Author's response

First of all, we are grateful to the referees for their detailed analysis of our work. In this new version we took into account the 3 constructive and thoughtful reviews. All the points raised by the referees were addressed, leading to significant modifications of the initial version. Complementary analyses were also carried out to address some technical points raised during the review process (cf. documents uploaded on the website of ESD).

As proposed by Reviewer 1, we added new data and explanation on the hydrophone method and on the relation between Power Spectral Density and sediment grain size.

The detailed review by Reviewer 2 was the basis for re-analysing the aDcp data presented in the first version. We reconsidered and modified the paper as suggested and we propose new developments to be performed in future studies on the basis of several comments.

As proposed by Reviewer 3, we modulated the conclusions focused on the dune tracking method and present new data detailing the influence of the vertical position of the hydrophone on acoustic power variation.

We consider that this paper is original and could now be a valuable contribution for colleagues working on the dynamics and quantification of sediment transport in large sandy gravel bed rivers. Specifically, we highlight the strong potential of the hydrophone method to quantify bedload sediments in this type of rivers and its ability to describe bedload fluxes associated with bedform migration (at different spatial scales).

We are convinced that the results and interpretations presented in this paper could be of interest for a wide audience of specialists working on river morphodynamics, sediment transport and sedimentology.

On behalf of the authors

J. Le Guern

Anonymous referee 1

General comments (GC)

GC1

"Specifically, in relation to the hydrophone method, I would like to see a clearer explanation of the method and some data that demonstrate the variability of the power-spectral density between hydrophone recordings."

We evaluated the variability of the PSD for all hydrophone records (n=448, Figure1a). This variability is illustrated by computing the central frequency and the peak frequency in the frequency band of bedload (15-350 kHz) (Figure 1b). The central frequency is affected by the grain size and vary between 50 and 270 kHz with a median value of 140 kHz. The first quartile and the third quartile are 127 and 154 kHz respectively that reflect a small variability of the central frequency. This is in agreement with the median grain size of sediment (D_{50}) distribution that vary between 0.3 and 3 mm with a median value of 0.9 mm (Q1=0.8 mm and Q3=1.1 mm).
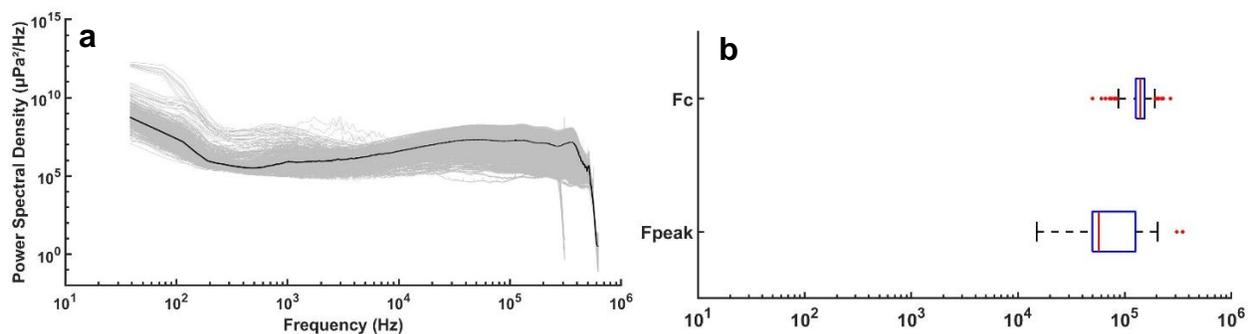

Figure 1: a) Power Spectral Density (PSD) for each hydrophone drift of the study (n=448), the median PSD in black; b) Distribution of the frequency peak and central frequency of each PSD.

GC2

"I would like the paper to describe if there are significant differences observed between the frequency spectrum of the studied sandy gravel-bed river and the frequency spectra in the comprehensive literature on gravel-bed rivers. How does the different grain size distribution affect the observed frequency spectrum?"

Different frequency spectrum from three different rivers are plotted in the Figure 2a. As expected by Thorne (1986), the central frequency decreases with the D_{50} (Figure 2a). We acquired 98 concomitant hydrophone records and sediment samplings and we compared these measured values to the theoretical relation of Thorne (2014) (Figure 2b). For the investigated range of D_{50}, the central frequency measured are in line with the theory (82% of the dataset are in the factor 2). For this range of grain size, Geay et al. (2020) found that rivers investigated did not follow Thorne's law because of the attenuation of high frequency sound waves in these steep rivers. The shift of the bedload frequency band towards high frequencies (figure 2a) means more bedload signal in this area of the PSD. However, the slope of the study reach is mild and the turbulence is relatively

low. Therefore, the high frequency sound wave are not attenuated and the acoustic power in high frequency region is not neglected. The use of hydrophone method in large lowland sandy-gravel bed rivers could be an asset considering the bedload frequency band of sands.
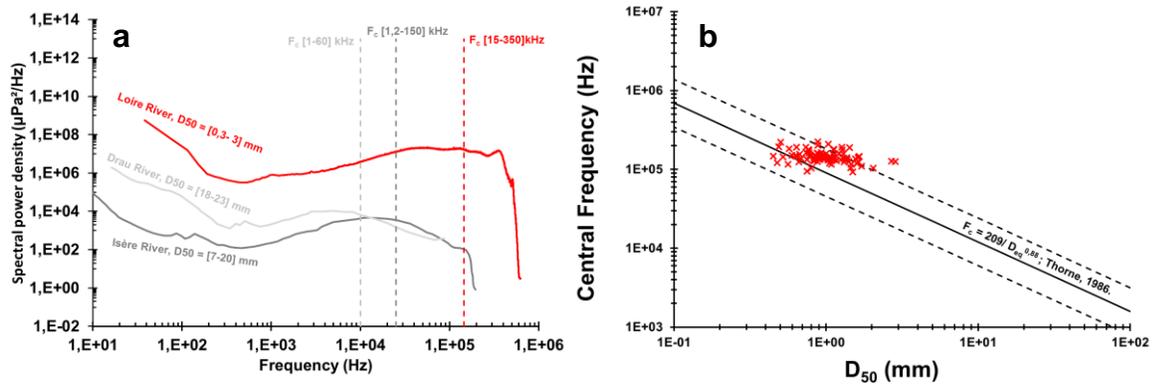

Figure 2: a) Comparison of PSD from 3 rivers with varying D_{50}; b) Comparison between measured central frequency of PSD of the Loire River and related D_{50} with relation of Thorne (1986).

We propose to add the figure 2a to the section 4.1. in order to provide PSD information related to other rivers, and also a paragraph: "*Moreover, the median PSD differ from the Isère River (Petrut et al., 2018) and from Drau River (Geay et al., 2017). These rivers are characterised by coarser sediments (see Fig. 5a) and the central frequency of the PSD are decreasing for an increasing D_{50}. These observations are in line with Thorne's (1986) theory. The central frequency of the median spectrum of the Loire is about 140 kHz. The frequency band of the bedload is shifted towards high frequencies due to finer grain size.*"

Specific comments (SC)

<u>*SC1*</u>

Lines 57-58 "to estimate the capacity of acoustic signals to detect the bedload axes on relatively wide cross-sections for various discharge conditions".

"Are you referring to estimating the accuracy of acoustic methods for measuring bedload transport rates for wide transects across a range of flow conditions?"

Yes, we are referring to the accuracy of acoustic methods to measure cross sectional variations of bedload fluxes over various discharge conditions. This sentence was modified in the revised version of the paper.

"to estimate the accuracy of acoustic methods to measure cross sectional variations of bedload fluxes for various discharge conditions;"

<u>*SC2*</u>

Line 116 "giving better results"

"How do you know that they are better results? Please provide a clearer description of the method used to determine data quality."

We just wanted to refer to Jamieson et al. (2011, p. 1066) observations here: *"A positive bias in bed velocity magnitude could be generated by interpolating bed velocity magnitude alone (i.e., without direction), where inconsistent bed velocity directions (which are typically higher in regions of zero to low bed velocity, where a moving bed is more difficult to detect with the ADCP) generates higher than expected bed velocity."*

We changed "*giving better results*" by "*limiting the over estimation*".

<u>*SC3*</u>

Lines 147-148 "and could allow a better understanding of the apparent bedload velocity gradient along bedforms."

"It's not clear to me what this means. Why would this give a better understanding of the bedload velocity gradient along bedforms? Is it because the ADCP footprint is relatively small compared with the bedform dimensions? If so, please state typical bedform dimensions relative to the footprint."

This is exactly what we mean. We propose to change this sentence by: "*This configuration permitted to fix the footprint for each beam to about 0.0046 m² and a distance 146 of 0.56 m between opposed beams. This allowed to describe apparent bedload velocity with a finer accuracy especially in the presence of bedforms of 0.2 m height and 3.9 m long (in average)."*

Lines 161-162 "The determination of a proxy to evaluate sediment transport directly from DTM measurements is difficult."

"Explain why it was difficult."

Because there are several dune parameters such as dune celerity, height or length that influence bedload rate and all these parameters are evolving according to discharge variations. We propose to modify this sentence by: "*The determination of a proxy to evaluate sediment transport directly from DTM measurements is difficult because dune migration is function of several parameters. A semi-empirical equation that integer these parameters was used to compare bedload transport rates with the reference measurement.*"

Lines 240-241 "and the integration of median PSD over a wider range of frequency in the present study."

"What is meant by median? I'm not sure what this is referring to and I think it needs to be explained more clearly. Is the PSD integrated across the median value of each frequency bin?"

To make things more clearly, we decided to add this comment to the final version: "*For each drift, a spectral probability density is computed (Merchant et al., 2013). Then a median Power Spectral Density is determined as done in Geay et al. (2017). Median PSD are preferred to mean PSD as it enables to filter rare and powerful acoustic events such as the hydrophone impinging the riverbed.*"

Lines 251-257 "The comparison can be made between indirect methods to discuss the acceptability of the BTMA reference. The apparent bedload velocity and the acoustic power are not well-correlated with mean dune morphological parameters (dune celerity and height). The aDcp method is measuring the apparent velocity of the grain being transported from the stoss to the lee side of a dune. It must be noted that apparent bedload velocity is higher than dune celerity with about a factor 100, whereas the grain size (D50) is smaller than dune height with the same order. Therefore, sediments that are 100 times smaller than dune height allows the dune migration with a celerity 100 times smaller than their own celerity."

"I'm not sure this section is necessary and I struggle to understand some of the arguments being made. What is meant by the acoustic power and apparent bedload velocity not being well correlated with the morphological parameters? Where is this shown in the data? The comparison between apparent bedload velocity and dune celerity is interesting but I think most readers would already understand that the bedload velocity is much smaller."

We wanted just to mention that relations between dune parameters (height and celerity) and acoustic methods are not consistent. We chose not to show these relations with figures to avoid overload of the paper. This can be illustrated by a short table of the coefficient of determination (COD) for each relation. We added this table.

Table 1: Coefficient of determination between dune parameters and acoustic methods (log values).

	P	V_a	$q_{s\ BTMA}$	H_{dune}	C_{dune}
P	-	-	-	-	-
V_a	0.76	-	-	-	-
$q_{s\ BTMA}$	0.70	0.48	-	-	-
H_{dune}	0.20	0.27	0.16	-	-
C_{dune}	0.22	0.24	0.36	0.22	-

We agree that the following sentence is confusing and not necessary. We wanted to show differences between the magnitude order of dune celerity and apparent bedload velocity, keeping in mind that the apparent bedload velocity participate to dune migration. Just to be clear, you mentioned that bedload velocity is smaller but this is the opposite, the bedload velocity is higher than dune celerity. We modified the text, keeping just this part of the sentence: "*It must be noted that apparent bedload velocity is higher than dune celerity with about a factor 100.*"

SC7

Lines 257-258: "On the other hand, the apparent bedload velocity is positively correlated with the acoustic power. The RMA regression model explains 76% of the dataset dispersion (Fig. 6a)."

Line 261 "and water discharge explain 71% of the dataset dispersion"

"What it meant by 'explains 76% of the dataset'? This needs to be accurately described. This paragraph needs to be re-written."
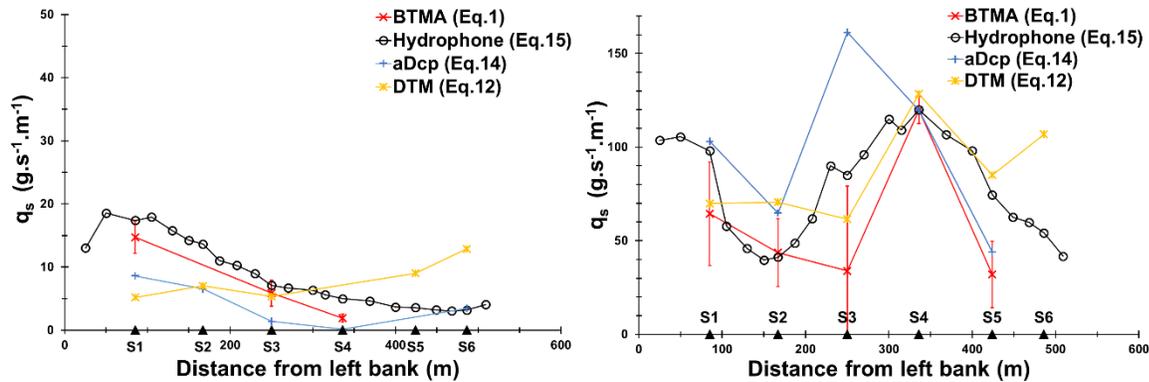
The COD (R^2) of the RMA regression is equal to 0.76 this mean that the equation of the RMA regression (model) determines 76% of point distribution (scattering). We modified these two sentences by: *"The COD of the RMA regression is equal to 0.76.*"; and "*The COD of the RMA regression established between BTMA bedload rates and water discharge is 0.71.*"

SC8

Line 286 "there was no reference measurements"

"Why not? Please explain. Can you please add labels, e.g. S1, to the BTMA data points in both panels of Fig. 8 to clearly identify the reference measurements? Why is there no S2 in Fig. 8a? Why are reference measurements missing from Fig. 8b?"

As mentioned in the material and method section, BTMA method is very time-consumming and it is not possible to measure all sampling points for each field survey. This is the case for S2, S5 and S6 on the figure 8a and also for S6 on the figure 8b. We added sampling points on figures 8a et 8b as requested:

BTMA (Eq.1)
Hydrophone (Eq.15)
aDcp (Eq.14)
DTM (Eq.12)

q_{ls} (g.s^{-1}.m^{-1})

Distance from left bank (m)

S1 S2 S3 S4 S5 S6

SC9

Line 294 I would like to see some discussion as to why the hydrophone method is producing larger values of unit bedload rate compared with the BTMA measurements. Just a thought, but could it be related to the omnidirectional hydrophone picking up higher noise magnitudes that are not directly below the boat?

A hydrophone senses every noises that are propagating in the water column. Therefore, the hydrophone can sense noises that are far away from its location. Noises are more and more attenuated with increasing distance (Geay et al., 2019). Particularly, when there is few bedload noises close to the hydrophone, the hydrophone can sense bedload noises that are generated far away. This behaviour could explain why the hydrophone tends to overestimate bedload fluxes when bedload fluxes are weak.

SC10

Lines 303-305 "For this survey, acoustic signals (i.e. acoustic power, apparent bedload velocity) followed the same evolution pattern as isokinetic samplers along the cross section except for S3."

"Again, why is this the case? Is it due to a lack of directionality?"

As the aDcp method, isokinetic sampler method is a "point/local" measurement which is function of the local grain size. The sediment grain size variability is important to explain the spatio-temporal variability in bedload sediment dynamics in large sandy gravel bed rivers. For instance, we made some measurements where one BTMA caught some sediment transported and the other one (located about 4 m away on the same cross section) was empty due to coarser grain size at this specific location. So, it is clear that an uncertainty is due to the limited spatial resolution of these two methods compared with hydrophone method that integer longitudinal variability of bedload (drifts are about 30 m long depending of flow velocity). Moreover, the presence of bedforms, specifically bars, increases the bedload (and grain size) variability (Venditti et al., 2012; Cordier et al., 2020).

"Line 325 Immediately downstream of the bar there are bedload transport values that are higher than those observed further downstream. Is this due to the omnidirectional hydrophone picking up bedload noise from the bar upstream?"

It is possible that the hydrophone senses the noise that are generated at the bar top and that acoustic power decreases smoothly when increasing the distance from the bar edge. The small increase of noise could be the noise of a dune that fall on the avalanche face (lee face) of the bar when the hydrophone was over the bar front, but we can only make assumptions here.

All other technical corrections mentioned by the referee were adopted in the final version of the paper.

References

Cordier, F., Tassi, P., Claude, N., Crosato. A., Rodrigues, S., and Pham Van Bang, D.: Bar pattern and sediment sorting in a channel contraction/expansion area: Application to the Loire River at Bréhémont (France). Advances in Water Resources, 140, https://doi.org/10.1016/j.advwatres.2020.103580, 2020.

Geay, T., Belleudy, P., Gervaise, C., Habersack, H., Aigner, J., Kreisler, A., Seitz, H., and Laronne, J. B.: Passive acoustic monitoring of bed load discharge in a large gravel bed river, J. Geophys. Res.: Earth Surf., 122 (2), https://doi.org/10.1002/2016JF004112, 2017.

Geay, T., Michel, L., Zanker, S., and Rigby, J. R.: Acoustic wave propagation in rivers: an experimental study. Earth Surface Dynamics, 7 (2), 537–548, https://doi.org/10.5194/esurf-7-537-2019, 2019.

Geay, T., Zanker, S., Misset, C., and Recking, A.: Passive Acoustic Measurement of Bedload Transport: Toward a Global Calibration Curve?: J. Geophys. Res.: Earth Surf., 125 (8), https://doi.org/10.1029/2019JF005242, 2020.

Merchant, N. D., Barton, T. R., Thompson, P. M., Pirotta, E., Dakin, D. T., and Dorocicz, J.: Spectral probability density as a tool for ambient noise analysis. The Journal of the Acoustical Society of America, 133, https://doi.org/10.1121/1.4794934, 2013.

Petrut, T., Geay, T., Gervaise, C., Belleudy, P., and Zanker, S.: Passive acoustic measurement of bedload grain size distribution using self-generated noise. Hydrol. Earth Syst. Sci., 22, 767-787, https://doi.org/10.5194/hess-22-767-2018, 2018.

Thorne, P. D.: Laboratory and marine measurements on the acoustic detection of sediment transport, J. Acoust. Soc. Am., 80(3), 899, https://doi.org/10.1121/1.393913, 1986.

Thorne, P. D.: An overview of underwater sound generated by interparticle collisions and its application to the measurements of coarse sediment bedload transport, Earth Surf. Dyn., 2 (2), 531–543, https://doi.org/10.5194/esurf-2-531-2014, 2014.

Venditti, J. G., Nelson, P. A., Minear, J. T., Wooster, J., and Dietrich, W. E.: Alternate bar response to sediment supply termination. Journal of Geophisical Research, 117, F02039, https://doi:10.1029/2011JF002254, 2012.

Anonymous referee 2

General comments (GC)

<u>GC1</u>

a)   *"In the text, there is no info about the standard deviation of the particle's sizes of the bedload material. I would also like to see the range of the particle Reynolds number (Rep), the shear velocity or stress and the critical shear velocity or stress. It could be also interesting if the authors give information about the reference shear stress at the study site (Parker et. al 1982). Note that this is similar to the critical shear stress but not the same."*

As expected by the referee, we added some details about sediment characteristics in the text.

"*It varies between 0.3 and 3.1 mm with a standard deviation of 0.4 mm. The D_{90} is much variable with a median value of 3.3 mm ranging from 0.5 to 15.7 mm. Hydraulic conditions at the sampling points varied with discharge; the median water depth and flow velocity are 1.6 m and 0.9 m.s^{-1}, respectively.*"

In order to limit the number of figures in the paper, we proposed to add the figures 1a and 1b (please see below) in the present author's response. Colleagues interested will find all the details and figures in the present document available on the website.

Hydraulic conditions and sediment characteristics allowed to compute the Shields stress (τ^*) and its critical value (τ_c^*) for 312 samples with the following equations (figure 1b):

$$\tau^* = \frac{D.S}{RD_{50}};$$

Where D is the water depth, S is the slope and R is the submerged specific gravity of the sediment (1.65).

$$\tau_c^* = 0.22\, Re_p^{-0.6} + 0.06\, e^{(-17.77 Re_p^{-0.6})};$$

Where R_{ep} is the particle Reynolds number:

$$Re_p = \frac{\sqrt{RgD_{50}}\, D_{50}}{\nu};$$

With *g*, the gravitational acceleration (9.81 m.s^{-2}) and ν the kinematic viscosity of water (10^{-6} m².s^{-1}). The Shields stress related to the D_{50} of these 312 samples was higher than the critical Shields stress for all hydraulic conditions measured (figure 1b). The sediment (rather fine) is characterized by a significant mobility for various discharge and flow velocity conditions.
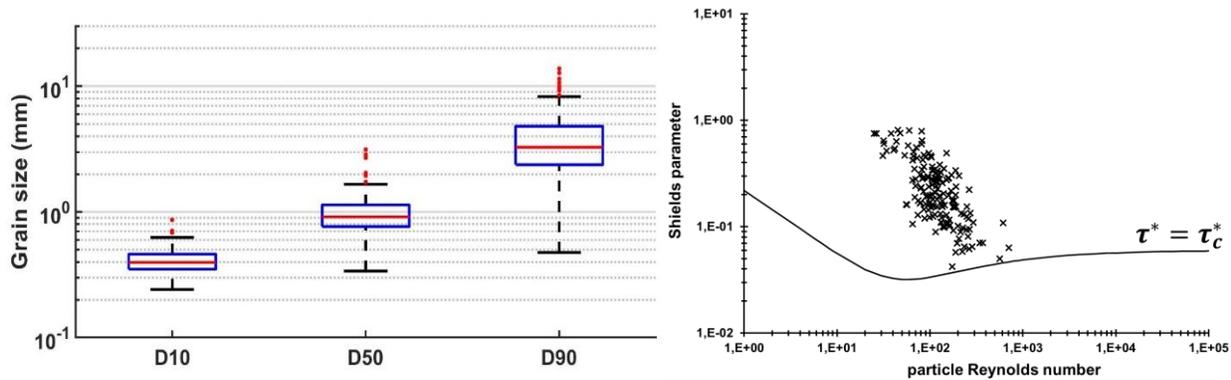
Figure 3: a) Distribution of grain size statistical parameters for n=447 samples. b) Shields parameter as a function of particle Reynolds number (D_{50}) with Shields curve of critical parameter.

Table 2: Distribution of characteristics sediment diameters for 447 samples and hydraulic conditions (water velocity V and water depth D) at sampling points.

	D_{10} (mm)	D_{50} (mm)	D_{90} (mm)	V (m.s^{-1})	D (m)	R_{ep}
Median	0.4	0.9	3.3	0.9	1.6	111.3
Minimum	0.2	0.3	0.5	0.2	0.4	25.1
Maximum	0.9	3.1	15.7	1.4	4.8	705.4
Standard deviation	0.1	0.4	2.3	0.2	0.9	86.0

b)   *"There are quite a few comments about the measurement capability of the instruments related to the different PSD of the bedload particles, but not details about the PSD of the sampled bedload surface."*

We presume that you refer to Particle Size Distribution/ Grain Size Distribution by mentioning PSD (and not to Power Spectral Density as in the paper). We are not sure to understand your comment. The GSD (grain size distribution) is mentioned in the *GC1a* point you raised.
Anyway, we did not investigate in this paper the specific potential link between GSD and the capability of the instruments. As a first approximation we've seen on some hydrophone records that for a given PSD the shape of the frequency distribution curve is different according to the GSD. This is a very interesting point we will investigate in another contribution.

c)   "I have another question about the boat/floating structure given on Fig2b: How was it fixed and how intensive was the drifting that you mention several times (in terms of meters)?"

The figure 2b of the paper illustrated the boat used for BTMA and aDcp sampling. As mentioned line 89, this 20 m long boat is fixed with 2 anchors, one in the front of the boat and 1 anchor at the back. The aDcp and its floating structure were fixed and tightened to the boat with a rope. There was no drifting of this structure during measurements (only small lateral movements of the boat were possible). We mentioned drifting only in the hydrophone protocol. Passive acoustic measurements were acquired with a smaller boat, turning off the engine and drifting at water velocity during more or less 30 seconds (about 50 m).

a)   "I was wondering how this calibration coefficient (α=2) was adapted. Does it take in consideration all the factors numbers above (check Banhold, et al., [2016] for detailed analysis of a very similar sampler in Germany)? It would be good to see a sentence regarding this issue. Calib coefficient of means that the samples underestimate 100 % of the real sediment transport."

b)   and d) "Based on which tests? I could not find the citation online…Boiten 2003? It looks like lecture notes, so inside it must be stated how these coefficients are defined and related to another study. And if some experiments exist, were those conducted in similar conditions as your study site?"

Boiten (2003) is a textbook describing lecture notes of Delf University (https://www.taylorfrancis.com/books/hydrometry-boiten/10.1201/9780203971093). We added the DOI in the reference list. Details of the calibration procedure are in two Dutch reports that I cannot find (Delft Hydraulics Laboratory, 1958 and 1969). But, some outlines are described in de Vries (1979). The first calibration tests consisted to compare the BTMA bedload catches with the average flume sediment transport. In a second time, tests were made by weighing catches in a sediment trap of the same size than sampler mouth. Sediment mixtures used during these tests was coarser than bedload of the Loire River with D50 varying between 2.5 mm and 6.4 mm. Both test series concluded to the same calibration coefficient of 2 (efficiency of 50%). This means that actual transport is two times higher than those measured with the sampler. The average efficiency of a basket sampler as the BTMA sampler is about 45% (Hubbell, 1964). These tests established a calibration curve linking unit bedload transport rates with caught volumes of sediments. It is mentioned in Boiten (2003) that this calibration factor was not including the possible losses of sediments finer than 0.3 mm (size of the mesh). The sand loss was estimated during flume experiments with a similar samples by Banhold et al. (2016) to 50% in average but with a mesh size of 1.4 mm and varied sediment mixture (D50=[0.8-10] mm). There is no detail about the hydraulic coefficient of the BTMA which is the ratio between velocity in the sampler and flow velocity in de Vries (1979). But, author argued that the BTMA construction promote the transport coefficient at the expense of hydraulic coefficient. In comparison, the Helley-Smith sampler tend to overestimate bedload with the ratio between the flume sediment transport and the sampler sediment transport that vary with water velocity from 1.2 to 2.6 (Helley and Smith, 1971). The BTMA was already tested on the Loire River (Gautier et al., 2008 ; Claude et al., 2012) and its results were compared with others estimation of bedload sediment rates.

The sampling error was estimated to 30% with 10 observations, 20% with 20 observations and 9% with 100 observations (Eijkelkamp, 2003). This stochastic uncertainty can be evaluated for each sample point with the standard deviation of unit bedload ($\sigma_{q_{s\,BTMA}}$) and the number of bedload samples (n) (Frings and Vollmer, 2017):

$$E_{q_{s\,BTMA}} = \sigma_{q_{s\,BTMA}}\ n^{-0.5} \ ;$$

The mean calculated value of stochastic uncertainty is about 30% of the mean bedload rates measured. It varies between 12% and 100% and increases when bedload transport decreases. The uncertainty given by the constructor is not far away from mean calculated value. To conclude, we are not able to describe with details how the calibration factor is fixed but we know that this calibration factor don't take into account the under-estimation of fine sands (less than 0.3 mm).

This sentence was added to the final version:

"*Suggested values of α and b were adopted from Boiten (2003) which mentioned that the trap efficiency factor not include the possible losses of sediment finer than 0.3 mm (mesh size opening). This calibration factor come from successive calibration procedures concluding to the same calibration coefficient of 2 (de Vries, 1979).*"

"*The uncertainty of unit bedload was calculated using the equation of the stochastic uncertainty from Frings and Vollmer (2017).*"

c)   "Is there a defined procedure for sampling? How many samples are sufficient to minimize the uncertainty (Frings, et al., 2017)?"

The defined procedure for sampling is described in Boiten (2003). We used 2 BTMAs to decrease the sampling time on each sampling points. At each sampling point 10 samples were collected with each BTMA (so 20 in total) and volumes of each samples were measured *in-situ* with a graduated cone (Imhoff cone). All samples volumes from each BTMA were merged for sieving analysis (leading to 2 sediment samples per sampling point; one for each BTMA). Then, the average volume of caught sediments from the 2 BTMAs was computed and converted into instantaneous unit bedload rates using equation 1. This equation is extracted from the calibration curve mentioned above and adapted to convert volume into mass of sediment.

We added some details of the procedure in the final version of the paper: "*At each sampling point, 10 samples were collected with each BTMA (so 20 in total) and volumes of each samples were measured in-situ with a graduated cone (Imhoff cone). Collected volumes were integrated over at least 2 minutes. All samples volumes from each BTMA were merged for sieving analysis (leading to 2 sediment samples per sampling point; one for each BTMA). Then, the average volume of caught sediments from the 2 BTMAs was computed and converted into instantaneous unit bedload rates using Eq. (1):*"

e)   "What are the conditions when the sampler would malfunction (e.g. weak transport conditions or too abundant)? There are many examples of low transport rates could be underestimated or overestimated by the physical samplers. Although most of the literature suggests that in weak transport conditions the samplers would over-estimate (e.g., due to digging in the bed), sometimes in sandy rivers, we have noticed many samples to under-sample the total transport (observing videos and total mass caught)."

We fully agree with your comment that in weak transport conditions the samplers tend to under-estimate bedload. Video records showed a very thin bedload layer for these conditions and the mouth of the sampler is not always well posed on the river bed, letting through particles underneath the mouth sampler. This occurs when sediments are coarser, when dunes are present and also immediately downstream a bar front where the riverbed (lee effect).During flood event, it is difficult to see the mouth's sampler on video records and adding a light was not a solution because there are lot of particles in saltation/suspension near the bed for these hydraulic conditions.

f)   g) "Given the Eq. 1 I was wondering if you have done PSD and weight/sieving analysis in the laboratory for each sample? If so, why not using only the dry mass? Besides the sieving, there is no information on how you have measured the other parameters, such as the volume V."

As mentioned in the response *GC2b*, we did sieving analysis on the total volume of caught sediments for each BTMA of each sampling point (2 grain size analyses per sampling point). Sampling volumes were merged (one sample for BTMA1, another for BTMA2 for each sampling point) so we are not able to use dry mass in laboratory. Bedload volumes are directly measured in the field with a graded cone.

h)   "In the appendices, you give information on how many samples are averaged in the given value. But there is no explanation of how the samples from the two BTMA are used in the processing. Please give information about that."

Please see our previous comment (*GC2b).*"

GC3

a)   "The post-processing procedure only partially follows some of the latest findings given in (Conevski, et al., 2019) and (Conevski, et al., 2020). Although the authors tried to implement some of the procedure there are some miss-steps. The de-spiking and the filtering procedure (discarding the raw BT velocities that are in the opposite direction of the flow) given in these two studies improved the correlation with the samples and gave more realistic velocities. It seems that the filtering procedure automatically excluded the negative velocities that occur in the so-called recirculation zone of the dunes (which you also mention). In addition, there is some sort of final filtering that involves analysis of the difference of the depth values registered by each beam, which helped to discard some samples that suffered from beam inhomogeneity (Conevski, et al., 2020)."

We are grateful for this very good idea and also for the reference. We agree that the filtering procedure mentioned in Conevski et al. (2020) would probably improve the results. Nevertheless, we did not have time to develop the code that could be used to process the dataset (initial submission of our paper: 21 of September 2020). We plan to test the filtering procedure detailed in Conveski et al. (2020) in a future paper.

b)   "The projection to the water direction (Wdir GPS) is not clear to me. How these two are defined do you use compass or GPS heading for Wdir and compass heading for Va? Why is this necessary? You say "To avoid compass and GPS issues, and to eliminate the effect of residual lateral displacement", but then you use again: Wdir GPS (GPS ref coordinate system)-bdir BT (compass heading EH). By doing this some of the issues given above are inherited again in the velocity. Also, it is not true that the bedload velocity should align with the water velocity especially in presence of bedforms. If one would like to project one velocity vector on another both should have then the same reference coordinate system or to contain information (rotation and /or translation) about the different reference coordinate system. As far as I understand you use GPS-ENU and Compass – ENU. Note that you are in a fixed position and you do not need water velocity referenced to the BT. Another thing is that if you project over the water velocity the real direction of the bedload is lost, and the final value decreased. Please correct me if I understood wrong."

You are right, there is a mistake here. We corrected $W_{dir\ GPS}$ by $W_{dir\ BT}$ in the equation and the results (figure 3a, plot and RMA equation). So all direction information come from the compass

heading. We agree that by doing this, we reduced the final bedload velocity magnitude. But, the idea was to get the flux velocity and not the bedload velocity magnitude. We are aware that this is an assumption to consider bedload in the same direction of the main flow direction mainly in presence of bar. Again, maybe the filtering procedure mentioned above would lead to better results than this method.

We added: "*This method involve the assumption that bedload is in the same direction than the main flow direction.*"

c)   "The concept of volume and surface scattering seems not to be considered in conceptualizing the discussions. Please check (Urick, 1983), also (Conevski, et al., 2019), characteristics (Conevski, et al., 2020, accepted). Note that the scattering of the surface formed by the immobile particles below the active layer has large impedance and therefore the echo coming from this scattering is dominant. The roughness of this surface could induce false velocity estimation. This error depends on the acoustic parameters of the instrument and the signal processing algorithm. Please have in mind that this is a complex two-phase (surface +volume) scattering process occurring between two scattering regions (Rayleigh + geometric), so it cannot be easily described."

d)   "In the same manner, the concept of internal processing and signal processing is not well addressed. Note that the RDI uses broadband codded pulse technology while M9 uses the pulse-coherent method with some internal modifications such as the use of HD smart pulse. Note that the RDIs BB is able to use much finer resolution systems (Brumley, et al., 1991) in the detection of the bottom signal, thus it should be able to define the better immobile bed surface. Note that the main purpose of the bottom tracking signal is to identify the riverbed and it is developed in that way (R. Lee Gordon, 1996). This is one of the reasons why RDI continuously reports lower velocities than M9, regardless of the frequency (Conevski, et al., 2020). This is also related to the acoustic sampling effect described in the same paper."

e)   "Although it is well known that the grain size is correlated with the backscattering strength even with the bedload particles (Conevski, et al., (2020, accepted); Shiel, 2010), the laboratory investigation (Conevski, et al., 2019; Conevski, et al., (2020, accepted); Conevski, et al., 2020) demonstrated that it is not the only reason for different Doppler / apparent bedload velocities especially in well-developed transport conditions."

We are grateful to all your above comments that would improve the discussion of the paper. We decided to add a sentence in the manuscript based on your comments (GC3 c, d, e):

"*The response of aDcp to bedload transport depends on several parameters The variation of the impulse frequency, the pulse length, beam focusing or associated internal signal processing (Broadband or Narrowband) can lead to different estimation of the apparent bedload velocity for the same sediment transport conditions (Conevski et al., 2020a). These parameters vary from a device to another (RDI/Sontek; Conevski et al., 2020b). As the aDcp pulse sample a volume of the riverbed (Rennie et al., 2002) which can lead to a biased estimation of V_a: 1) an underestimation in case of large roughness of the riverbed with most of the reflected pulse is scattered by the immobile particles below the active layer (Conevski et al., 2019); 2) an overestimation in case of high concentration of the bedload layer (Rennie et al., 2017) or sand particles became in suspension near to the riverbed (water bias, Rennie et Millar, 2004). Even if a general trend seems to be highlighted by the river comparison (figure 3a) with an increasing*

bedload rate as grain size increase for a constant V_a, the relationship between grain size and V_a cannot be easily described in response to all variables mentioned above. One explanation of this trend could be that suspended sands could contribute to the bottom tracking signal without being caught by the sampler (Rennie et al., 2017)."

f)   "In the results section, there is no distinction in the Figures between the data measured by the RDI 1200 kHz and the M9 (also which frequency). In Latosinski et .al. (2017) and Conevski et al. (2020) it is clearly indicated that different ADCPs give different mean apparent bedload velocities, and it is not strictly dependent on frequency, but also to the acoustic sampling and resolution, pulse length, cell-profiling, signal processing etc. The laboratory tests are given in Conevski et al. (2019) and Conevski, et al. (2020) also confirm and elaborate on these issues."

The scope of this paper is not to study the impact of frequency or pulse length on the apparent velocity. In that way, we did not mention these details in the figure 3a. But you can find details in the appendix B. About 55% of the dataset was measured with the RDI Rio Grande (1.2 kHz). You can find below the figure 3a adapted with the reference of the used device. It is clear that if we had been able to use the same device during all measurements it would be better.
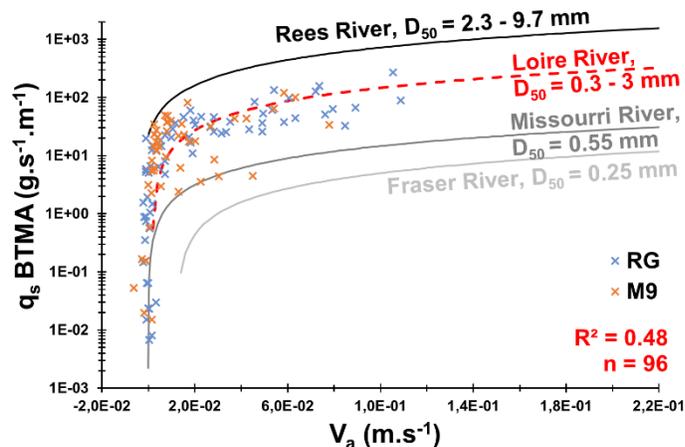
Figure 4: Adaptation of figure 3a of the original paper with point color in function of aDcp device (Rio Grande/M9).

g)   "The timing of 3-10 minutes is enough to get a stable apparent velocity at a single position (Conevski et al., 2019; Rennie, et al., 2002)."

When aDcp measurement is associated to BTMA sampling, the aDcp records during all BTMA sampling time. So, the timing mentioned in your comment is easily reached. Nevertheless, Rennie et al (2002) mentioned that the sampling duration have to be about 25 min to get a reliable estimation of apparent bedload velocity. Whereas Conevski et al. (2019) argued that 3-4 min was sufficient if no bedforms were present. We added a sentence in the manuscript as well as the references:

"*The sampling time needed to get a stable apparent velocity is not well defined and ranging between 3 min in case without bedforms (Conevski et al., 2019) and 25 min (Rennie et al., 2002). In our study the sampling time ranged between 5 and 190 minutes.*"

We agree that filtering procedure could eliminate beam heterogeneity due to bedforms. Here mean that the aDcp calibration differ from a river to another. Even if Conevski et al. (2020) have found high correlation coefficient between bedload rate and apparent bedload velocity for two rivers (Elbe River and Oder River), these two rivers are very similar in term of grain size characteristics, shear velocity, water depth or slope. As mentioned L. 381, there is no general calibration equation linking bedload rates and apparent bedload velocity. Every river have its own calibration curve and the strong dependence of V_a on bedload grain size lead to complicate calibration procedure in heterogeneous sediments (Rennie et al., 2017). We discarded this sentence in the discussion and added comment: "*Recent works have been carried out on two rivers (Elbe, Oder) very similar to the Loire River in term of grain size characteristics, flow and shear velocity, and water depth (Conevski et al., 2020a). Even if the correlation between apparent bedload velocity and bedload rates is significant, this calibration equation was obtained from two very similar rivers. Despite these observations, there is no general agreement between bedload rates and apparent velocity (Rennie and Villard, 2004; Rennie et al., 2017).*"

It is a very interesting comment but it is out beyond our competence domain. We believe that these investigations should be carried out experimentally (in a flume study). We are not able to provide an estimation of SSC because the paper mainly focuses on bedload.

This is a good idea but we think that it is difficult to perform these measurements in field conditions. In our case, the positioning of the aDcp precisely over dune stoss side is difficult because dunes are rather small. Moreover, the sampling points are defined before going on the field (GPS coordinates) to be sure to compare diachronically the same sampling points (to analyze the evolution of bedload for various discharge conditions). We understand that using van Rijn's approach implies that roughness is assumed to be $3D_{90}$. We could consider the effective roughness defined by the same author (van Rijn, 1984b) as the sum of the grain roughness and the form roughness in order to compute the Chézy coefficient and then the shear velocity related to form roughness. Unfortunately, we don't have bedform geometry for each BTMA/aDcp comparison. This makes impossible this computation but remains a very nice field of investigation.

It seems that the formulas you use for the critical shear velocity do not involve the hiding effect function. Note that you have sand and gravel in the bed mixtures, and therefore a fractional transport. This might not be crucial for only determining the active layer thickness but it is worth checking because the active layer thickness might increase due to the decrease of the critical shear velocity of the gravel particles, in a presence of sand (Wilcock, et al., 2001; Curran, et al., 2005).

As you mentioned, the shear velocity here is computed to assess the transport stage parameter (Van Rijn, 1984) and then the active layer thickness. Taking into account the hiding effect could be a good idea to see the influence of the composition of the riverbed on the active layer thickness, specifically in reaches characterized by coarse and poorly sorted sediments. Actually, we think that is not crucial in this paper because 1) the proportion of gravels in our sediment mixtures remains not sufficient to induce a hiding effect (16% of small gravels and 84% of sand in average) and 2) The size of the gravels remains quite fine (median value of $D_{90}=3.3$ mm).

Porosity.

Which value do you use for the porosity of bedload concentration in the Eq. 4? A constant bedload concentration of 0.65 is utterly false. Frozen-core bed surface or substrate samples were taken from riverbed have a porosity of 0.4, which simply means the concentration of 0.6. In the same context perfectly, packed particles without consistent shape hardly reach porosity less than 0.3. In the kinematic model, the concentration value is the value of the concentration in the instantaneous (dynamical) bedload active layer. This value shall be at a maximum lower than 0.6; Van Rijn (1984) reports this as maximum bedload concentration. Rennie and Villard (2004) reported 0.1-0.15 in their model. Note that this is not true for the porosity of the dunes; it is also stated in the paper that you make a distinction between the "dynamical active layer" (I am calling it instantaneous, or max saltation height) and the event -scale exchange layer (e.g., this includes the height of the dunes)

We used the constant value of bedload concentration (0.65), defined by van Rijn (1984) as a maximum. Of course, it varies according to hydraulic conditions. It also differs between bedforms (e.g. dunes and ripples; Simons et al., 1965). Here, our goal was not to determine the exact concentration or the thickness of the active layer precisely but to compare methods. So, we simplified the equation by considering concentration of the bed as a constant. However, we

decided to take into full consideration your suggestions relative to the bedload concentration and apparent velocity. Please see SC8 for full details.

GC4

a)   "The river morphology and the presence of bars (that actively migrate) suggest the dune formation throughout this cross-section would be 3-dimensional, which means that I would not expect straight line dune wavefront. Is it correct?"

Partially. Actually, the shape and angle of the dune front will mainly depend on flow depth, velocity and turbulence (and obviously the Froude number). The 3D dunes (sinuous crest) will develop mainly at high flow velocities with reduced water depth. The flow velocities on the site are relatively low (0.8 m.s^{-1} in average) due to the slope and width of the channel investigated. So, most of the bedforms present on the site are characterized by a relatively straight crest (2D dunes defined by Ashley, 1990). Nevertheless, 3D dunes can be rarely present close to obstacles for instance.

b)   "On the other hand, the DTM method that you use is built based on well-developed 2D dunes. Could you please elaborate on this issue? I am afraid that maybe only the dune measurements in the thalweg zone would be most reasonable to use in the comparison with the other two techniques?"

We understand your comment. On the Loire River the DTM method was used on several sites because the 2D dunes were recognized on all these sites (multibeam echosoundings). Please see the following publications: Claude et al., 2012 & 2014; Rodrigues et al., 2012 & 2015; Wintenberger et al., 2015. If dunes are present in the thalweg for all discharge conditions, they develop rapidly on the back of bars since these latter are submerged (Le Guern et al., 2019). That means that most of the riverbed is covered by dunes when bars are submerged. Although the significant width of the channel makes difficult to perform multibeam echosoudings, we did a couple of them, in 2018, on some longitudinal tracks located on the back of bars present in the study site (please see figure below). As it can be seen on this figure, the dunes are clearly 2D dunes for the 3 longitudinal tracks presented.
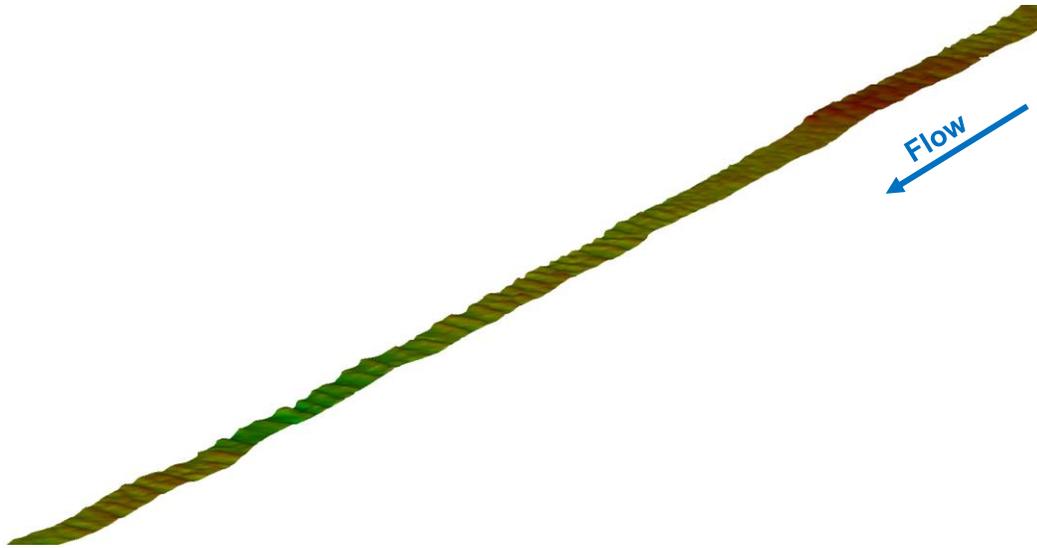
Figure 5: View of a multibeam longitudinal track measured on the study site (scale=about 300 m long to 4 m width).

c)   "Why not compare the bedload transport integrated over the entire cross-section from the hydrophones and ADCPs vs the dune tracking of only several longitudinal profiles (e.g., close to the thalweg)."

We guess that the elements given previously (see GC4b) show that this would probably lead to a bad estimation because the dunes present close to the talweg will have different morphological features (height, length) and celerity than the dunes present on other parts of the riverbed (stoss side of bars for example).

GC5

a)   "What is the advantage of drifting of the hydrophones? Isn't there less ambient noise (not bedload flux noise) if they stay fixed? Ok, I read in the Geay et al. (2020) that it is actually the opposite… I did not find a comparison of both cases. How was drifting performed and why the waves hitting the boat or produced by the boat did not affect the measurements?"

By moving at the water velocity, the ambient noise is reduced in comparison with a fixed position. The drift is performed as follow:

-    positioning of the boat upstream the sediment transport gauging cross section,

-    drifting of the boat at flow velocity by turning off the motor engine. The hydrophone record is stopped after the boat crosses the sediment transport gauging cross section.

The sound of waves and the boat are reduced when drifting. Among all our hydrophone records (448), we noted that, only one survey (performed during high wind conditions) was characterized by waves oriented upstream and increasing the ambient noise. However, these noises do not affect acoustic power measurements because they produce noise in the low frequency band (<2 kHz) and we integrated bedload noise in higher frequency band (15-350 kHz). We also checked

if the lower limit for integration of the PSD affects the overall acoustic power recorded (Artus and Bouchard, 2021). The results of this work is clear: we miss a very small part of the bedload signal by integrating at 15 kHz. Among all the measurements, only one sample is greater than 1% of the reference acoustic (Figure 2). As expected, this sample appears when we decrease the integration lower limit to 1 kHz and it correspond to the windy weather conditions that we mentioned above.
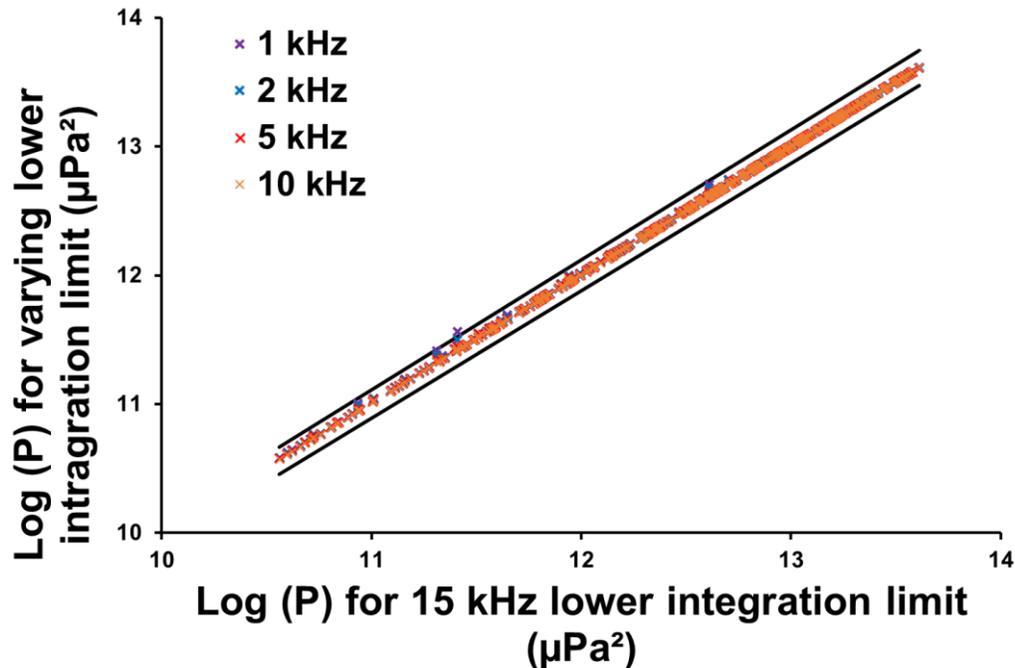

Figure 6: Correlation between acoustic power integrated above 15 kHz and above varying lower limits (1, 2, 5 and 10 kHz).

b)   "What was the ambient noise level at the presence of the bars? How was the acoustic power correlated with these morphological changes?"

Immediately downstream of bars, bedload sediment transport is mostly equal to 0. In consequence, the noise of bedload particle impact should be very small or null. We illustrated this phenomenon with figure 10a (L. 339). In this figure, acoustic power is converted into bedload transport rate. The figure shows that an acoustic power of 1.1014 µPa² characterize the crest of the bar while it is equal to1.1012 µPa² downstream of the bar front. The acoustic power is not null in this area because the hydrophone records noise from the stoss side of the bar (located upstream) where sediment transport is active. So, the hydrophone measured a decrease of acoustic power downstream the bar fronts but this decrease of noise is under-estimated because it integers some noise coming from upstream. This phenomenon is accentuated by the low noise level downstream the bar.

c)   "Was the signal sensitive to the dune's height change?"

We saw that the signal is sensitive to change in sediment transport induced by location on the dune (figure 10b): the signal increases towards dune crest, but we don't have enough data to

know if the hydrophone is sensitive to the dune's height change. Dune morphology changes with hydraulic conditions so we need data from different hydraulic conditions to explore this question. Actually, work is in progress on this specific point in our team.

d) "What was the bias of this signal towards the presence of gravel particles?"

Coarse particles produces noise at lower frequencies than fine particles. As a consequence, the Power Spectral Density of a coarser sediment mixture differs from a finer one. The median grain size is negatively correlated with central frequency of the spectrum (Thorne, 1986). As proposed by another referee, we decided to add a figure in the new version of the paper to illustrate this dependency.

Specific comments (SC)

SC1

"L21-L23: Please correct the English. Something like: Although the parameters that control the self-generated noise still require analysis…"

"Although further work is needed to identify the parameters controlling sediment self-generated noise …"

SC2

"L23-24: The sentence does not contribute in the abstract and it is not clear. Does it mean that they are not able to measure if bedform migration is not happening?"

It means that aDcp and hydrophone are both able to measure variation of sediment transport associated with the presence of dunes. "*Moreover, aDcp and hydrophone measurement techniques are sufficiently accurate to continuously measure bedload variations associated with dune migration.*"

SC3

"L 48: Are you talking about multi-beam?"

Here we refer to both single and multi-beam echo-soundings. We propose to add the following sentence: "*Moreover, these bathymetrical surveys are carried out simultaneously with sediment sampler measurements.*"

SC4

"L 79: "adopted here as reference". Maybe delete and start new sentence giving information that these data are used as reference data because this method so far is the most reliable… or something similar."

Agreed. We propose this sentence: "*Direct measurements of bedload sediment transport rates were performed using isokinetic samplers. This conventional approach was used to evaluate three indirect acoustic methods:*"

SC5

"L92-L100: See general comment 2. All sub comments."

Please see our response to these comments.

"L102: You say two ADCPs. On Fig 2 I see only one and, in the description, I understood that only one was available per campaign."

"L102 "positioned" could be replaced with installed."

You are right, there is a mistake in the previous version of the manuscript. There was only one aDcp mounted on the boat but aDcp changed from a campaign to another (Sontek, RDI).

"*Simultaneously with the BTMA measurements, an aDcp was installed on the boat (Fig. 2b). Measurements were performed using a Sontek Riversurveyor M9 (bi-frequency, 1 and 3 MHz) or a Teledyne RD Instruments Rio Grande (1.2 MHz).*"

"L103: Which mode (IC-incoherent or SmartPulse HD?) was used in each case. Please add in the appendices (see, general comment 3 d.)."

RDI aDcp uses coherent pulse system (Broadband) whereas Sontek aDcp use SmartPulse HD system that switches between Broadband and Narrowband system to provide the best data resolution. Most of the time, we used the Sontek device in manual mode fixing the pulse system to incoherent mode. We added the information in appendices B (P.21).

"L124-L130: These are both kinematic models and essentially both describe a unit volume with its specific density moving with unit velocity. The first one is an assumption that the ADCPs are more sensitive to the particles and particle numbers and that by assuming characteristic particle Va would represent average in all 4 acoustic footprints. Considering g.c. 3c it is rather strange that you get better matching for Eq4 in Fig3b. But then the assumptions for Eq. 5 and the filtering are not correctly applied."

We applied this equation based on a recent literature review (Holmes, 2010; Latosinski et al., 2017; Villard et al., 2005). As a starting point, we chose to fix the bedload concentration to the maximum bedload concentration according to van Rijn (1984) to compare to the computations performed for BTMA and DTM bedload rates. The thickness of the active layer was defined by van Rijn's (1984) equation as the saltation height. There was a mistake in the previous version of the manuscript on equation 6 (d_s=0.3 $D_*^{0.7}$ $T^{0.5}$ D_{50}). Fortunately this mistake was only in the text and not in the dataset. According to yours seminal comments (GC3j, SC19, SC20, SC33, SC35) we decided to analyze the effect of the apparent bedload velocity projection on kinematic models (Eq. 4 and Eq. 5). As expected, the concentration of bedload layer is probably overestimated when it is fixed to maximum bedload concentration ($c_0 = 0.65$), so, following your suggestion, we computed bedload concentration according van Rijn (1984):

$$c_b = 0.18 \frac{T}{D_*} \, c_0;$$

Computed bedload layer concentration varies between 0.005 and 0.1 (0.03 in average). Bedload layer thickness (d_s) was computed with Eq. (6) and varies between $1D_{50}$ and $7D_{50}$ ($5D_{50}$ in average). In consequence Eq. (5) underestimates bedload rates. If we consider that van Rijn equations allow a good estimation of c_b and d_s, V_a seems to be underestimated and bedload rates computed using Eq. (5) underestimate BTMA bedload rates (24% of the dataset in the discrepancy ratio). By considering apparent bedload velocity without projection over flow direction, the kinematic model is able to better estimate BTMA bedload rates (41% of the dataset in the discrepancy ratio). We already seen in the paper that projected V_a allowed a good estimation of bedload rates using Eq. (4). By using raw apparent bedload velocity, only 33% of the dataset are in the discrepancy ratio against 54% with projected V_a. So, as explained in the comment GC3b, the projection of apparent bedload velocity on flow direction allowed to filter incoherent values but it reduced the velocity magnitude. Consequently, Eq. (5) underestimates bedload rates and the use of raw apparent bedload velocity with a higher magnitude allowed to better estimate BTMA transport rates. Although, the application domain of Eq. (4) does not correspond to the characteristics of the Loire River, the decrease of projected V_a seems to compensate the overestimation of bedload rates when the raw apparent bedload velocity is used.
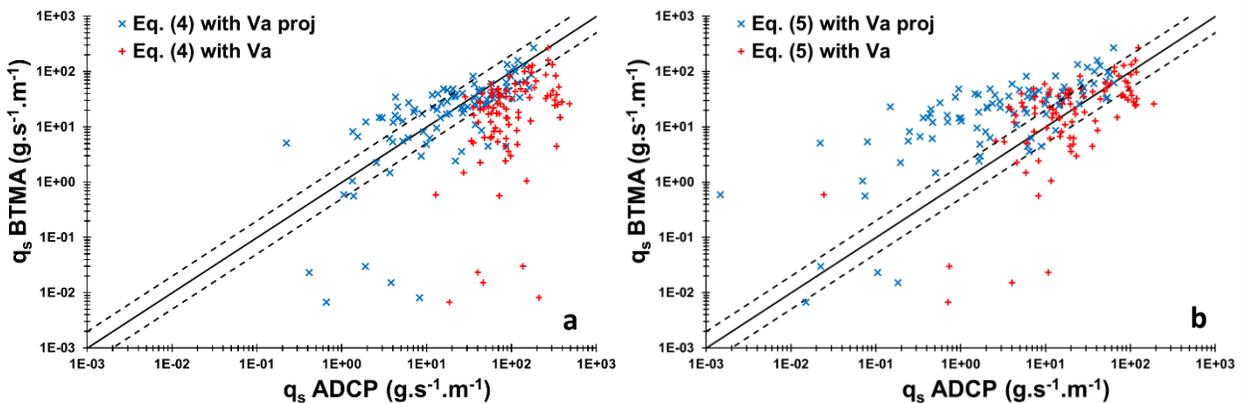

Figure 7: sensibility of V_a post-processing to estimate BTMA bedload rates with both kinematic models employed in the paper, a) Equation 4; b) Equation 5.

These results bring us to consider these two options of apparent bedload calculation in both kinematic model. Therefore, we could discuss the availability of our data processing and its consequence for using in kinematic models. The calibration between BTMA and aDcp with raw apparent bedload velocity is not possible and data need a filtering procedure as you proposed in GC3a. We propose to add:

- Equation of bedload concentration with appropriate modification in the Materials and Methods section (paragraph 3.2).
- Figures5a and 5b in the Results section 4.1. with the description of these results. "*Computed bedload layer concentration (Eq. 7) varies between 0.005 and 0.1 (0.03 in average). Bedload layer thickness (d_s) (Eq. 6) ranges between $1D_{50}$ and $7D_{50}$ ($5D_{50}$ in average). Bedload rates computed using Eq. (5) underestimate BTMA bedload rates with only 24% of the dataset in the discrepancy ratio (Figure 4b). By considering apparent bedload velocity without projection over flow direction, the kinematic model (Eq. 5) better estimates BTMA bedload rates with 41% of the dataset in the discrepancy ratio. Conversely, using raw apparent bedload velocity in the Equation 4, leads to only 33% of*

the dataset in the discrepancy ratio against 54% with projected V_a. According to these results, the Equation 4 better describes the sampler bedload rates with projected apparent bedload velocity whereas raw apparent bedload velocity are preferred with the Equation 5."

- The discussion of these new results in the paragraph 5.1. "*The results shown in Fig. 4a suggest that the equation 4 estimates sampler bedload rates if the projected bedload velocity is used. This kinematic model doesn't take into account the thickness or the concentration of the bedload layer and assumes that bedload transport never exceeds the size of a single particle assessed as uniform in terms of grain size (Rennie et al., 2002). These assumptions seems not appropriated for a sandy-gravel bed river. The active layer thickness should increase as suspended bed material load increase. Nevertheless, results are in agreement with BTMA bedload rates (Figure 4a). This can be explained by an underestimation of the apparent bedload velocity when it is projected along flow direction. On the other hand, Van Rijn (1984) defined the bedload layer thickness equal to the saltation height. The computed values of bedload layer thickness are coherent with other estimations made on comparable rivers (Conevski et al., 2020a). The equation 5 better estimates sampler bedload rates using the raw bedload velocity (Figure 4b). If we consider that c_b and d_s are well estimated by van Rijn equations, these results confirm that the projection of the apparent bedload velocity decreases the bedload velocity magnitude when the bedload direction differs from flow direction (e.g. bed slope effects). The influence of bedload velocity projection appear to be important when bedload are computed using kinematic models. Nevertheless, the calibration curve established seems to be in line with other studies. Although, the application domain of equation 4 does not correspond to the conditions of the Loire River, the decrease of projected V_a seems to compensate the overestimation of bedload rates when the raw apparent bedload velocity is used. This is the opposite for Eq. 5 that takes into account bedload layer thickness and concentration. In this case, the projection of V_a leads to underestimate bedload rates. Further works need to be done to improve the post-processing of V_a by recently published filtering procedures (Conevski et al., 2019 and 2020a) and to estimate its effect on calibration curve and kinematic models.*"

SC9

"L143-L150: Did you get better results in these surveys, when ADCP was close to the bed?"

In this configuration, we did not compare aDcp measurements with BTMA sampled volumes. This protocol was used to investigate the capability of aDcp to describe bedload variation over dunes. We did not compare both configuration of the aDcp. It could be interesting to do this work by measuring with immerged aDcp and a floating aDcp to see if the protocol allow a better description of bedload fluxes by reducing the footprint. We supposed that this is the case, especially with the small dunes of the Loire River (footprint vs. dune size).

SC10

"L150: Large window like that would do simple de-spike low pass filter. Not the noise but the outliers."

That is correct. We changed "*noise*" by "*outliers*".

"L151: All negative values refer to the average values, if they were negative you discard from the dataset? Or you applied some kind of filtering on the raw data as given in g.c. 3a? If so please cite Conevski et al 2019."

We removed negatives values from the dataset after the data processing. We did not use any filtering on the raw data but, as mentioned above, it could be very interesting to test it in further works. We added your reference in the conclusion.

SC12

"L170: So why did you adopt 0.5?"

We mentioned line 182 that the discharge coefficient is equal to 0.5 for a perfect triangular dune shape. We tested two options according to our literature review (0.33 and 0.57).

SC13

"L171: The accuracy of the echo sounding shall be very high. The problem is in the post-processing and identification in the bedforms, isn't it? See g.c 4a.

The accuracy of the echosounder is good but when you combine the accuracy of the DGPS, the echosounder and the navigation, the overall accuracy of the echosounder decreases. If you consider the RTK GPS accuracy and beam overture (6°) for the echosounder, it could be several centimeters. Based on our experience of the Loire River and on the papers above-mentioned, we chose 10 cm as a minimum dune height. It is nearly impossible to follow lower dunes especially during the celerity calculation post-processing. We modified the sentence by: "*Considering the accuracy of the bathymetrical echosounding relative to the dune size, the sinuosity of dune crests, and the representativeness of dune celerity, only profiles with a mean dune height of 0.1 m and more than 10 dunes are considered*".

SC14

"L174: What are the dimensions of the hydrophone?"

The hydrophone is 27 cm long and have a diameter of 3.8 cm.

SC15

"L186: I assume the sampling was not synchronized?"

Hydrophone drifts and BTMA sampling are not synchronized. We added this information to the paragraph 3.4.

"L196: Is it the mean of all STDs?"

Yes, "*mean value of 33 g.s^{-1}.m^{-1}*" is the mean value of all standard deviation.

"L198: Again. Which of the two BTMA samplers were used in the average value?"

Both BTMA samplers were used.

"L203-204: I am not sure if this comparison makes sense. To me, it looks that your RMA reg model could fairly represent the entire dataset (Rennie + yours, in the range of R2~0.5). Would be interesting to see. Check g.c. 3e."

By referring to your comment GC3e we agree that the variation of the apparent bedload is not only due to sediment grain size. We added a more general comment to this figure: "*The RMA regression presented here describes fairly well the dataset already published on several world large rivers*".

But, Rennie et al. (2017) showed this trend of increasing bedload rate with increasing sediment size (decreasing percentage of sand in the mixture).

"L213: Why did not you calculate for each point separately? What does constant means? For all data? If this is the case then it makes sense eq 4 to perform better."

We discarded this sentence because actually, as you mentioned in a specific comment above (*SC8),* the first kinematic model doesn't take into account the bedload thickness and concentration so it is not true to consider these parameters as constant. In the second model (Eq. 5), we initially considered bedload concentration as a constant (0.65) and bedload thickness as a variable (Eq. 6). So, only this last parameter was computed for each point. Please see also SC8 for related response.

"L219-220: Could you clarify this sentence. Do you mean Eq5 is better than Eq4 or the opposite? Eq.4 does not assume any active layer thickness, but average particle size."

We mean that, according to our data, Eq. 4 better estimates bedload rates obtained using BTMA than Eq. 5. Rennie et al. (2002) mentioned that this model assumes that the bedload transport never exceeds the size of one particle. We thought that considering the active layer thickness proportional to the median sediment grain size reflected this assumption.

$$q_s ADCP = \frac{4}{3} \frac{D_{50}}{2} \ \rho_s V_{a\ proj}; \tag{4}$$

$$q_s ADCP = 0.66 \ D_{50} \ \rho_s V_{a\ proj}; \tag{4}$$

And by comparing with Eq. (5) 0.66 could be assimilated to c_b and D_{50} to d_s

$$q_s ADCP = \ c_b \ d_s \rho_s V_{a\ proj}; \tag{5}$$

That is why we made the assumption of bedload layer thickness proportional to median grain size in the Eq. (4).

SC21

"Do you have something to say about the points down in the middle of the Fig3b?"

These points are related to very weak transport conditions reached at low flows (Q<300 $m^3.s^{-1}$). It could be due to an underestimation of bedload with samplers as discussed in the general comment GC2e. We added this sentence to the paragraph 4.1: "*Some outlier data are observed for BTMA bedload discharge lower than 0.1 $g.s^{-1}.m^{-1}$. These points correspond to low flow conditions for which bedload samplers could under-estimate bedload fluxes (gap between the sampler mouth and the riverbed more significant)*".

SC22

"L221-230: But that is reasonable. See g.c. 5a-b"

OK. See GC5a and b.

SC23

"L243: Does it involve the bedload concentration or just the velocity?"

The transport stage parameter is calculated using equations mentioned in 3.2. Equations 7, 8, 9, 10 and 11 do not involve bedload concentration but mean flow velocity, shear velocity, grain size and sediment density ratio.

SC24

We modified this paragraph accordingly to comments addressed by other referees. We added your comment that makes the sentence clearer. "*The apparent velocity measures the velocity top layer or dynamical active layer (sediment transported over dune), whereas the dune celerity is the mobility of the exchange event active layer, according to Church and Haschenburger (2017)*".

SC25

Here we explained that it did not work in this specific case because of the presence of bars and the difference of integration transport scales between the methods. But, even in the presence of bars, dunes are well-developed.

SC26

Only data with a corresponding BTMA measurement were used in the calibration of equations 14 and 15. So, aDcp measurements of figure 8 are all used to calibrate equation 14, whereas only few points of the total hydrophone dataset were used to calibrate equation 15.

SC27

Exactly. We changed that term according to your comment.

SC28

We used Eq. 14 and Eq.15 to calculate transport rates from raw apparent bedload velocity and raw acoustic power respectively. Even if these equations were calibrated with long averages values, the goal here was to check if aDcp and hydrophone were able to record instantaneous bedload variations that would be consistent with bedform migration. We did not use raw values because we decided to keep the work done in the above paragraphs. We thought it would make the article more consistent. Do you think that calibration with instantaneous values is possible?

No, we only suggest that the hydrophone is sensitive to the bedload variation induced by the front of the bar in a similar way to a dune crest (lee effect) but at a different scale.

SC30

"L352: "bedload transport occurs over a very low thickness", please reformulate to: in weak bedload transport conditions the BTMA sampler most likely performed with reduced efficiency … and cite literature."

"*Moreover, in weak bedload transport conditions, the BTMA sampler most likely performed with reduced efficiency initially calibrated to 50% (van Rijn and Gaweesh, 1992; Gaweesh and van Rijn, 1994; Banhold et al., 2016)*."

SC31

"L378-384: This is very superficial information given in a way to suppress the ADCPs under the hydrophones. And it is only partially true. Please correct according to the general comments 3 and cite accordingly."

Our idea was not to suppress the aDcp under the hydrophones. To our knowledge, to date there is no general equation available to compute bedload transport from apparent bedload velocity estimated on different rivers. We just wanted to highlight that, even if the results obtained from hydrophone calibration have to be taken with caution, it seems that a general tendency is emerging from Geay et al. (2020) and our results. Even if investigations were made in different rivers and using different hydrophones. Nevertheless we believe that the elements you gave in GC3 could seriously improve results given by the aDcp.

SC32

"L386: After the filtering (Conevski, et al., 2019) negative values should not exist. But if there is beam heterogeneity this adcp value is rather incorrect and shell be eliminated not used as null. Also if somehow all the beams are sampling the dunes trough (lee side) the measurement is also incorrect if one aims to estimate the bedload transport over a cross-section."

We agree that filtering could improve these results, especially instantaneous apparent bedload velocity. Here, we computed bedload with the calibration equation for all dataset even those which have been made without associated BTMA sampling. Where negative values were observed, we interpolated values from surrounding sampling points to compute the total bedload of the section. In consequence, we agree that our explanation cannot be use anymore.

SC33

As mentioned above, Rennie et al. (2002) argued that this model assumes the transport never exceeds the size of a single particle. We referred to low flow conditions where we observed a kind of very thin bedload layer (about 1 particle in height). We are not able to measure the bedload thickness accurately from our videos in field condition. Nevertheless, we agree that considering the bedload thickness of a single particle size can be an underestimation of the actual thickness. This is why we try to discuss this model for high flow conditions.

SC34

We agree that saltation (suspended bed material load) is part of the bedload. We understand that the text was unclear on this specific point and decided to modify according to your suggestion: "*The active layer thickness should increase with discharge that also increases suspended bed material load*".

SC35

Please see GC8 for reformulation of this part of the discussion.

SC36

We agree. At line 448 we added: "*Moreover, physical samplers sample the dynamical active layer, thus more comparable to the hydrophones and aDcps*".

SC37

This is the effective mean time needed to do the different protocols. For the aDcp method, we estimate this time for a survey without comparison with BTMA. During these surveys, aDcp was mounted on a small boat which is easier to handle and measurements were made on 6 sampling points. The time mentioned for aDcp method take into account the time needed to position and anchor the boat at each sampling point (comment added). We based the sampling duration from literature available at the start of the study (Rennie et al., 2002; 25 minutes) and adjust to 10 minutes considering the variability of the sampling time in several studies summarized by Rennie et al. (2017). Same estimations were made for others methods.

SC38

"L440: Could you elaborate on how this is not valid for the hydrophones if they are in a fixed position? In the same way, by use of the RTK one could move the ADCPs and estimate the BT velocity, or just subtract the boat velocity from the BT velocity measured by another device or the boat itself. Or if it is certain that one dune has passed below the ADCPs then using the filtering (Conevski, et al., 2019) the data shell be fairly accurate."

The hydrophone was not used in a fixed position (please see our comments above). We made drifts of varying distance (depending of flow velocity) in order to reduce the ambient noise generated by water. These drifts integrate several dunes. As mentioned above we will use the filtering method proposed by Conevski et al. (2019) in our future works to improve these measurements (not possible during this study).

SC39

"L452-456: Could you please reformulate. Do not really understand how the hydrophones are sensing the bars. I assume you mean change of the dunes which is just changing of the instantaneous transport and the shape of the dune… It could be another structure, not necessarily a bar… Is this correct?"

Yes. We wanted to underline the role of the lee effect exerted by bars on dune development downstream of a bar front (Reesink et al., 2014). We simplified the text:

"*Moreover, as for the aDcp, the hydrophone also detects the theoretical pattern of bedload transport rates associated with bedform migration. As shown by Reesink et al. (2014), the lee effect generated by bar fronts influences the development of dunes downstream. Specifically, the hydrophone is able to record the decrease of the acoustic power immediately downstream the bar front and its progressive increase (traduced by the development of dunes at about 11h06, Fig. 10a). In the present study, dunes smaller than 0.4 m*"

SC40

"L462: This is also partially true. The accuracy is within this range, but this stands for moving boat measurements. The laboratory data suggest lower values until 1mm/s check (Conevski, et al., 2020)."

In Rennie et al. (2017) this accuracy is extracted from fixed field measurements. We acknowledge that in controlled conditions such as in laboratory studies, this precision can be lower. In our case, the scattering of the apparent bedload velocity for velocity lower than 1 cm.s^{-1} (figure 3a) allowed us to argue in this sense. But, filtering procedure that you mentioned could improve the accuracy by discarding points for this range of apparent velocity.

SC41

"L469-L471: Please correct according to the general comments for the ADCP. If strongly correlated with the grain size, please show the correlation."

The term strongly correlated is maybe too strong, but regarding figure 3a, for the same apparent bedload velocity, the bedload rate is higher when D$_{50}$ increase. We changed "*strongly correlated*" by "*dependent to*".

References

Ashley, G. M.: Classification of large-scale subaqueous bedforms: a new look at an old problem, Journal of Sedimentary Petrology, 60, pp. 160-172, https://doi.org/10.2110/jsr.60.160, 1990.

Banhold, K., Schüttrumpf, H., Hillebrand, G. and Frings, R.: Underestimation of sand loads during bed-load measurements- a laboratory examination, in: Proceedings of the international conference on Fluvial Hydraulics (River Flow 2016), 11-14 July 2016, Saint Louis, USA, 2406 pp., 2016.

Boiten, W.: Hydrometry, IHE Delft Lecture Note Series, A.A. Balkema Publishers, Netherland, 256 pp, https://doi.org/10.1201/9780203971093, 2003.

Church (2019

Claude, N., Rodrigues, S., Bustillo, V., Bréhéret, J. G., Macaire, J. 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, https://doi.org/10.1016/j.geomorph.2012.07.030, 2012.

Claude, N., Rodrigues, S., Bustillo, V., Bréhéret, J. G., Tassi, P., and Jugé, P.: Interactions between flow structure and morphodynamic of bars in a channel expansion/contraction, Loire River, France, Water Resour. Res., 50, https://doi.org/10.1002/2013WR015182, 2014.

Conevski, S., Guerrero, M., Ruther N.,and Rennie, C. D.: Laboratory investigation of apparent bedload velocity measured by ADCPs under different transport conditions, J. Hydraul. Eng., 145 (11), https://doi.org/10.1061/(ASCE)HY.1943-7900.0001632, 2019.

Conevski, S., Guerrero, M., Winterscheid, A., Rennie, C. D., and Ruther N.: Acoustic sampling effects on bedload quantification using acoustic Doppler current profilers, Journal of Hydraulic Research, https://doi.org/10.1080/00221686.2019.1703047, 2020a.

Conevski, S., Guerrero, M., Rennie, C. D., and Ruther, N.: Towards an evaluation of bedload transport characteristics by using Doppler and backscatter outputs from ADCPs, Journal of Hydraulic Research, https://doi.org/10.1080/00221686.2020.1818311, 2020b.

de Vries, M.: Information on the Arnhem Sampler (BTMA), Internal Report n°3-79, Delft University of Technology, Department of Civil Engineering, Fluid Mechanics Group,1979.

Eijkelkamp: Operating instructions: Bedload Transport Meter Arnhem, Giesbeek, Netherland, 8 pp., 2003.

Frings, R. M., and Vollmer, S.: Guidelines for sampling bed-load transport with minimum uncertainty, Sedimentology, 64 (6), 1630-1645, https://doi.org/10.1111/sed.12366, 2017.

Gautier, J., Rodrigues, S., Peters, J., Peeters, P. and Jugé, P.: Transport solide en Loire moyenne lors des crues. Justification des mesures in situ et quantification. La Houille Blanche, 5:71–78, 2008.

Gaweesh, M. T. K., and van Rijn, L. C.: Bed-load sampling in sand-bed rivers, J. Hydraul. Eng., 120 (12), 1364-1384, https://doi.org/10.1061/(ASCE)0733-9429(1994)120:12(1364), 1994.

Geay, T., Zanker, S., Misset, C., and Recking, A.: Passive Acoustic Measurement of Bedload Transport: Toward a Global Calibration Curve?, J. Geophys. Res.: Earth Surf., 125 (8), https://doi.org/10.1029/2019JF005242, 2020.

Helley, E., J., and Smith, W.: Development and calibration of a pressure-difference bedload sampler, U.S. Geological Survey Open-File Report, 1971.

Holmes, R. R. Jr.: Measurement of Bedload Transport in Sand-Bed Rivers: A Look at Two Indirect Sampling Methods, U.S. Geological Survey Scientific Investigations Report, 2010-5091, 2010.

Latosinski, F. G., Szupiany, R. N., Guerrero, M., Amsler, M. L., and Vionnet, C.: The ADCP's bottom track capability for bedload prediction: Evidence on method reliability from sandy river applications, Flow Measurement and Instrumentation, 54, 124-135, https://doi.org/10.1016/j.flowmeasinst.2017.01.005, 2017.

Le Guern, J., Rodrigues, S., Tassi, P., Jugé, P., Handfus, T., Duperray, A., and Berrault, P.: Influence of migrating bars on dune geometry, in: Book of Abstracts of the 6th Marine and River Dune Dynamics conference, 1-3 April 2019, Bremen, Germany, 157-160, 2019.

Reesink, A. J. H., Parsons, D. R., and Thomas, R. E.: Sediment transport and bedform development in the lee of bars: Evidence from fixed- and partially-fixed bed experiments, in: Proceeding of River Flow 2014 - seventh International Conference on Fluvial Hydraulics, Lausanne, Switzerland, 3-5 Septembre 2014, 8 pp., 2014.

Rennie, C. D., Millar, R. G., and Church, M. A.: Measurement of Bed Load Velocity using an Acoustic Doppler Current Profiler, J. Hydraul. Eng., 128 (5), 473-483, https://doi.org/10.1061/(ASCE)0733-9429(2002)128:5(473), 2002.

Rennie, C. D., and Millar, R. G.: Measurement of the spatial distribution of fluvial bedload transport velocity in both sand and gravel, Earth Surf. Process. Landforms, 29, 1173-1193, https://10.1002/esp.1074, 2004.

Rennie, C. D., Vericat, D., Williams, R. D., Brasington, J., and Hicks, M.: Calibration of acoustic doppler current profiler apparent bedload velocity to bedload transport rate, in: Gravel-Bed Rivers: Processes and Disasters, Oxford, UK: Wiley Blackwell, 209–233, https://doi.org/10.1002/9781118971437.ch8, 2017.

Rodrigues, S., Claude, N., Jugé, P., and Bréhéret, J. G.: An opportunity to connect the morphodyamics of alternate bars with their sedimentary products, Earth Surf. Process. Landforms, 37, 240-248, https://doi.org/10.1002/esp.2255, 2012.

Rodrigues, S., Mosselman, E., Claude, N., Wintenberger, C. L., and Jugé, P.: Alternate bars in a sandy gravel bed river: generation, migration and interactions with superimposed dunes, Earth Surf. Process. Landforms, 40 (5), 610-628, https://doi.org/10.1002/esp.3657, 2015.

Simons, D. B., Richardson, E. V., and Nordin, C. F. Jr.: Bedload Equation for Ripples and Dunes, U.S. Geol. Survey Prof. Paper, 462-H, https://doi.org/10.3133/pp462H, 1965.

Thorne, P. D.: Laboratory and marine measurements on the acoustic detection of sediment transport, J. Acoust. Soc. Am., 80(3), 899, https://doi.org/10.1121/1.393913, 1986.

Van Rijn, L. C.: Sediment Transport. Part I: Bed Load Transport, J. Hydraul. Eng., 110, 1431-1456, https://doi.org/10.1061/(ASCE)0733-9429(1984)110:10(1431), 1984.

Van Rijn, L. C., and Gaweesh, M. T. K.: New Total Sediment-Load Sampler, J. Hydraul. Eng., 118 (12), 1686–1691. https://10.1061/(ASCE)0733-9429(1992)118:12(1686), 1992.

Villard, P., Church, M., and Kostaschuk, R.: Estimating bedload in sand-bed channels using bottom tracking from an acoustic Doppler profiler, Spec. Publs int. Ass. Sediment, 35, 197-209, https://doi.org/10.1002/9781444304350.ch12, 2005.

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, https://doi.org/10.1016/j.geomorph.2015.07.032, 2015.

Referee 3: Paul Grams

General comments

"I think the conclusions regarding dune tracking are more specific to how the method was applied in this study and to this field site. Thus, I don't think that general conclusions about the accuracy or utility of that method can be claimed."

You are right, DTM results are more specific to the study site especially to the dune scale that can vary with the presence of bars (Le Guern et al., 2019a). The application of the DTM was employed to have a second method (usually used in sandy-gravel bed rivers) to compare with aDcp and hydrophone measurements. Results would be improved if the study was centered around this method but to compare all these methods it is very time consuming in the field and protocols have to be adapted.

"The authors state that acoustical power was not affected by water depth, but no data are presented to demonstrate that conclusion. Because acoustical power is expected to be affected by water depth. That should be demonstrated, rather than asserted."

Please see Specific comment SC12.

Specific comments (SC)

<u>SC1</u>

"L. 52-54: Poor sentence construction. Rewrite. Likely needs two sentences."

"*In this work, we compare the efficiency of active and passive acoustic techniques to quantify bedload transport. The investigation took place in a reach of the Loire River (France) characterized by the presence of migrating bars and superimposed dunes (Le Guern et al., 2019b).*"

<u>SC2</u>

"L. 57: What are "bedload axes"? Need to define or use more descriptive term."

We agree that this term was unclear. We modified the sentence by: "*to estimate the accuracy of acoustic methods to measure cross sectional variations of bedload fluxes for various discharge conditions*".

<u>SC3</u>

"L. 58-59: Not clear how this is different from #1"

The third point aims at exploring bedload variations associated to bedforms with an aDcp and a hydrophone. Are they sufficiently accurate to record these bedload variations? The first point is related to the calibration of several acoustic methods with a direct measurement method without linking bedload to the presence of bedforms. We modified this sentence in order to be clearer: "*3) to investigate the capabilities of hydrophones and aDcps for capturing bedload variations along bedforms.*"

<u>SC4</u>

"L. 76: Why "theoretical" hydrophones?"

This is theoretical position of hydrophone drifts because when we drift we were not able to follow exactly these theoretical lines. However, we agree that this term can be confusing. We decided to delete it.

<u>SC5</u>

"L. 88: need full description of sampler (dimensions, weight, bag size, mesh size, etc.). Or reference to where that description can be found."

"*Samplers consists of a sampling basket mounted on a frame. The sampling basket have a rectangular mouth of 0.05 m high and 0.085 m wide. Complete description of the sample can be found in de Vries (1979).*"

The reference is a textbook. We added the DOI in the reference list. Details of the calibration procedure are in two Dutch reports that we cannot find (Delft Hydraulics Laboratory, 1958 and 1969). But, additional elements are described in de Vries (1979). The first calibration tests consisted to compare the BTMA bedload catches with the average flume sediment transport. In a second time, tests were made by weighing catches in a sediment trap of the same size than sampler mouth. Sediment mixtures used during these tests was coarser than bedload of the Loire River with D_{50} varying between 2.5 mm and 6.4 mm. Both test series concluded to the same calibration coefficient of 2 (efficiency of 50%). This means that actual transport is two times higher than those measured with the sampler. The average efficiency of a basket sampler as the BTMA sampler is about 45% (Hubbell, 1964). These tests established a calibration curve linking unit bedload transport rates with caught volumes of sediments. It is mentioned in Boiten (2003) that this calibration factor was not including the possible losses of sediments finer than 0.3 mm (size of the mesh). The sand loss was estimated during flume experiments with a similar samples by Banhold et al. (2016) to 50% in average but with a mesh size of 1.4 mm and varied sediment mixture (D_{50}=[0.8-10] mm). There is no detail about the hydraulic coefficient of the BTMA which is the ratio between velocity in the sampler and flow velocity in de Vries (1979). But, author argued that the BTMA construction promote the transport coefficient at the expense of hydraulic coefficient. In comparison, the Helley-Smith sampler tend to overestimate bedload with the ratio between the flume sediment transport and the sampler sediment transport that vary with flow velocity from 1.2 to 2.6 (Helley and Smith, 1971). In the new version of the manuscript we present the BTMA as a pressure differential sampler.

"*Suggested values of α and b were adopted from Boiten (2003) which mentioned that the trap efficiency factor not include the possible losses of sediment finer than 0.3 mm (mesh size opening). This calibration factor come from successive calibration procedures concluding to the same calibration coefficient of 2 (de Vries, 1979).*"

The calibration approach is the approach which attempt to correlate bedload transport rates calculated from BTMA measurements with apparent bedload velocity measured by aDcp (Rennie et al., 2017).

"L. 113-114: If the boat was static, wouldn't it be better to assume zero boat velocity than to assume the boat velocity equaled the bed (bedform) velocity?"

The boat was almost static in the GPS referential (except some negligible lateral movements) but in the bottom track referential was mobile, due to mobile bed. When GPS data were missing, we considered the boat static in the GPS referential and we computed apparent bedload velocity directly from boat displacement in the bottom track referential.

"L. 127-128: There is no ds in equation 4. Is there a typo in the equation, or in the sentence? It seems like there should be d_s in the equation. If there is a constant ds, what is it?"

You are right, this was not clear. According to other referee comment, we changed the part of the method by separating the 2 kinematic models and explaining that the first one assumes that maximum bedload thickness is a single particle:

"$q_s ADCP = \frac{4}{3} \rho_s \, r \, V_{a\,proj} \times 10^3$;   *(4)*

Where $r = D_{50}/2$ is the particle radius, D_{50} is the median sediment diameter (m), ρ_s is the sediment density (2650 kg.m^{-3}). In this model, it is assumed the maximum bedload thickness is a single particle.

$q_s ADCP = V_{a\,proj} \, d_s \, c_b \, \rho_s$;   *(5)*

Where c_b is the concentration of the active transport layer considered as the saltation height (van Rijn, 1984), and the van Rijn (1984) formulation was adopted to compute the active layer thickness (d_s) as a function of the hydraulic condition and sediment grain size:"

"L. 159-160: This is the only mention of interval time and it's not clear how many repeat profiles were measured to measure bedform celerity?"

Bedform celerity was estimated from two repeated profiles for each sampling point. You can find mean interval time for each survey in the appendix D.

"L. 187-188: As this is a new method, this should be explained in more detail. Especially because acoustic power is affected by transmission losses that are a function of distance. It may be that the losses are negligible in the range considered here, but that should be demonstrated rather than just asserted."

Several tests were carried out to ensure that these acoustic power variations are not due to the variation of the distance between the hydrophone and the riverbed. A total of 20 drifts were performed with two hydrophones on the same horizontal location. The vertical location was

constant for hydrophone P1 (0.4 m depth) and variable for hydrophone P2 (from 0.4 m to 1.6 m depth). 3 drifts were done for each P1 and P2 configurations (except for P1 and P2 at the same depth, 2 drifts). When P1 and P2 are at the same depth, the acoustic power ratio is about 1.06, so the two hydrophones measured approximately the same acoustic power. This ratio doesn't evolve linearly with the proximity of river bed (depth ratio increasing). Even if the largest acoustic power ratio is observed for the largest depth ratio, it varies between 0.98 and 1.15 with a mean standard deviation of each configuration of 0.03 (see figure 1). This means that acoustic power can be considered as independent of the listening depth until 85% of the water depth.
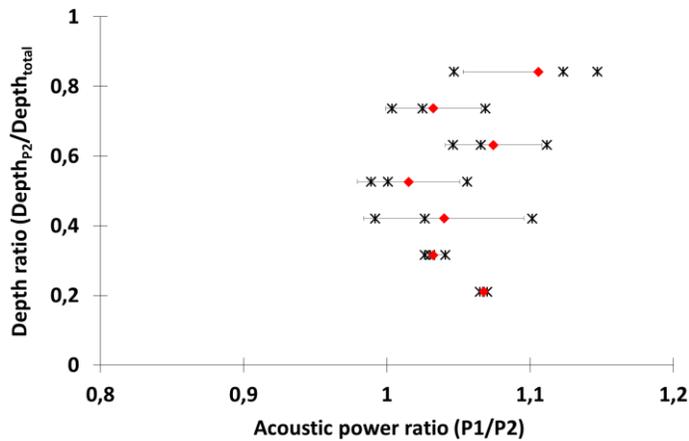

Figure 8: Variable vertical position of hydrophone in the water column and associated acoustic power variation (red dots are the mean values of acoustic power).

SC13

"L. 238: This should be in terms of qs as the dependent variable."

We presented the equation with the acoustic power as the dependent variable to make the comparison with Geay et al. (2020) easier. This did not influence the result of the regression because we applied a Reduced Major Axis (RMA) regression. Effectively, a classical regression aim to examining the dependent variable (Y axis) and the deviation of Y values from the fitted line are minimized. The RMA minimizes the deviations of the observations from the fitted line in both X and Y directions (Davis, 2003).

SC14

"L.307-308: While I think it's valid to use the hydrophone to look at the spatial distribution as done above, I don't think the direct comparison with the BTMA measurements is useful, because the hydrophone is calibrated to the BTMA. Some will "overestimate" and "underestimate" by definition of the calibration."

Here, we considered that our calibration equation was verified and we compared results of this equation with BTMA measurements. Of course, the overestimation or underestimation depends of the calibration equation but we believe that it is interesting to see when the calibration could differ from BTMA sampling.

<u>SC15</u>

"L. 346-347: Recently, Leary and Buscombe (2020), Ashley et al. (2020) and probably others have argued that repeat surveys with multibeam sonar are suitable for accurate measurements."

Here we wanted to highlight that the samplers were the only direct measurement of bedload in the field that are used as a reference measurement. We propose to reformulate this sentence: "*Despite their lack of accuracy and their low spatial representativeness, isokinetic samplers allow a direct measurement of bedload and represents the only reference measurement of bedload in the field.*"

We agree that the multibeam echosounders are helpful for the measurements mentioned here. The main problem is due to the geometry of the Loire at this site (wide and shadow) that makes the multibeam echosounding pretty difficult to perform. This was already done on the Loire (Claude et al., 2014; Wintenberger et al., 2015) but on study sites that were smaller.

<u>SC16</u>

"L. 371-373: The idea of a "global" calibration" for bedload by acoustic power is intriguing, but I have not yet had a chance to read the referenced paper in detail (it was only published this year). Given the wide variation in conditions that I think most would expect to work against a global calibration, a bit more discussion about why this might be possible would be helpful."

The question of the possible existence of a global calibration curve was, as you mentioned, addressed in Geay et al. (2020). It could be possible because the intensity of acoustic power is proportional to the number of particle impacts, the size of these particles and the impact velocity. Theoretically, if we consider two different sediment samples (one finer than the other) that produce the same acoustic power and the same bedload rate, the finer one produce more noise at high frequencies and less noise at low frequencies. This is the opposite for the coarser sample. The high frequency noise produced by the finer sample is compensated by the low frequency noise of the coarser sample. Moreover, for the same bedload rates, the fine sample is composed by more particles than the coarser but can produce the same energy as energy is related to the number of impact and the size of sediment. These assumptions do not take into account the problem of acoustic wave propagation (Geay et al., 2019).

<u>SC17</u>

"L. 406: Not clear what exactly is meant by this? Instantaneous?"

Here "*punctual*" means local measurement. In other words, the DTM measure bedload over a longitudinal distance with potentially different hydraulic and sediment transport conditions whereas BTMA measure bedload at a sampling point that reflect local conditions.

<u>SC18</u>

"L. 408: This is highly dependent on how the method is executed. If dune tracking is done at higher temporal frequency, it can provide"

Yes, by extending the temporal frequency of DTM method, the registered bedload variation could be at the same temporal scale for both methods. But, the special scale will always differ.

References

Banhold, K., Schüttrumpf, H., Hillebrand, G. and Frings, R.: Underestimation of sand loads during bed-load measurements- a laboratory examination, in: Proceedings of the international conference on Fluvial Hydraulics (River Flow 2016), 11-14 July 2016, Saint Louis, USA, 2406 pp., 2016.

Boiten, W.: Hydrometry, IHE Delft Lecture Note Series, A.A. Balkema Publishers, Netherland, 256 pp, https://doi.org/10.1201/9780203971093, 2003.

Claude, N., Rodrigues, S., Bustillo, V., Bréhéret, J. G., Tassi, P., and Jugé, P.: Interactions between flow structure and morphodynamic of bars in a channel expansion/contraction, Loire River, France, Water Resour. Res., 50, https://doi.org/10.1002/2013WR015182, 2014.

Davis, J., C.: Statistics data analysis in geology, 3rd ed., John Wiley & Sons, New York, 2003.

de Vries, M.: Information on the Arnhem Sampler (BTMA), Internal Report n°3-79, Delft University of Technology, Department of Civil Engineering, Fluid Mechanics Group,1979.

Geay, T., Michel, L., Zanker, S., and Rigby, J. R.: Acoustic wave propagation in rivers: an experimental study. Earth Surface Dynamics, 7 (2), 537–548, https://doi.org/10.5194/esurf-7-537-2019, 2019.

Geay, T., Zanker, S., Misset, C., and Recking, A.: Passive Acoustic Measurement of Bedload Transport: Toward a Global Calibration Curve?, J. Geophys. Res.: Earth Surf., 125 (8), https://doi.org/10.1029/2019JF005242, 2020.

Helley, E., J., and Smith, W.: Development and calibration of a pressure-difference bedload sampler, U.S. Geological Survey Open-File Report, 1971.

Hubbell, D., W.: Apparatus and Techniques for Measuring Bedload, U.S. Geological Survey Water-Supply Paper 1748, 1964.

Le Guern, J., Rodrigues, S., Tassi, P., Jugé, P., Handfus, T., Duperray, A., and Berrault, P.: Influence of migrating bars on dune geometry, in: Book of Abstracts of the 6th Marine and River Dune Dynamics conference, 1-3 April 2019, Bremen, Germany, 157-160, 2019a.

Le Guern, J., Rodrigues, S., Tassi, P., Jugé, P., Handfus, T., and Duperray, A.: Initiation, growth and interactions of bars in a sandy-gravel bed river, in: Book of Abstracts of the 11th Symposium on River, Costal and Estuarine Morphodynamics, 16-21 November 2019, Auckland, New-Zealand, 226 pp., 2019b.

Rennie, C. D., Vericat, D., Williams, R. D., Brasington, J., and Hicks, M.: Calibration of acoustic doppler current profiler apparent bedload velocity to bedload transport rate, in: Gravel-Bed Rivers: Processes and Disasters, Oxford, UK: Wiley Blackwell, 209–233, https://doi.org/10.1002/9781118971437.ch8, 2017.

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, https://doi.org/10.1016/j.geomorph.2015.07.032, 2015.