were detrended using the bedform tracking tool (BTT) developed by van der Mark et al. (2008). This tool detrends each ~~BEP~~ bed elevation profiles (BEP) using a weighted moving average~~and extracts bedform height and wavelength data. This produced 2, 720 individual repeat multibeam bedload transport estimates (and daily averages from the 40 bed elevation profiles)for March and 1, 740 individual bedload transport estimates (and 20 daily averages) for July.~~. After the BEPs are detrended, the BTT determines the zero upcrossing (i.e. points at which the profile positively crosses zero) and zero downcrossing (i.e. points at which the profile negatively crosses zero). The locations of crests and troughs are determined in the original BEP as follows: a crest is located at the maximum value between a zero up- and zero downcrossing; and vice versa, a trough is located at the minimum value between a zero down- and zero upcrossing. Bedform height is calculated at the vertical distance between crest and downstream trough. Bedform wavelength is calculated as the distance between two successive crests. For a more detailed explanation of the BTT please refer to van der Mark et al. (2008) and **?**.

Whereas repeat multibeam analyses can be carried out in two dimensions (Nittrouer et al., 2008; Abraham and Pratt, 2010; Abraham et al., 2011; Shelley et al., 2013), analyses were deliberately carried out using one-dimensional transects oriented with flow direction, so any anisotropic effects in flux (caused by dunes not aligned perpendicular to the flow) affected repeat multibeam, single beam and multiple single beam results equally.~~We have chosen not to incorporate the ISDOTTv2 dune differencing method (Abraham and Pratt, 2010; Abraham et al., 2011; Shelley et al., 2013) into the analyses presented herein as the resources to do so are not currently publicly available. We do, however, incorporate the missing triangles correction formulated by Shelley et al. (2013) into our calculations of bedload flux (see section 2.3, Equation 2).~~

Virtual single beam and multiple single beam echosounders were placed at the downstream end of each repeat multibeam BEP (Fig. 1D). ~~multiple~~ Multiple single beams have four virtual beams, one of which is the same beam location as the single beam virtual echosounders. Two different beam spacings were explored: 1) 0m-0.56m-1.16m-1.74m and 2) 0m-1.74m-3.48m-5.22m.

2.3 Calculating Bedload Transport

Bedload transport, q_b (m^{3}s^{-1}), was calculated using the Simons et al. (1965) formulation based on the two-dimensional Exner equation (Paola and Voller, 2005) for bed sediment mass conservation, assuming triangular dunes:

$$q_b = (1-p)V_c\frac{H}{2} - q_e - q_0 \tag{1}$$

where p is the porosity of the sand (0.35 was used here) and q_0 is a constant of integration (set to zero here; see McElroy and Mohrig (2009) for a discussion of the potential physical meaning of this term). The original formulation of (1) has been validated and extended by numerous studies (e.g. Willis and Kennedy, 1975; Engel and Lau, 1980; Havinga, 1983), most recently by Shelley et al. (2013) who proposed the addition of the q_e term, defined as:

$$q_e = \frac{V_c^2 \Delta t H}{2\lambda} \tag{2}$$

where Δt is the change in time between successive surveys. Physically, q_e represents an area of underpredicted transport in the original Simons et al. (1965) formulation. Shelley et al. (2013) developed q_e to account for that missing area, which is typically a small part of the flux, perhaps negligible within overall error in flux.. Note that Eq. (1) is averaged over a field of dunes ~~.~~ (i.e. the bed elevation profile).

The primary variables in the above equations are calculated differently for each type of bed elevation profile. For repeat multibeam, H and λ are calculated directly using the BTT. V_c is calculated using a cross-correlation of two consecutive bed elevation profiles (McElroy and Mohrig, 2009; Engel and Lau, 1980). Using Eq. 1, we calculated 2,720 individual repeat multibeam bedload transport estimates for March and 1,740 individual bedload transport estimates for July. These individual bedload transport estimates were then averaged to generate 40 average daily bedload transport estimates in March and 20 average daily discharge estimates in July.

Single beam data ~~consists~~ consist of time-varying elevation only (Fig. 1A, 1F) therefore λ must be estimated independently. This might be done by measuring dune wavelengths in the field (for example, by wading, SCUBA, or using a boat-mounted sonar or ADCP) while installing or maintaining the echosounder. To simulate such an exercise, we use the daily average wavelength calculated by the BTT from the repeat multibeam survey directly upstream of the virtual single beam echosounder. Celerity (V_c) is:

$$V_c = \frac{\lambda'}{T} \tag{3}$$

where T is the period, and λ' is the estimated average wavelength.

For multiple single beam data, average period and height can be measured directly from the bed elevation profiles, whereas (V_c) may be estimated in one of three different ways. The first, “original method~~"~~”, is the same as Eq. (3), in which each beam is treated as a separate BEP to produce four estimates of transport that are then averaged. The second “cross-correlation method~~"~~” is to use a cross-correlation of bed elevation profiles measured by two different beams to find the spatial offset or ‘lag’, l, between translated dunes:

$$V_c = \frac{D}{l * \Delta t} \tag{4}$$

where D is the distance between sensors. In a field situation, this is constrained by practical considerations, but here we are free to vary D to evaluate its effects. This method produces six estimates of bedload transport (from six pairs of four beams), as does the third, “manual method", in which velocity is:

$$V_c = \frac{D}{t_{m2} - t_{m1}} \tag{5}$$

where t_{m1} and t_{m2} are manually picked times at which a crest appears at each beam.

2.4 Sinusoid Model of ~~Growing~~ Bedform Growth and ~~Shrinking Bedforms~~Decay

To test the affects of ~~bedform disequilibrium (i.e.~~ the adjustment of bedform height and length to changing flow conditions~~)~~, a simple sinusoid model was used to simulate time-varying dune height and wavelength. Each detrended bed elevation series was approximated by:

$$\eta = A \sin(B + Cx) \tag{6}$$

Dune growth/decay was controlled by varying A (amplitude) and C (wavelength). Dune translation was controlled by B (shift). Dune wavelength was estimated from dune height according to the regressions of the relationship between bedform height and wavelength for each survey day (Figure 2C). Using Eq. (6), sinusoid single beam bed elevation profiles are constructed from the synthetic elevation series, η, at a single location, x. These profiles were then used to calculate synthetic single beam bedload transport estimates using Eq. (1) and (3). The average λ and H were used in these calculations to replicate the methods used for the virtual echosounder experiment. Synthetic repeat multibeam bedload transport rates are ~~then~~ calculated using Eq. (1). We then take the ratio of synthetic multibeam bedload transport estimates to synthetic single beam bedload transport estimates and use that ratio as a correction fact for our actual measurements (i.e. multiply the actual single beam estimates by the ratio determined in the sinusoid model).

3 Results

3.1 Dune Field Characteristics

Bedform height ~~(wavelength)~~ averaged 0.17m ~~(2.38m)~~ and 0.36m ~~(~~in March and July, respectively. Bedform wavelength averaged 2.38m and 5.1m ~~)~~ in March and July, respectively. Dune geometry was highly variable during both survey days, with standard deviations of bedform height (wavelength) of 0.05m (0.4m) in March and 0.2m (2.7m) in July. Discharge along the Colorado River in Grand Canyon fluctuates daily as a result of daily release flows from Glen Canyon Dam. In response to changes in discharge, bedform size almost doubled over the course of the survey in March and almost halved over the course of the survey in July (Fig. 2B). Due to the greater discharge, the bedforms in July are larger in height and wavelength (Fig. 2C) compared to those in March. A space-time plot of bed elevations shows bedform heights and wavelengths increasing over the duration of sampling in March (Figure 2D). Bedform crest traces become less frequent as bedform troughs deepen. In ~~contract, a space time~~ contrast, a space-time plot of bed elevations in July shows ~~bedfrom~~ bedform heights and wavelengths decreasing throughout the survey period (Figure 2D).

3.2 Repeat Multibeam v. Single Beam

We consider the repeat multibeam-derived bedload estimates to be the most accurate because the superior spatio-temporal coverage of these data allow for simultaneous resolution of V_c, H, and λ. Single beam-derived daily bedload transport rates are underestimated relative to repeat multibeam in March, and overestimated in July (Fig. 2E). This could be caused either by mischaracterization of V_c, H, or λ in either repeat multibeam or single beam calculations, or in both.

The most likely source of error in the repeat multibeam calculations occurs when calculating V_c. To investigate whether cross-correlation correctly measured translation of dunes, l was manually calculated from repeat multibeam bed elevation profiles ~~and~~ by picking the locations of crests and tracking them and then used to calculate bedform celerity. This showed that cross-correlation-derived V_c were underestimated in both March and July (Fig. 3A). This underestimation is much larger in July, when dunes were adjusting to decreasing flow conditions and deforming at a greater rate. This result indicates that caution should be exercised when using cross-correlation to derive V_c, especially during higher transport stages. The regressions between manual and cross-correlation computed V_c (Fig. ~~6~~3A) are used to calculate a lag-corrected celerity and lag-corrected bedload transport rates for the repeat multibeam data (Fig. ~~5~~2E). Correcting repeat multibeam estimates for cross-correlation lag errors results in 1.6% and 33.9% error for March and July, respectively. Percent error will be expressed relative to repeat multibeam-corrected lag flux estimates for the remainder of this paper.

Even with the lag-correction applied, discrepancies exist between repeat multibeam and single beam ~~flux~~ bedload transport estimates due to errors estimating V_c from estimated wavelength and observed period. In March, period computed from single beam data is overestimated relative to repeat multibeam period, causing V_c, and therefore transport, to be underestimated. The opposite is true for the July data (Fig. 3B).

These discrepancies in observed period are likely linked to the bed responding rapidly to unsteady flows during each survey (Fig. 2A), with changes in discharge causing commensurate changes in H (Fig. 2B) and λ (Fig. 2C). This suggests that dunes ~~in disequilibrium (i.e. growing/shrinking)~~ adjusting to unsteady flow conditions apparently distort the period observed in the single beam data, which would invalidate the assumption made in Eq. (3) that the daily average wavelength (or any invariant measure of wavelength) is representative.

3.3 Single Beam Correction from Sinusoid Model

To test the above hypothesis, synthetic repeat multibeam and single beam ~~flux~~ bedload transport estimates calculated from ~~the growing~~bedform growth/~~shrinking sinusoid models~~ decay madels are compared. Single beam bed elevation profiles of growing and ~~shrinking~~ decaying sinusoids display significantly different distributions of period compared to the assumption of constant bedform wavelength (Fig. 3C). When bedforms are ~~in disequilirbium the period recored by the signle~~ adjusting to changes in flow conditions the period recorded by the single beam profile changes to either be longer or shorter depending on whether the dunes are growing or ~~shrinking~~decaying, which effects bedload transport measurements. As dunes grow or ~~shrink~~decay, the ratio of synthetic repeat multibeam to synthetic single beam bedload transport increases or decreases, respectively. ~~The maximum sinusoid repeat multibeam-to-single beam ratio for growing (shrinking)~~ In the sinusoid model, the maximum ratio of synthetic repeat multibeam to synthetic single beam bedload transport rates for growing and decaying dunes is 1.2 ~~(~~and 0.75~~)~~, respectively. Applying these ratios as correction factors to the single beam estimates (i.e. multiplying the single beam estimates by these ratios) generates sine-corrected single beam transport estimates (Fig. 2E), resulting in a decrease of the discrepancy between repeat multibeam and single beam derived bedload transport rates from 45.3% to 27.7% in March and from 38.9% to 10.7% in July.

3.4 Repeat Multibeam v. Multiple Single Beam

Another potential practical solution to minimizing the distortion of period in single beam surveys is to use a multiple single beam echosounders in a spatial array (Figure 1). By increasing the spatial resolution of bed elevation data, multiple estimates of bedload may be obtained, as well as multiple options for computing V_c (Eq. (3) through (5)), two of which (Eq. (4) and (5)) do not require a priori estimation of bedform wavelength (λ'). We expect the period recorded by each beam to be similarly affected by growing/~~shrinking~~ decaying dunes as were the single beam periods. We therefore apply the same sine correction from above to multiple single beam flux estimates calculated with the “original method~~"~~”. Fig. 4A shows these results for the beam spacing of 0, 0.56, ~~-1.16, and -1.74~~ 1.16, and 1.74 m for the three methods for computing celerity, and Fig. 4B shows the bedload transport estimates using a larger beam spacing and Eq. (5) only. The original method of calculating celerity (Eq. (3)) produces an average percent error of 13.3% and 15.8% in March and July, respectively; suggesting that increasing the number of beams and incorporating a sinusoid correction can mitigate discrepancies with repeat multibeam estimates.

The cross-correlation method (Eq. (4)) systematically over-estimates bedload transport in both March (43.4% error) and July (108.4% error), suggesting that the lag is systematically underestimated, and hence overestimating celerity. The manual method (Eq. (5)) yields a small mismatch between multiple single beam and repeat multibeam derived bedload in March (1.3% error) but a 62.9% error in July. This could be related to beam spacing, because the bedforms (and bedload mismatches) in the July data are much larger than those in March. This could cause greater celerity because only between 10 and 30% of the dune wavelength is being captured by the multiple single beam with the smaller sonar spacing, increasing to 30-100% with the larger spacing of 0-1.74-3.48-5.22 m (Fig. 4B). However, increasing beam spacing does not fully resolve discrepancies between repeat multibeam and multiple single beam bedload estimates (36.6% error; Fig. 4B), suggesting another factor is contributing to the observed discrepancies, most likely temporal resolution.

Using a linear interpolation we increase the temporal resolution of the data from 6 to 3 minutes. At this new sampling frequency, the original method yields a 2% error in March, but continues to overestimate bedload transport in July (67.5% error; Fig. 4C). Increasing the temporal resolution of the data results in more accurate estimates of lag. The cross-correlation method yields a 6.8% error in March and a 16.3% error in July (Fig. ~~4D~~4C), suggesting that the temporal resolution of the multiple single beam data will cause variation in cross-correlation-derived estimates of V_c.

3.5 Lateral variations in calculated bedload transport rates

We also consider how bedload transport rates vary with distance across the channel (Fig. 5). During the March survey, mean bedload transport rates were slightly higher on river left compared to river right (Fig. 5A). We see a similar pattern in mean bedform height (Fig. 5B) but no recognizable pattern in mean bedform celerity (Fig. 5C). In the July survey, mean bedform transport rates are slightly larger on the on river right than on river left (Fig. 5A). This is again reflected in the bedform height distributions (Fig. 5B) but not the bedform celerity distributions (Fig. 5C). We also observe changes in the shape of the distributions of bedform transport rates and beform heights with distance across the channel. Near the banks of the channel, our bedform transport rate and bedform height distributions have a much larger range whereas these distributions have a much tighter range in the center of the channel. This is mostly likely caused by varying amounts of superimposed bedforms across the channel and suggests that the bedform decay process can be heterogeneous in the cross-channel direction.

The ratio of single beam to repeat multibeam bedload transport rates also changes with lateral position along the channel (Fig. 6). For this analysis we compare daily mean lag-corrected repeat mutlibeam bedload transport rates and daily mean single beam bedload transport rates. In March, we find the highest agreement between the two methods in the center of the channel, with high discrepancies between the two methods near both banks. In July, the highest agreement between the two methods is also in the center of the channel although only proximity to one bank (river left) showed significant discrepancies in bedload estimates.

4 Discussion

Bed elevation profiles recorded by repeat multi-, single-, and multiple ~~single- beam~~ single-beam sonar methodologies produce different estimates of bedload transport, but practical steps can be taken to reduce the mismatch.

Significant errors in ~~fluxes computed~~ bedload transport rates can arise for two main reasons: (1) cross-correlation derived repeat multibeam bedform celerity estimates can show systematic bias, and (2) dunes can grow/shrink in response to unsteady flow conditions or varying sediment supply (Martin and Jerolmack, 2013; Rodrigues et al., 2015). Caution should be exercised when using cross-correlation to derive dune celerity, especially during higher transport stages and for relatively large time increments between successive measurements. It is good practice to check lags estimated using cross-correlation with manual measurements in order to compile a relationship that can be used to correct for systematic bias in estimated lag (Fig. 3A).

Using single beam bed elevation profiles, as dunes grow, transport is underestimated because period is overestimated. As dunes ~~shrink~~decay, transport is overestimated because period is underestimated (Fig. 3B). It is therefore important to understand the time scales over which dunes size is responding to flow in order to assess the relative effect period distortion may be having on the bedload estimates. A sinusoidal growth model is proposed that accounts for geometric effects on bedload ~~flux~~ transport in unsteady flows, using measured dune heights and translations and a scaling relationship to predict dune wavelength from ~~its~~ dune height (Fig. 3C). Such a scaling relationship could be compiled over time for a specific single beam deployment and applied retroactively to entire time-series of bed elevation profiles. The sinusoid model could be applied in any operational setting where temporal variations in dune wavelength and a dune height-wavelength scaling relationship exist.

A less generally applicable extension to this procedure could involve modeling the spatio-temporal evolution of the bed more explicitly using Fourier series (e.g. Guala et al., 2014). Guala et al. (2014) demonstrate a frequency dispersion in the relationship between dune celerity, V_c, and wavelength, λ, because small dunes tend to move faster than larger ones. This does not bias our computed bedload fluxes from multibeam data since we use time-series of bedform statistics from $\eta(x,y,t)$, however it does place limits on any calculation of equivalent statistics from $\eta(t)$ because it requires assuming a model that relates average V_c with average λ, or rather that the functional form between them does not vary in time, which may not be strictly true.

In this study, accounting for temporal changes in dune geometry accounted for 28.9 (March) and 134.8 (July) tons/day in daily bedload rates computed using single beam, or 17.6% (March) and 28.3% (July) compared to lag corrected repeat multibeam-derived rates. It is worth noting that real single beam echosounders can operate at a finer temporal scale than we are able to approximate in our virtual experiment. Our analysis indicates that single beam measurements made at a fine enough scale could greatly reduce the error in bedload flux estimates.

Increasing the spatial resolution of the bed elevation data by using a multiple single beam system does not necessarily improve upon single beam transport estimates. Multiple single beam transport estimates do not suffer from distortions in period caused by changing dune wavelength but are sensitive to both beam spacing and sample frequency (Fig. 4). Ideally, sonar beams should be spaced such that a large proportion of the dune wavelength is sampled (Fig. 4B), although this is not always practical, especially in shallow water. If dune wavelengths change significantly according to flow, designing sampling to be optimal for a particular wavelength would not be recommended. A more effective approach to maximizing multiple single beam-derived bedload accuracy is to adjust sampling rate (Fig. 4C and D), calibrated in relation to a known range of dune migration rates. This is especially helpful for dune celerity estimates based on cross-correlation (Fig. 4A). We found the most accurate way to measure dune celerity from multiple single beam data is to measure time elapsed between successive dune crests.

Our analyses also suggest that lateral position along the channel can significantly affect single beam bedload transport estimates (Figs. 5 and 6). When mounting single beam or multiple single beam systems, its important to be cognizant of proximity to channel margins. We found that the center of the channel yielded the most comparable magnitudes of bedload transport rates between sampling methods. In our study area, the center of the channel is very near the thalweg of the channel. In rivers in which the thalweg differs significantly from the center of the channel, we suggest mounting the echosounder in a position near the thalweg if possible.

5 Conclusions

~~In summary, repeat~~ All practical methods for estimating bedload from dune observations suggested to date require a time-series of bed elevation changes associated with dune migration, either $\eta(t)$ using a single beam echosounder or $\eta(x,y,t)$ using a multibeam sonar. In lieu of a parameterization for bedform migration speed appropriate for field scales, forced by routinely measurable quantities such as discharge, ambiguity between wavelength and period can negatively affect bedload transport estimates from single beam echosounders. Here, a simple model to simulate dune wavelength changes was sufficient to resolve much of the ambiguity. Repeat multibeam-derived elevation time-series are a more accurate means with which to estimate bedload than using single beam or multiple single beam, because the superior spatio-temporal coverage of these data allow for simultaneous resolution of V_c, H, and λ. However, there are significant practical advantages to using single beam or multiple single beam systems over repeat multibeam, and their capacity to monitor bedload over long periods may in some situations outweigh any disadvantages to do with greater errors in instantaneous bedload flux. We have offered a case study and practical guidelines to maximizing the efficacy of comparing bedload transport estimates derived from different sampling methodologies, which collectively will guide design of optimal bed sampling strategies for tracking dunes in rivers.

Code and data availability. Data and related codes are available at doi.org/10.5967/M02J6904.

Author contributions. Kate Leary designed the virtual experiments, developed code, and performed the simulations. Daniel Buscombe provided invaluable insight and helped with data analysis. Kate Leary prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was funded by the Glen Canyon Dam Adaptive Management Program administered by the U.S. Bureau of Reclamation, through a cooperative agreement with the U.S. Geological Survey Grand Canyon Monitoring and Research Center. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. government. Thanks to Matt 'The Colonel' Kaplinski for leading the field surveys and processing all multibeam data, and to Bob Tusso, Erich Mueller, and Tom Ashley for field support. Dave Topping, Brandon McElroy, Paul Grams, Dave Dean, Mark Schmeeckle, and Kelin Whipple provided useful discussion.

References

Abraham, D. and Pratt, T.: Missouri River bed-load computations at Kansas City and Washington, ERDC-CHL Data Report, 2010.

Abraham, D., Kuhnle, R. A., and Odgaard, A. J.: Validation of bed-load transport measurements with time-sequenced bathymetric data, Journal of Hydraulic Engineering, 137, 723–728, 2011.

Ashworth, P. J., Best, J. L., Roden, J. E., Bristow, C. S., and Klaassen, G. J.: Morphological evolution and dynamics of a large, sand braid-bar, Jamuna River, Bangladesh, Sedimentology, 47, 533–555, 2000.

Blom, A., Ribberink, J. S., and de Vriend, H. J.: Vertical sorting in bed forms: Flume experiments with a natural and a trimodal sediment mixture, Water Resources Research, 39, 2003.

Buscombe, D., Grams, P., and Kaplinski, M.: Characterizing riverbed sediment using high-frequency acoustics: 1. Spectral properties of scattering, J. Geophys. Res. Earth Surf., 119, 2647–2691, 2014.

Buscombe, D., Grams, P., and Kaplinski, M.: Compositional signatures in acoustic backscatter over vegetated and unvegetated mixed sand-gravel riverbeds., J. Geophys. Res. Earth Surf., 122, 1771–1793, 2017.

Church, M.: Bed material transport and the morphology of alluvial river channels, Annu. Rev. Earth Planet, 34, 325–354, 2006.

Claude, N., Rodrigues, S., Bustillo, V., Bréhéret, J., Macaire, J., and Jugé, P.: Estimating bedload transport in a large sand-gravel bed river from direct sampling, dune tracking and empirical formulas, Geomorphology, 179, 40–57, 2012.

Curran, J. C., Waters, K. A., and Cannatelli, K. M.: Real time measurements of sediment transport and bed morphology during channel altering flow and sediment transport events, Geomorphology, 244, 169–179, 2015.

Dinehart, R.: Bedform movement recorded by sequential single-beam surveys in tidal rivers, J. Hydrol., 258, 25–39, 2002.

Edwards, T. and Glysson, G.: Field methods for measurement of fluvial sediment, 03-C2, U.S. Geological Survey: Information Survices, 1999.

Emmett, W.: A field calibration of the sediment-trapping characteristics of the Helley-Smith bedload sampler, U.S. Geol. Surv. Prof. Pap., pp. 1–44, 1980.

Engel, P. and Lau, Y.: Computation of bedload using bathymetric data., Proc. Am. Soc. Civ. Eng., 106, 369–380, 1980.

Gaeuman, D. and Jacobson, R.: Field assessment of alternative bed-load transport estimators, Journal of Hydraulic Engineering, 133, 1319–1328, 2007.

Gomez, B.: Bedload Transport, Earth Sci. Rev., 31, 89–132, 1991.

Gray, J., Laronne, J., and Marr, J.: Bedload-surrogate monitoring technologies., U.S. Geological Survey Scientific Investigations Report 2010-5091, pp. 1–37, 2010.

Guala, M., Singh, A., BadHeartBull, N., and Foufoula-Georgiou, E.: Spectral description of migrating bed forms and sediment transport, J. Geophys. Res. Earth Surf., 119, 123–137, 2014.

Hackney, C. R., Darby, S. E., Parsons, D. R., Leyland, J., Aalto, R., Nicholas, A. P., and Best, J. L.: The influence of flow discharge variations on the morphodynamics of a diffluence–confluence unit on a large river, Earth Surface Processes and Landforms, 43, 349–362, 2018.

Havinga, H.: Bedload discharge coefficient, J. Hydraul. Eng., 109, 157–160, 1983.

Hoekstra, P., Bell, P., van Santen, P., Roode, N., Levoy, F., and Whitehouse, R.: Bedform migration and bedload transport on an intertidal shoal, Continental Shelf Research, 24, 1249–1269, 2004.

Huizinga, R.: Bathymetric and velocimetric surveys at highway bridges crossing the Missouri River near Kansas City, Missouri, June 2–4, 2015, U.S. Geological Survey Scientific Investigations Report 2016–5061, pp. 1–93, 2016.

Julien, P., Klaassen, G., Ten Brinke, W., and Wilbers, A.: Case study: bed resistance of Rhine River during 1998 flood, Journal of Hydraulic Engineering, 128, 1042–1050, 2002.

Kaplinski, M., Hazel, J.E., J., Grams, P., Kohl, K., Buscombe, D., and Tusso, R.: Channel mapping river miles 29-62 of the Colorado River in Grand Canyon National Park, Arizona, May 2009, U.S. Geological Survey Open-File Report 2017-1030, pp. 1–35, 2017.

Martin, R. and Jerolmack, D.: Origin of hysteresis in bed form response to unsteady flows, Water Resour. Res., 49, 1314–1333, 2013.

Martin, Y. and Church, M.: Re-examination of Bagnold's empirical bedload formulae, Earth Surf. Proc. Land., 25, 1011–1024, 2000.

McElroy, B. and Mohrig, D.: Nature of deformation of sandy bed forms, J. Geophys. Res. Earth Surf., 114, 2009.

Moulton, M., Elgar, S., and Raubenheimer, B.: Improving the time resolution of surfzone bathymetry using in situ altimeters, Ocean Dynamics, 64, 755–770, 2014.

Nittrouer, J. A., Allison, M. A., and Campanella, R.: Bedform transport rates for the lowermost Mississippi River., J. of Geophys. Res.: Earth Surf., 113, 2008.

Nolan, K., Gray, J., and Glysson, G.: Introduction to Suspended Sediment Sampling, USGS Scientific Investigations Report 2005-5077, 2005.

Paola, C. and Voller, V.: A generalized Exner equation for sediment mass balance, J. Geophys. Res, 110, F04 014, 2005.

Parsons, D., Best, J., Orfeo, O., Hardy, R., Kostaschuk, R., and Lane, S.: Morphology and flow fields of three-dimensional dunes, Rio Paraná, Argentina: Results from simultaneous multibeam echo sounding and acoustic Doppler current profiling, Journal of Geophysical Research: Earth Surface, 110, 2005.

Rodrigues, S., Mosselman, E., Claude, N., Wintenberger, C. L., and Juge, P.: Alternate bars in a sandy gravel bed river: generation, migration and interactions with superimposed dunes, Earth Surface Processes and Landforms, 40, 610–628, 2015.

Schmitt, T. and Mitchell, N. C.: Dune-associated sand fluxes at the nearshore termination of a banner sand bank (Helwick Sands, Bristol Channel), Continental Shelf Research, 76, 64–74, 2014.

Shelley, J., Abraham, D., and McAlpin, T.: Removing systemic bias in bed-load transport measurements in large sand-bed rivers, J. Hydraul. Eng., 139, 1107–1111, 2013.

Shugar, D. H., Kostaschuk, R., Best, J. L., Parsons, D. R., Lane, S. N., Orfeo, O., and Hardy, R. J.: On the relationship between flow and suspended sediment transport over the crest of a sand dune, Río Paraná, Argentina, Sedimentology, 57, 252–272, 2010.

Simons, D. B., Richardson, E. V., and Nordin, C. F.: Bedload equation for ripples and dunes, US Goveernment Printing Office, Washington D.C., 1965.

Ten Brinke, W., Wilbers, A., and Wesseling, C.: Dune Growth, Decay and Migration Rates during a Large-Magnitude Flood at a Sand and Mixed Sand–Gravel Bed in the Dutch Rhine River System, Fluvial sedimentology VI, pp. 15–32, 1999.

Van Den Berg, J.: Bedform migration and bed-load transport in some rivers and tidal environments, Sedimentology, 34, 681–698, 1987.

Van der Mark, C., Blom, A., and Hulscher, S. J.: Quantification of variability in bedform geometry, Journal of Geophysical Research: Earth Surface, 113, 2008.

Van Rijn, L.: Sediment transport, Part I: bedload transport., J. Hydraul. Eng., 110, 1431–1456, 1984.

Villard, P. and Church, M.: Dunes and associated sand transport in a tidally influenced sand-bed channel: Fraser River, British Columbia, Can. J. Earth Sci., 40, 115–130, 2003.

Wilbers, A. W. E. and Ten Brinke, W.: The response of subaqueous dunes to floods in sand and gravel bed reaches of the Dutch Rhine, Sedimentology, 50, 1013–1034, 2003.

Willis, J. and Kennedy, J.: Sediment discharge of alluvial streams calculated from bedform statistics., Tech. rep., Iowa Institute of Hydraulic Research Report No. 202, University of Iowa, Iowa City, Iowa, 1975.

Wintenberger, C. L., Rodrigues, S., Claude, N., Jugé, P., Bréhéret, J.-G., and Villar, M.: Dynamics of nonmigrating mid-channel bar and superimposed dunes in a sandy-gravelly river (Loire River, France), Geomorphology, 248, 185–204, 2015.

Wren, D. G., Barkdoll, B. D., Kuhnle, R. A., and Derrow, R. W.: Field techniques for suspended-sediment measurement., J. Hydraul. Eng., 126, 97–104, 2000.

Yang, C.-S.: On Bagnold's sediment transport equation in tidal marine environments and the practical definition of bedload, Sedimentology, 33, 465–486, 1986.

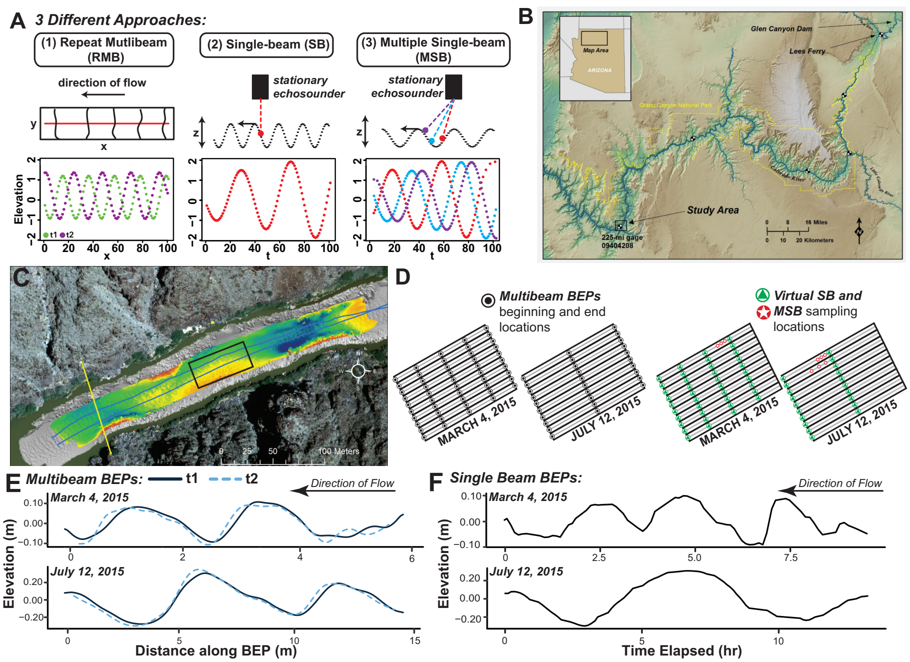

Figure 1. (A) Schematic of three common field methodologies for collecting bed elevation profiles (repeat multibeam, single beam, and multiple single beam) and the types of bed elevation profiles produced by each method. (B) Location of study area on the Colorado River in Grand Canyon National Park. Map adapted from Kaplinski et al. (2017). (C) Map of study reach. Yellow line indicates the location of the Diamond Creek USGS gage. Grey ~~section~~ area indicates area mapped with single multibeam survey. Colored area indicates area over which repeat multibeam surveys were collected (colors indicate elevations where red is high and blue is low). The blue lines that bisect the repeat multibeam survey area indicate the track lines the boat drove along in order to obtain each survey. Black rectangle indicates area in which bed elevation profiles were extracted. (D) ~~BEP~~ Bed elevation profile extraction map for multibeam~~and~~, virtual ~~single-~~ single beam, and virtual multiple ~~single-~~ single beam. Examples of repeat multibeam (E) and ~~single-beam~~ single beam (F) bed elevation profiles for March and July data sets.

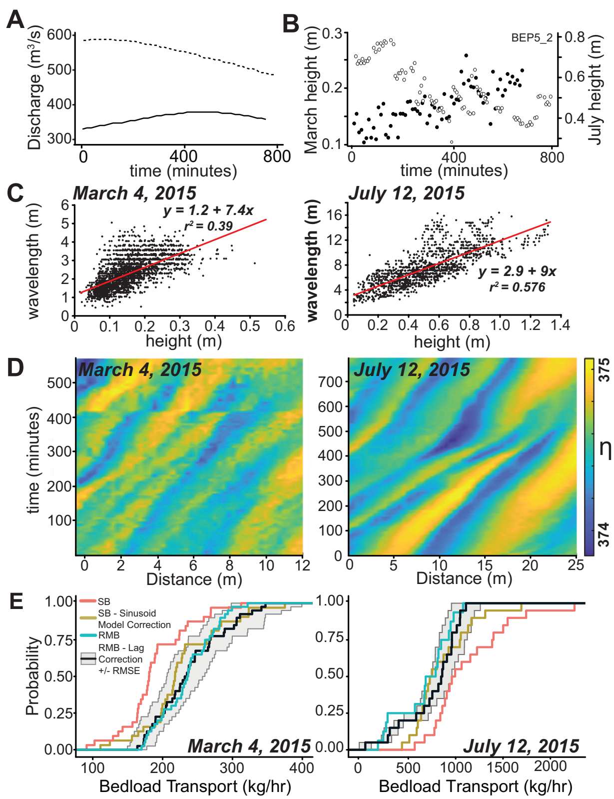

Figure 2. (A) Discharge during the sample time period. Dashed line is July data, solid line is March data. (B) Example of ~~bed form~~ bedform height varying with time from ~~BEP 5_2~~a single bed elevation profile. Open circles indicate July data, closed circles indicate March data. (C) ~~Height~~ Wavelength versus ~~wavelength~~height. Red line indicates linear regression of the data. (~~D~~E) Cumulative density plots of single beam and repeat multibeam bedload transport estimates with added corrections for mischaracterized lag (repeat multibeam) and period (single beam).

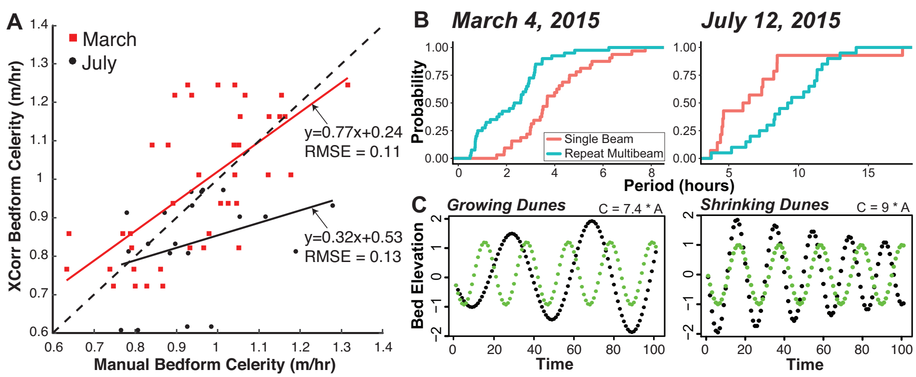

Figure 3. (A) Bedform celerity calculated using a manually picked lag versus a lag ~~produce~~ produced using a cross-correlation algorithm. The best fit linear regression ~~of the data~~ for each survey day serves as a correction factor for repeat multibeam celerity ~~estiamtes~~estimates. Dashed line represents ~~x=y~~ 1:1 relation. RMSE signifies the root mean squared error. (B) Cumulative density plots of period measured from ~~single-~~ single beam and repeat multibeam ~~BEPs~~bed elevation profiles. (C) Sinusoid model showing what ~~the signle~~ single beam ~~BEP~~ bed elevation profiles would look like if dunes ~~were growing~~grow/~~shrinking~~ decay (black) or if dunes ~~remained~~ remain the same size through time (green).

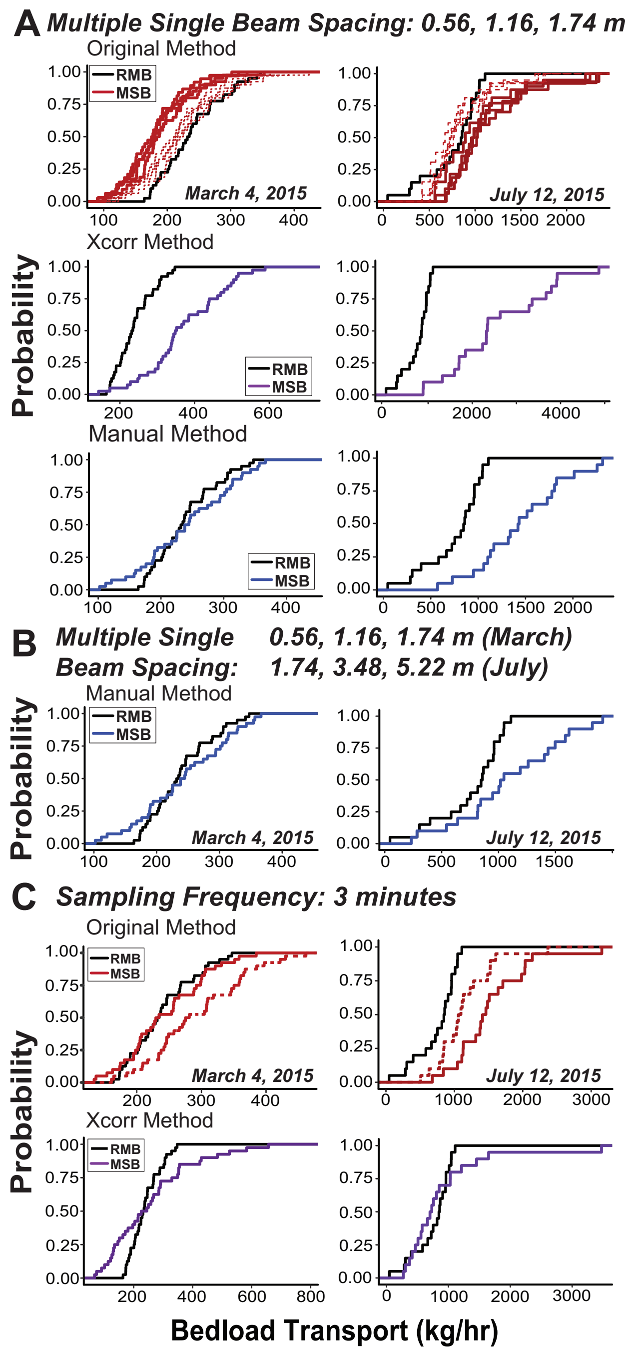

Figure 4. CDFs of lag-corrected repeat multibeam (RMB; black) and multiple single beam (MSB) bedload transport estimates using the original (red), cross-correlation (Xcorr; purple), and manual methods (blue) to calculate bedform celerity for the ~~MSB~~ multiple single beam profiles. Dashed red lines are sine-corrected estimates. (A) Multiple single beam beam-spacing of 0 - 0.56 - 1.16 -1.74 meters for both March and July. (B) Multiple single beam beam-spacing of 0 - 1.74 - 3.48 - 5.22 meters for July. (C) Bedload transport estimates with 3 minute sampling frequency.

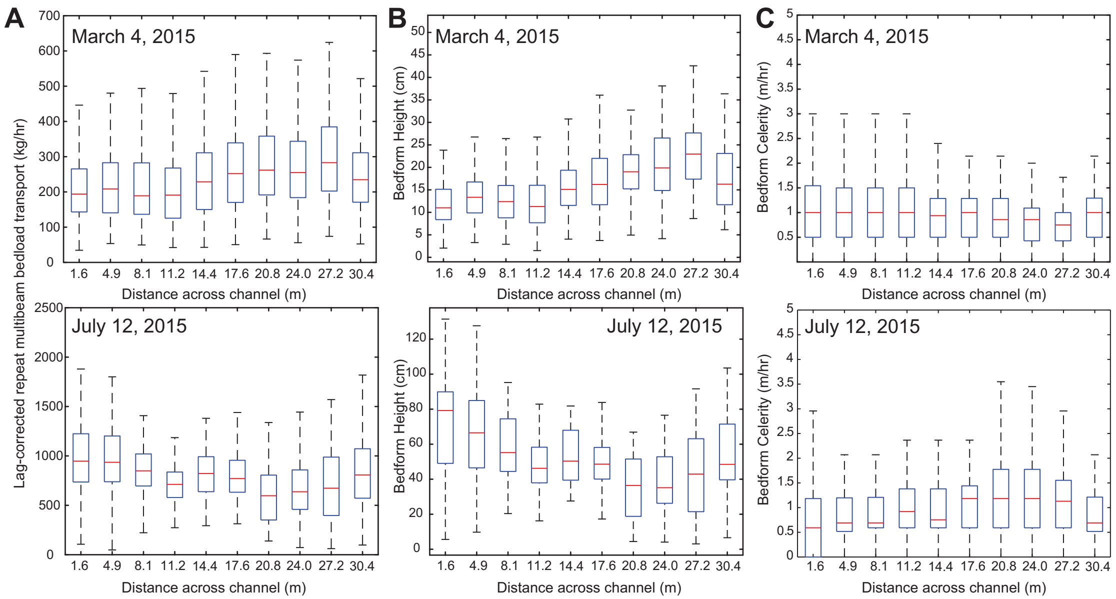

Figure 5. Boxplots of lag-corrected repeat multibeam bedload transport estimates (A), bedform heights (B), and bedform celerities (C) with varying distance across the channel (from river right to river left).

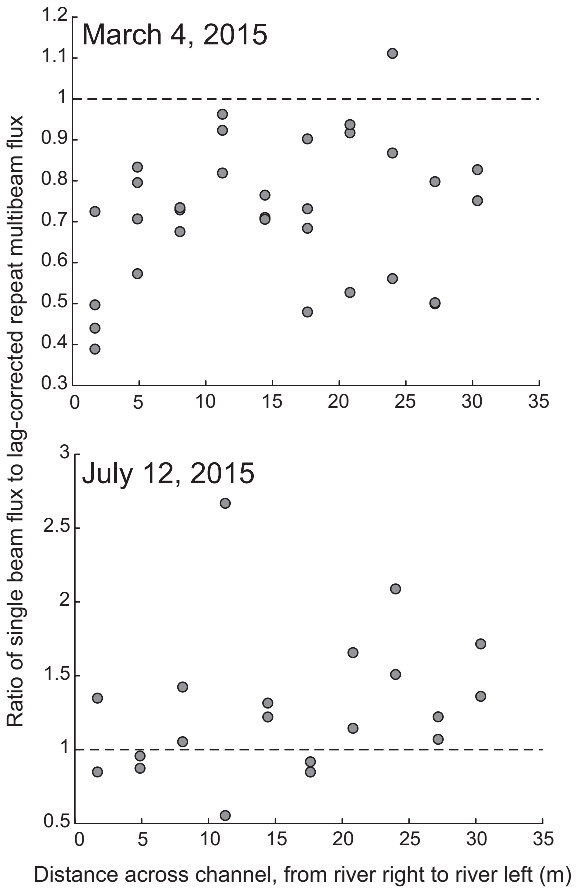

Figure 6. The ratio of daily mean single beam bedload transport estimates to daily mean lag-corrected repeat multibeam bedload flux estimates with varying distance across the channel (from river right to river left). Dashed line indicates a ratio of 1, where we expect comparable magnitudes of bedload transport estimates between the two methods.

We would first like to thank the anonymous reviewer for their thoughtful review. In the below document, the reviewers comments are in black; our responses to reviews are in blue italics.

This manuscript compares three different techniques to track bedforms and estimate bedload transport rates. This paper could be very useful for scientists who consider estimating bedload transport rates by bedform tracking, even though the paper does not include new methods. In general, the introduction, discussion and conclusion are very clear and informative. However, the methods and results are sometimes more difficult to read and need extra sentences to explain the concepts and how the conclusions are derived from the results. See my comments below.

Specific comments:
- P2, L30: "(also called altimeters. . .)", depending on the importance of this message, should this be mentioned earlier in the text and not between brackets?
We have removed this information from the text, because we feel "stationary single beam echosounder" is well understood.

- What is the difference between the second and third research question at the end of the introduction? It reads like it is the same question, but then the other way around.
The second question asks if changes in bedform size and shape affect measurements from different sampling methods. The third question asks how it affects the measurements. We have reworded these in the text for clarity.

- What is the possible influence of the study area on the results? In the introduction there is a distinction between shallow and deep rivers when mentioning the practical use of the multibeam and single beam, is the study area shallow or deep?
The study area is quite deep (6-9 m depending on discharge) so multibeam is a practical choice. This information has been added to the revised manuscript.

In the results, it is mentioned that there is a daily discharge variation that influences the bedform dimensions, how extreme are these discharge variations compared to other rivers and would this influence the advice in the discussion?
On a daily timescale, these changes are pretty significant compared to other rivers. That being said, sediment and water discharge conditions vary continuously in natural rivers causing bed morphology to often be out of equilibrium with prevailing flow conditions. As we state in our introduction, numerous field studies suggest that bedform disequilibrium is likely the norm rather than an exception in natural river systems (e.g., Frings and Kleinhans, 2008; Julien et al., 2002; Ten Brinke et al., 2009; Wilbers and Ten Brinke, 2003). It is for this reason that the sinusoid model was developed, to account for non-stationarity in the flow causing increases or decreases in bedform dimensions in time. The sinusoid model is suggested as a utility for such situations to minimize the error in single beam style estimates of bedload flux from stationary echosounders in unsteady flow.

- Is there an effect expected of using virtual single and multiple single beam profiles based on the multibeam data, instead of measuring it separately and thus independently in the field?
There are arguments both ways: (1) The benefit of this virtual experiment is that we know the virtual single beam echosounders are measuring the exact same bedforms the multibeam is measuring, so independent measurements might have more error. (2) That being said, real single beam echosounders can operate at a finer temporal scale than we are able to approximate in our virtual experiment. As shown by our virtual experiment, temporal resolution makes a big difference in bedload flux estimates. If independent single beam measurements were made a fine enough scale, this could greatly reduce the error. We also were limited by our field site, which has no bridge access. So at this location we were not able to take independent single beam measurements. We have added text to the discussion in regards to this question.

- Section 2.2, L11-14: fluxes caused by dunes that are not aligned perpendicular to the flow are ignored to be able to compare the results between multibeam and single beam. How much is this expected to influence the estimated bedload transport? Is this taken into account in other multibeam studies? The effect of varying dune dimensions due to disequilibrium with the flow is taken into account, should transport direction be taken into account as well?
Single beam echosounders would not be able to assess transport in other directions besides streamwise. However, the multibeam data are chosen specifically to be comparable to single beam data. As such, it isn't in the scope of any paper to look for directionality in a single beam trace.

- Section 2.2, L14: "we have chosen not to incorporate the ISDOTTv2": add a short explanation of what this method entails.
We have removed this section of the text on the advice of another reviewer.

- Section 2.2: I think the readability of this section could be improved by removing some of the information between brackets and incorporate it in the sentence. E.g. line 9-10, line 13. This might be a personal preference, but in general it feels like there is important information between brackets throughout the paper and therefore this information seems less important and less clear. Another example is the definition/cause of bedform equilibrium in the first sentence of section 2.4. I think some definitions and explanations will be clearer when this is explained in extra sentences.
We have reworded and reorganized this section for clarity.

- Section 2.3, L28: what is the physical meaning of q_e and why does it need to be added? Could you add a short explanation?
Q_e represents an area of underpredicted transport discussed by Shelley et al. (2013). The area represented by C in Shelley et al. (2013) figure 1 (see below) is not accounted for using the original method of Simons et al. (1965). q_e is the area of triangle D and therefore accounts for that missing portion. We have added a short explanation of q_e to the text. It is a very small contribution to total computed flux.

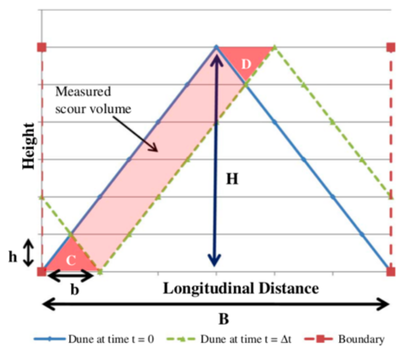

Fig. 1. Dimensions of a triangular sand dune

- Section 2.3: Is it possible to calculate an estimated average wavelength from the time series since you can estimate celerity from this? Would it differ a lot from the spatial estimate?
It's not possible to calculate an average wavelength from the single-beam timeseries because it only measures elevation through time. Guala et al (2014) showed that bedform space-time substitution in this way cannot work; imposing a relationship between the wavenumber and frequency spectra breaks down because small bedforms travel faster on average than large bedforms.

- Section 2.4: this section misses an explanation of why the bedform disequilibrium is determined Even though this is mentioned before, it would help the reader to repeat this here shortly. Furthermore, it is explained how equation 6 is used to calculate synthetic bedload transport estimates, but not how this is used to determine bedform disequilibrium.
We are empirically accounting for bedform adjustment to changing in flow (i.e. bedform growth and bedform decay). We have updated the text to reflect this.

- Where are figures 5 and 6?
Those figure references were for a previous version and were mistakenly left in this version. We apologize for the confusion and have corrected the manuscript to reflect the correct figure references.

- Section 3.2, L 24-26: how are the lag-corrected bedload transport and celerity calculated? And the errors? This might be visible in figure 5 and 6, but the pdf only shows figures 1 to 4.
Please see figure 3A for the regressions and r-squared values mentioned in page 6 lines 22-23.

- Section 3.3: I don't really understand yet how the sinusoid model is used to correct the data. I think this would be clearer if the method section 2.4 explains this better. What do you mean with the ratio between synthetic multibeam and synthetic singlebeam?
We take the ratio of synthetic multibeam bedload transport estimates to synthetic single beam bedload transport estimates and use that ratio as a correction fact for our actual measurements (i.e. multiply the actual single beam estimates by the ratio determined in the sinusoid model). We've added and reworded the text to make this more clear.

- Section 3.3, L 16: is this compared to the multibeam that is corrected for cross- correlation lag errors?
This is in the sinusoid model. We have reworded for clarity.

- Figure 4B: There is only one line for the multiple single beam? Shouldn't there be more lines for different spacings?

We only use a different spacing for the July data. You can compare the CDFs to Figure 4A for the smaller beam spacing.

Technical corrections

We have corrected the below technical corrections in the main text.

- P1, L14: There is a "?" instead of a source
- Figure1C: I do not see the grey section that indicates the area that is mapped with the single multibeam survey.
- Section 2.2, L8: Did you define BEP before this? You can for example add "(BEP)" at line 2 of this section
- Figure 2: there seems to be a caption missing to panel D.
- Figure2B: what is BEP5_2?
- Figure2C: "height vs wavelength" shouldn't this be "wavelength vs height" (Y vs X)? - Figure3A caption: "estimates"
- Figure3C caption: ""single"
- Section 3.3, Line 11: "disequilibrium" and "single"
- Discussion line 30: is "(July)" missing after the 28.3%?

We would first like to thank Robert Mahon for their thoughtful review. In the below document, the reviewers comments are in black; our responses to reviews are in blue italics.

The authors present a systematic comparison of bedform bedload measurement techniques using a unique dataset. Using field data, as opposed to flume data as is often the case, the authors are able to investigate some of the complexities associated with systems evolving under unsteady flow conditions. The ultimate outcome of this paper can inform decisions on both multibeam sampling and processing strategies as well as the placement of single beam echosounder instrumentation on rivers to monitor bed- load flux. Thus the results of this paper are broadly relevant to river managers as well as to academic geomorphologists.

The overall flow and structure of the manuscript are quite clear. Figures are well placed into the manuscript context and are appropriate for fully describing the nature of the work. While I have no concerns that fundamentally call into question the nature of the science being done, there are a number of points which the authors could clarify or analyses that could be bolstered by more complete discussion. These comments are below:

I would like to see a description of the methods used to extract height and wavelength data from the BTT toolbox as it is a fundamental operation to the analysis in the paper. There are several methods for calculating these parameters, each of which have their respective advantages and disadvantages so it would be good for the authors to describe why the calculations employed in this toolkit are appropriate to their system.

We have added the following to our description of the BTT: "After the BEPs are detrended, the BTT determines the zero upcrossing (i.e. points at which the profile positively crosses zero) and zero downcrossing (i.e. points at which the profile negatively crosses zero). The locations of crests and troughs are determined in the original BEP as follows: a crest is located at the maximum value between a zero up- and zero downcrossing; and vice versa, a trough is located at the minimum value between a zero down- and zero upcrossing. Bedform height is calculated at the vertical distance between crest and downstream trough. Bedform wavelength is calculated as the distance between two successive crests. For a more detailed explanation of the BTT please refer to van der Mark et al. (2008) and van der Mark et al. (2007)."

Were bed elevation surveys corrected for apparent dilation as a function of the time between start and end of each multibeam survey? If not, was this considered and determined to be a negligible effect? See McElroy dissertation 2009, p. 44 (URI: http://hdl.handle.net/2152/1511).

The average difference in time between surveys was around 10 minutes and departures from this were also order 10 minutes, therefore, while we acknowledge that the effect noted by the reviewer and McElroy (2009) is real, we determine it to be a negligible effect in the present study.

A figure demonstrating the cross-correlation results would be good to show, as a lot of discussion is based on issues resulting from velocity calculations.

See figure 3A.

In Page 5 Line 6 the method for estimating wavelengths for the singlebeam experiment is described as the daily average from the repeat multibeam. I wonder if this introduces potential for extra accuracy for this method that may not be possible in a situation in which a single beam fixed echosounder would be employed.

Yes, this is most likely the case. Our single beam flux estimates are probably more accurate than what one might get using a different estimate of bedform wavelength.

I would suggest more discussion of when a situation would arise where you have a measurement or a daily average of bedform wavelengths but only a single beam profile to estimate flux from. An alternative formulation might be to estimate wavelength using a height-wavelength relationship such as Bradley and Venditti, 2017 as this might be a more realistic representation of a likely application (i.e. a deployed single beam sensor established for continuous monitoring).

Analyses of bedform fields throughout the Colorado river in Grand Canyon reveals that the Bradley and Venditti (2017) relations are a poor fit to observations. Because the bedform field is likely not in equilibrium with the flow due to daily fluctuations in discharge, i.e. because the dunes are always adjusting to flow, correlations between instantaneous bedform height or length and flow are not as robust as they otherwise would be. A better approach is to use mulitbeam measurements at multiple flows to develop a site-specific model for bedform dimensions as predicted by flow.

What did the manual process entail for determining bedform velocity? Were you picking crests and tracking them? Looking at the slopes of the forms in the $\eta(x,t)$ field (e.g. in Figure 2D)? It would be critical to determine whether the manual method itself includes any potential sources of bias in order to interpret its relation to the cross-correlation results.

We picked crest locations and tracked them. We have added this information to the main text.

I wonder if other methods for calculating bed velocity might be more appropriate than the cross-correlation method for this application, particularly given the unsteady flow conditions investigated. One example from Ganti et al., 2013 (doi:10.1002/jgrf.20094), their eq. 5 to compute the local velocity based on dividing the temporal change in local elevation by local slope at all points on the bed.

The above mentioned method from Ganti et al. (2013) would most likely not apply to this data set. While computing local velocities is appropriate in flume experiments where the change in time between each successive bed elevation profile is 45 seconds, applying this method to field data with both a coarser spatial and temporal resolution would likely result in large errors (likely larger than those associated with the cross-correlation method).

Were any physical bedload samples collected during the multibeam campaigns to compare with the ranges of flux measurements?

No physical bedload samples were collected. We rather doubt the reliability of measurements from a bedload sampler lowered from a cableway suspended high above the water surface through 7m of the water column, into a field of dunes up to a meter high moving up to a meter per second.

Some discussion is warranted of whether the bedform bedload equation of Simons et al., is even geometrically appropriate in situations where bedform growth/decay is occurring. I don't believe they considered this in their original work, and I am not aware of any later publications that show the validity of this method for non-steady bedform fields.

Aside from the assumption that dunes are appropriately triangular, there is nothing in the Simons et al derivation that suggests it is not appropriate for application in time-discerete fashion such as here. That is to say, despite the growth and decay in dunes, the instantaneous bedform flux as predicted by their instantaneous geometry and celerity is appropriate and has been applied before to unsteady flows. While it is true that inferring wavelength from time-series of bed elevation measurements is made more difficult by spatially accelerating and decelerating dunes, the issue is a more in the implementation of the theory governing Simons et al, rather than the theory itself.

Along similar lines I would encourage the authors to consider incorporating, or at least explaining the inappropriateness for their application, the insights from Guala et al. (2014, their Section 4 paragraph 2 in particular; doi: 10.1002/2013JF002759) in joint averaging of the elevation and velocity values.
We agree the study of Guala et al is pertinent so we have added the following to the revised Discussion: "Guala et al. (2014) demonstrate a frequency dispersion in the relationship between dune celerity, Vc, and wavelength, λ, because small dunes tend to move faster than larger ones. This doesn't bias our computed bedload fluxes from multibeam data since we use time-series of bedform statistics from η(x,y,t), however it does place limits on any calculation of equivalent statistics from η(t) because it requires assuming a model that relates average Vc with average λ, or rather that the functional form between them doesn't vary in time, which may not be strictly true."

While somewhat outside the scope of the review of the paper itself, I should note that the license type given to the dataset and code hosted in the SEAD repository is potentially quite restrictive to some river management uses and researchers, given that it does not allow commercial use or any derivatives. This may be less important for the data itself, but it may heavily limit the use of this work to have code that cannot be modified. A share alike restriction, for example, would make this more accessible.
We have updated the license to allow for commercial use and derivatives.

Line Comments: The following line specific comments are non-critical to the science of the manuscript and are meant to help improve readability or clarity.
We have corrected the below Line Comments in the manuscript.

Page 1 Line 2: "remains elusive" is relatively non-concrete and feels dismissive of the wealth of literature and practice on field-scale bedload measurement techniques spanning half a century or more.
Page 1 Line 14: references are missing at "(e.g. ?)"
Page 1 Line 20: References such as Simons et al., 1965 and others don't explicitly derive from the exner equation, per se. They are derivations of mass conservation but not necessarily predicated on Exner's formulations.

Page 2 Line 1: Simons wasn't the first to show this, as written. For example, Bagnold 1941, Chapter 13 derives a similar formulation, albeit with some geometric inaccuracies. I suggest simply removing the word "first" from the sentence.

Page 2 Line 9: remove comma after "...discharge conditions,"

Page 2 Line 12: is there a reference for "...bedload flux estimated from translating dunes remains one of the most accurate..."?

This claim has previously been made by Wilbers and Ten Brink (2003) and Nittrouer et al. (2008) who said the measurements came with "... relatively high accuracy so long as the dune geometry and translation distances are large relative to the positioning error" which is the case here. The relative disadvantages of direct sampling are well documented (e.g. Holmes and Holmes, 2010), for example direct sampling disturbs the flow and therefore the rate of bedload, and the sampler must be placed squarely on the bed surface to adequately sample, therefore the presence of dunes causes error.

Page 4 Line 14: ISDOTTv2 is not a familiar/common tool since it is not public. If you wish to include this statement, it would be good to describe what that tool is and why it would be useful here. Otherwise I would suggest removing it.
We have removed this statement.

Page 4 Line 16: please describe the "missing triangles" correction.

Page 4 Line 22: consider rewording as it states a 1965 reference is based on a 2005 reference.

Page 7 Line 14: "...for growing (shrinking) dunes is 1.2 (0.75)." I suggest rewording to "...for growing and shrinking dunes is 1.2 and 0.75, respectively."

Figures: for figures 1, 2 and 4 there are abbreviations used which would be helpful to have defined in figure captions so the reader doesn't have to remember or find from the text. BEP, RMB, SB and MSB are all used. Additionally, Xcorr and RMSE are used but not defined in captions or in the text body.

We would first like to thank Kory Konsoer for their thoughtful review. In the below document, the reviewers comments are in black; our responses are in blue italics.

In "Estimating Sand Bedload in Rivers by Tracking Dunes: a comparison of methods based on bed elevation time-series", the authors present a systematic comparison for different approaches for estimating bedload transport based on dune migration. The methods compared rely on repeat multibeam echo sounding surveys from a reach of the Colorado River during two different field campaigns that exhibit different discharges. The multibeam surveys provide the base data, and three different subsets from the data are selected. The three datasets used in the comparison are, 1) longitudinal transects of bed elevation from the full multibeam surveys, which provide spatial data series, 2) extraction of bed elevation at a single point over time (temporal), and 3) extraction of bed elevation at multiple points over time (temporal). The authors also include synthetic sinusoidal signals that are used to evaluate bedform dynamic of growing/shrinking size that would occur during unsteady flows.

Overall the paper is well written and organized, and the presentation of the results is very clear. The topic of this paper is also of great importance as river scientists still struggle with determining best practices for quantifying bedload transport rates. However, I would recommend addressing a few issues related to the methods and discussion before the manuscript should be accepted for final publication. I outline these below.

Although the data are measured using a multibeam echo sounder, the dataset is not fully utilized and instead only bed elevation profiles are extracted. Thus, the comparisons are essentially spatial series of single beam, stationary single beam, and stationary multi-single beam. It is stated that the reason for this is to account for anisotropy among the different methods equally (page 4, lines 11-14), which is understandable. However, as is stated more than twice throughout the manuscript, multibeam surveys are considered the most accurate due to the high spatiotemporal resolution, yet are not being used to their full potential.

Why have you decided not to include the full three- dimensionality of the multibeam survey when considering sediment transport? If you consider this to be most accurate, then you could conceivably have a fourth method using the repeat multibeam surveys as two dimensional differencing compared to the three "single beam" methods presented in the paper.
This is essentially the ISDOT method (Abraham et al., 2011), which requires there to be conservation of mass over the survey area (all sediment eroded from the area is deposited in the same area). Additionally, this method is designed for bedforms moving at a constant speed, with little to no deformation, and little to no suspended sediment. In our field data, the dunes change speed throughout the day, change shape significantly, and suspended sediment is available. Although the ISDOT method works well in a flume setting, we don't feel that it is applicable to our field data.

Similarly, it appears as though all the repeat multibeam bed elevation profiles have been averaged into a single value for the area of interest. Why not keep these separate and evaluate the comparisons spatially?
The repeat multibeam profiles are only averaged at each location, so within the area of interest there are 40 daily bedload transport estimates. We compute a daily average at each location because it is directly comparable to the measurements made by single beam and multiple-single beam echosounders. The CDFs in this paper illustrate the distribution of daily average bedload transport estimates for the entire area of interest.

From the bed elevation raster shown in figure 1 there appears to be quite a difference in elevation and bedform size from the left bank (higher bed elevation) to right bank (lower bed elevation). Is there a systematic difference in the comparisons from left to right? If so, is it related to bedform dimensions?
We have added a section to the results to address this point. Please see section 3.5 in the updated manuscript.

This spatial information would be extremely relevant for the discussion section. In particular, one of the topics I felt was missing from the discussion was how the findings of this study can be used to provide insight on where stationary single beam sensors could be installed. My understanding is that most single beam sonars are attached to bridge piers or off banks/docks. If a spatial component of comparison is included in this paper, it would be possible to inform deployments in future studies. Do your comparisons show less agreement between the methods closer to the bank? These are questions easily answered from your dataset without much additional analyses.
We have added a paragraph to the discussion section to address this point. Please see page 10, line X6

Could you provide more information on how the cumulative density plots are prepared? It is stated on page 4 line 30 that Eq. 1 is averaged over a dune field. There is no mention of how the CDF are prepared. How many bed elevation profiles are needed before a 'stationary' average bedload transport rate is obtained? How far apart do the lines need to be? Answers to these questions could help guide surveys using boat- mounted single beam sonars. (it is stated that this is not of concern for the paper, however the extracted profiles from the multibeam survey is essential that).
Equation 1 produces a bedload transport estimate that is the average for the entire bed elevation profile. This is because we are using an average bedform height and average dune celerity. Therefore, each timestep at each location has one bedload transport estimate. We then average all timesteps at each location for a daily average bedload transport rate. Thus CDFs for

repeat multibeam July data contain 20 estimates of daily bedload transport while repeat multibeam CDFs for March contain 40 daily bedload transport estimate. Single beam CDFs contain 20 and 40 bedload transport estimates for July and March data respectively.

There is reference to a figure 5 and figure 6 on page 6, but figures are only 1-4. I have attached an annotated pdf with other technical issues. Please see for grammar and other comments.
Thank you and apologies for the confusion. Those figure references were for a previous version and were mistakenly left in this version.

Please also note the supplement to this comment: https://www.earth-surf-dynam-discuss.net/esurf-2019-38/esurf-2019-38-RC3- supplement.pdf
We have corrected the grammatical and spelling errors highlighted in this supplement.