Response file to comments

Referee #1, Jens Turowski

1. First, the interpretation of the data frequently rests on subjective judgements or undisclosed criteria and arguments. For example, the authors identify trends or 'significant' differences without explaining their criteria or providing suitable statistical tests. Further, errors and uncertainties are not reported or discussed. Given the often large fluctuations, the trends, effects and differences claimed by the reviewers are mostly hard to spot, or they cannot be judged against the uncertainties in the measurements. This makes it hard for the reader to fully understand and believe the conclusions. I ask the authors to provide uncertainty estimates for the key parameters, and to re-assess their interpretation in light of these uncertainties, best with suitable statistical tests.

We appreciate the reviewer for pointing this out! We make efforts to conduct statistical analyses in the revised manuscript, and we think that this helps improve our paper considerably. The methodology that we implement is described below.

An analysis of the uncertainty of bed elevation standard deviation is now conducted by analyzing the DEMs of the flume floor. The uncertainty of bed surface GSD is now analyzed by implementing the Wolman method 5 times for each measurement. The uncertainty of the light table measurement is analyzed by comparing the light table results with the sediment trap results.

The variations in the data are analyzed with the coefficient of variation (cv, defined as the standard deviation over the mean value). The correlation coefficient r is calculated to study the relation between two parameters (e.g. sediment transport rate and duration of conditioning flow).

The methods we used are explained in detail in the text. See Lines 186-197 of the manuscript with Track Changes.

2. Second, the reader is not guided through the entire argument and often, the punchlines are not explicitly delivered. For example, in the discussion section 4.1, the authors finish their back-calculation of the threshold of motion with the statement "…indicating that only the slope effect cannot explain the observed range of tau_c* (line 354). The obvious question to ask is: if it is not slope, what is it? I think the authors are trying to address this question in the following paragraph (starting on line 360), but this paragraph contains merely some further statements about the data. As a reader, I am unsure what features of the data I should particularly aware of, how they are interpreted and how this leads to the conclusion of the authors. There are similar problems in other parts of the manuscript. For example, the authors claim that they observe reduced transport rates after the conditioning phase (e.g.,. line 376: "Our flume experiments also show a reduced sediment transport rate in response to the implementation of conditioning flow."), which presumably rests on data shown in Fig. 6c and Fig. 7. These data show large scatter and behavior contrary to the expectation. They and their interpretation are not picked up in the discussion other than as a stepping point to the implications. Statistical tests of presumed similarities and differences are missing, as are error bars. To address these issues, I ask the authors to develop a clear logic with explicitly stated punchlines. In the best case, they can formulate an expectation (e.g., condition leads to an increase in the threshold of motion), a null hypothesis (e.g., threshold is constant), and a suitable statistical

test to discriminate them. Then, they can walk the reader through the various observations until the conclusion is reached.

For Line 354, we calculate the cv (coefficient of variation) to compare the variability of the τ_c^* obtained by two different methods. The relation of Lamb et al. (2008) considers the effect of bed slope on the threshold of motion. Other effects that could influence the threshold of motion were previously discussed in more detail in Section 4.2. We rephrase the text as follows so that the reader can understand more easily. "Besides, the τ_c^* values predicted by the Lamb et al. (2008) relation show little variability among different experiments, compared with the values back-calculated with equation (1) based on experimental data. More specifically, the cv values are 0.032 at t = 15 minutes and 0.031 at the end of the conditioning phase for τ_c^* predicted by Lamb et al. (2008) relation, but become 0.10 at t = 15 minutes and 0.12 at the end of the conditioning phase for τ_c^* back-calculated with equation (1) using measured data. Such discrepancies could be ascribed to the fact the relation of Lamb et al. (2008) considers only the influence of bed slope, without considering the effects of other mechanisms like organization of surface texture, infiltration of fine particles, etc. These potential effects are discussed in more detail in Section 4.2." See Lines 457-464 of the manuscript with Track Changes.

For Line 376, what we meant is that the implementation of a (long) conditioning duration will lead to a reduction of sediment transport at the beginning of the subsequent hydrograph. This effect, however, gradually diminishes with the increase of flow intensity and sediment supply during the hydrograph. This conclusion is supported by Figure 6(b) and Figure 7. We rephrase the text to make this clear. See Lines 502-503 of the manuscript with Track Changes. Quantification analysis is added in Section 3; for example, Lines 343-364 of the manuscript with Track Changes.

As for the issue of uncertainties, please see our reply to Question 1 and other related questions below (both reviewers provided constructive comments on this issue).

3. 44 to me it makes sense to differentiate the terminology somewhat. The terminology chosen in the cited papers seems to be motivated by the available data and the chosen approach. For example, Mao looked at the effect of flood events, while Masteller et al. focused more on continuous discharge data. Further, a change in flow does not directly translate into a change in stress, there are non-linearities and feedbacks involved (e.g., non-steady flow, re-organization of the bed). It does make sense to use a stress approach (and therefore terminology) for experiments, but stress information is much less reliable for field measurements.

The referee's comment is very helpful. We agree with this! We put part of this comment in the paper. See Lines 46-50 of the manuscript with Track Changes.

4. 69 Maybe the effect of sand supply on gravel transport should be mentioned somewhere in this paragraph; see for example Curran and Wilcock, JHE 2005 (DOI: 10.1061/(ASCE)0733-9429(2005)131:11(961).

We add the following sentence in this paragraph. "Besides, the supply of fine sediment (during high discharge events) is also widely observed to enhance the mobilization of coarse sediment (Wilcock et al., 2001; Curran and Wilcock, 2005; Venditti et al., 2010)." See Lines 84-85 of the manuscript with Track Changes.

5. 75 also Lenzi, (Step-pool evolution in the Rio Cordon, Northeastern Italy. Earth Surface Processes and Landforms 26: 991–1008, 2001), and Turowski et al., 2009 (DOI: 10.1002/esp.1855).

We add these two references in the paper. See Lines 79-80 of the manuscript with Track Changes.

6. 89 To investigate the study objectives…

Done.

7. 107 …not fed…

Done.

8. 107-108 The approach and reasoning here needs some more detail. What were the simulations? What kind of trial experiments? How was the final feed rate chosen?

We rephrase the text as "For each step of the hydrograph, the feed rate of sediment was specified to be close to the transport capacity of the flow. Determination of the sediment supply rates was facilitated by a numerical model which was calibrated for similar experimental conditions (Ferrer-Boix and Hassan, 2014)." See Lines 120-122 of the manuscript with Track Changes.

9. 136-141 Unclear, lacks detail.

We add more details in the text. Please see our reply to Questions 10, 12, and 13.

10. 137 How were images merged? How did you deal with image distortion due to merging or lense distortion effects?

We add the following sentence in the text. "To avoid distortion effects due to image merging, the width of the image strips that were stitched to get a composite image was specified as just 2 cm." See Lines 154-155 of the manuscript with Track Changes.

11. 137 perform

Done.

12. 139 I don't understand how particle sizes were measured.

Basically we used the Wolman count method to calculate the grain size distribution. This was done by identifying the grain sizes from the photograph (with a 5 cm grid superimposed on the photograph). We rewrite the text as follows, "The particle size distribution of the bed surface was estimated using the Wolman (point count) method, by identifying the grain size of particles at the intersections of a 5 cm grid superimposed on the photograph. Individual grains were identified by color. For each experiment, the grain size distribution of the bed surface was calculated at different times to quantify its changes during the experiment." See Lines 155-161 of the manuscript with Track Changes.

13. 140 How were changes quantified?

In each run, photographs were taken at different times throughout the experiment to calculate the bed surface grain size distribution. Information about the measurement frequency was given in the paper. See Lines 177-180 of the manuscript with Track Changes. The changes can thus be quantified directly by comparison of the grain size distributions at different times. We rewrite the text as "For each experiment, the grain size distribution of the bed surface was calculated at different times to quantify its changes during the experiment." See Lines 159-161 of the manuscript with Track Changes.

14. 144 Please add some information on accuracy and precision of this method.

Zimmerman et al. (2008) have conducted a detailed analysis of the accuracy of this method. They reported that the light table method is accurate and precise for sediment from 2 mm to 45 mm after calibration. Here we also analyze the accuracy of this method for the experiments presented in this paper. What we do is as follows.

To estimate the uncertainties of the light table method, we compare the data measured by the trap and the data measured by the light table, in terms of both sediment transport rate and the characteristic grain sizes of sediment load. This is stated in Section 2 of the main text. See Lines 193-195 of the manuscript with Track Changes.

For the total sediment transport rate, the light table data and the trap data show good agreement. More specifically, the light table overestimates the total sediment transport rates by 4% on average (111 samples and a standard deviation of 14.5%). 70 out of 111 samples show an accuracy of ±10% and 93 out of 111 samples show an accuracy of ±20%. We present the results in both the main text (Lines 271-276 of the manuscript with Track Changes) and the Supporting Information (Section S2).

Uncertainties of the bedload characteristic grain sizes are as follows. The D_{50} of bedload measured by light table show relatively good agreement with that measured by trap. The light table overestimates the D_{50} by 3% on average (111 samples and a standard deviation of 40.1%). The accuracy of the values of D_{10} and D_{90} of the bedload is not as good as that of D_{50}. The light table underestimates D_{10} by 20% on average (111 samples and a standard deviation of 39.0%), and overestimates D_{90} by 30% on average (111 samples and a standard deviation of 26.5%). We present the results in both the main text (Lines 366-372 of the manuscript with Track Changes) and the Supporting Information (Section S2).

15. 156 What does "slowly" mean in this context? Please also explain how an effect of the rising discharge after measurements was avoided.

We explain this issue in the manuscript. We rephrase the text as "For each measurement of DEM/Wolman, we stopped the pump instantaneously and let the flow slowly lower and then stop to allow for the bed to be scanned by a laser and photographed. The time interval between the stop of the pump and the stop of the flow was about 3 to 4 minutes. To avoid the influence of the following rising discharge, all subsequent measurements were taken after the flow became stable." See Lines 180-183 of the manuscript with Track Changes.

16. 168 why the average over the cross section, instead of, for example, the thalweg? Maybe this should be described and justified in the method section.

We add the following sentence in the paper. "The DEM over the cross section is used here to study the overall aggradation/degradation of the channel." See Lines 207-208 of the manuscript with Track Changes. For reference, we also add the DEM of bed surface at different time during the experiment. This is presented in the Supporting Information (Section S1)

17. 177 How were trends assessed here?

The "word" trend might be misleading. We rewrite this sentence as "…, with no evident aggradation/degradation being observed". See Line 219 of the manuscript with Track Changes. We also calculate the mean difference of bed elevation to support the conclusion. Results are added in Table 1.

18. 180 How exactly was this assessed? What features did you look for to identify a bedform?

Our determination of the bedform is visually based on the DEM as well as the direct observation of the channel bed. We realized that the statement was not accurate without further quantification. Therefore, we remove the sentence in the text to avoid misunderstanding. See Lines 223-224 of the manuscript with Track Changes. We also suggest the reviewer to refer to the DEM that we add in the Supporting Information (Section S1).

19. 191 Here, the authors implicitly identify the standard deviation of the bed with the bed roughness. Bed roughness is a technical term in fluvial hydraulics, and although there is evidence that the standard deviation scales with roughness, the terms are not directly equivalent, as they are used here.

We change the "bed roughness" to "standard deviation of bed elevation" as the referee pointed out. Our recently published paper on WRR (Chen et al., 2020) supports the idea that standard deviation of bed elevation is a good descriptor for bed roughness of gravel-bed rivers. We rewrite the beginning of this paragraph as follows "Figure 4 shows the temporal variation of the standard deviation of bed elevation, which is often scaled with the bed roughness for gravel-bed rivers (see Chen et al. (2020) for a detailed discussion on this topic), over the length of the erodible bed during the experiment." See Lines 231-236 of the manuscript with Track Changes.

20. 193 what does "almost constant" mean in this context? Can you make this statement quantitative?

We calculate the coefficient of variation (cv = sigma/mean) to quantify the temporal variation. Results are presented in the manuscript as follows. "The standard deviation of bed elevation becomes quite stable during the remaining conditioning phase, as well as during the hydrograph phase, despite the fact that degradation is evident as the flow approaches its peak value. For the standard deviation of bed elevation during the conditioning phase, we calculate the coefficient of variation (cv) for REF2 (15), which has the longest conditioning phase. The result shows a value of 0.038 from t = 15 minutes to the end of the conditioning flow. For the standard deviation of bed elevation during the hydrograph phase, we calculate the cv for all experiments. The results show that the values of cv vary between 0.031 and 0.075." See Lines 237-243 of the manuscript with Track Changes.

21. 198 please add in symbols to mark the time of the actual measurements.

We replot Figure 4 according to the suggestion of the reviewer.

22. 198 is it possible to add error bars to these plots?

For the standard deviation of bed elevation, we estimate the uncertainties as follows. We scanned the floor of the flume twice and calculated the standard deviations of the scanned DEM. The floor of the flume was horizontal and flat, with no sediment on the bed. Theoretically, the standard deviation of the DEM should be zero. Therefore, the calculated standard deviations of the flume floor could be regarded as an estimation of the uncertainty of the calculated values during the experiment.

The way we estimate the uncertainties is explained in Section 2 of the paper. See Lines 186-190 of the manuscript with Track Changes. According to this method, we estimate the uncertainties of the results in Figure 4 to be in the range of 1.6~2.5 mm, which is close to the vertical resolution of the laser scans (1 mm). We present our estimation in the caption of Figure 4.

23. 204 What does "relatively stable" mean in this context? How was stability assessed?

We calculate the coefficient of variation (cv = sigma/mean) to quantify the temporal variation of the variables. Results are given in the text. See Lines 256-258 of the manuscript with Track Changes.

24. 205 What does "relatively constant" mean in this context? How was constancy assessed?

Again, we calculate the coefficient of variation. Results are given in the text. See Lines 260-262 of the manuscript with Track Changes.

25. 206 Interpretation, to discussion.

Thanks for pointing this out. We now move this interpretation to the Discussion Section. See Lines 509-512 of the manuscript with Track Changes.

26. 209 Please mark the time of the measurements on the plots and add error bars.

We replotted Figure 5. Time of the measurements are marked, and range bars are added to estimate uncertainties.

27. 228 Fig. 6c shows presumably not a derivative, but a ratio of discrete changes. Please change the notation accordingly (for example, by using delta symbols). Please add error bars.

We do not think we are showing a ratio of discrete changes. We tried to quantify the trend by linear regression. Here is what we said in the paper, "Such intra-step variations of sediment transport rate are investigated in Fig. 6(c), with the x axis being the averaged sediment transport rate of each step Q_{sa} and the y axis being $d(Qs/Q_{sa})/dt$. The value of $d(Qs/Q_{sa})/dt$ is estimated by linear regression."

As for the uncertainties of the light table method, please see our reply to Question 14.

28. 229 During each of the hydrograph steps, there seems to be decline of transport rates over time. This may be due to a transient adjustment of the bed to the changed hydraulic conditions. It does not seem to me here that an equilibrium is achieved. Can the different stages then be meaningfully compared? How would a transient adjustment affect the interpretations?

Yes, this is what we tried to quantify in Figure 6c. A negative value of discrete change denotes a declining trend, whereas a positive value denotes an increasing trend. We agree with the reviewer that these are transient adjustments due to changed water and sediment supply, and the equilibrium is not achieved. Actually, the decreasing/increasing trends exist just because the bed is not in equilibrium. We explain a bit more in the text. See Lines 325-326 of the manuscript with Track Changes. The sediment transport in different stages were compared in Figure 6b and in the text (Lines 299-308 of the manuscript with Track Changes).

29. 237 …of the sediment transport rate…

Done.

30. 238 How did you assess whether it agrees? What does agreement mean in this context?

When we say "agree with", we mean that the adjustments of sediment transport and the adjustments of bed elevation show similar patterns during the hydrograph phase. In order to make our statement clear, we rewrite it as follows. "Such adjustments of sediment transport rate are consistent with the process of channel deformation shown in Fig. 3. That is, for both sediment transport and channel deformation, results of REF7 (0.25) deviate from other experiments in Step 1 (larger sediment transport rate and more degradation in REF7 (0.25)), but collapse with other experiments in the following three steps." See Lines 304-308 of the manuscript with Track Changes.

31. 246 Please explain how you detected trends and give corresponding statistics. The interpretation is a little difficult, since there are no error bars for the data.

The trends were detected based on linear regression, as we have stated in the text. See Line 311 of the manuscript with Track Changes. A negative value of $d(Qs/Q_{sa})/dt$ corresponds to a decreasing trend, and a positive value of $d(Qs/Q_{sa})/dt$ corresponds to an increasing trend.

32. 260 How did you arrive at this assessment? REF7 is around 50% larger than REF3! It would make sense to add error bars to the measurements and a statistical test to actually show that there is no difference.

Our previous statement might be misleading. What we want to express was that a longer duration of conditioning flow does not lead to a reduced sediment output during the subsequent hydrograph. We add a comparison of the data and also a calculation of the correlation coefficient, which helps support our idea. The text is rephrased as follows. "It can be seen that the effect of conditioning duration on the total sediment output during the entire hydrograph phase is not evident: a longer duration of conditioning flow does not necessarily lead to a smaller (or larger) sediment output. The largest sediment output occurs in REF7 (0.25), which is 55% larger than the sediment output in REF3 (10) which has the smallest output, but is about the same as (only 4% larger than) the sediment output in REF6 (15). We further calculate the correlation coefficient between the total sediment output and the duration of conditioning flow, and obtain a value of r = -0.14, indicating that there is almost no correlation between the two parameters." See Lines 330-336 of the manuscript with Track Changes.

As for the error bar, there is no error bar for the trap data. Data shown in Figure 7 are not calculated values, but are the material weighted in the sediment trap. Errors could be introduced due to the 0.25 mm mesh size of the tail box, the resolution of the scale, etc. However, we think these errors are negligible. The word "calculate" at the beginning of this paragraph could be misleading. We rephrase the text as "Sediment collected in the trap/tailbox at the flume outlet allows us to plot the total amount of sediment output during each step of the hydrograph." See Lines 327-328 of the manuscript with Track Changes.

33. 265 please add uncertainties.

The uncertainties of the trap data cannot be evaluated, since the data do not correspond to calculated values but instead correspond to the material weighted in the sediment trap. Please see our reply to Question 32 (the second paragraph).

34. 270 with a longer conditioning phase leading…

Done.

35. 274 How did you arrive at this interpretation? Maybe this is just an effect of the scale of the plot? It would be good to add error bars to the data here and some suitable statistical test.

To better illustrate our idea, we quantitatively compare the data between different experiments. We also calculate the correlation coefficient *r* to estimate the effect of stress history on sediment transport. We rephrase the text. See Lines 343-364 of the manuscript with Track Changes.

As for the error bars, again we state that there is no error bar for the trap data. Please see our reply to Questions 32 and 33.

36. 290-298 how did you establish significance and what does 'more significant' mean in this context / how was this assessed?

We use the coefficient of variation (cv = sigma/mean) to support our idea. The value of cv is calculated for both the conditioning phase (after t = 10 hour, the beginning of the conditioning sees a drop in D_{l10} so that the cv is not appropriate to quantify the fluctuation) and the first two steps of the hydrograph phase. Results are presented in the text. See Lines 381-397 of the manuscript with Track Changes.

37. 300 Please give some indications of the uncertainties of these measurements.

In our reply to Question 14, we explain the uncertainties of light table measurements (both total transport rate and characteristic grain sizes) in detail. We do not repeat the explanation here. Please refer to our reply to Question 14.

38. 303-308 this needs more detail if it is relevant for the central message of the paper. If not, consider deleting it.

We put more detail concerning fractional sediment transport in the Supporting Information.

39. 304 The use of the term 'equal mobility' has become ambiguous. Originally, it meant that the grain size distribution of the transport material matches the grain size distribution of the material found on the bed. However, it is now often used to mean an equal threshold of motion for all grain sizes. It is unclear from the context here which meaning is intended.

In our study, "equal mobility" means the first and original definition. That is, the grain size distribution of the sediment load matches the grain size distribution of be sediment on bed surface. We add the definition of "equal mobility" in the text. See Line 404 of the manuscript with Track Changes. This definition is also added in the Supporting Information when sediment mobility is discussed.

40. 334 transport rate

Done.

41. 365 please add error estimates to these calculations.

We now estimate the uncertainty associated with the calculation of τ_c^*. The methodology is explained in the text. See Line 466-476 of the manuscript with Track Changes.

For the τ_c^* values back-calculated with Equation (1) (i.e., Meyer-Peter and Muller type relation), the estimated uncertainties are presented in Table 1. For the τ_c^* values calculated with the Equation (5) (i.e., the equation of Lamb et al. (2008)), the uncertainty is less than ±1%, and is therefore not presented in Table 2 (but is explained in the text).

42. 388 predictions

Done.

43. 397 This would imply that the stronger trends should be seen in d90 rather than d50, right? How does this expectation compare to the data?

The reviewer might misunderstand. With experiments, Masteller and Finnegan (2017) found that the most drastic changes during the conditioning flow are the reorientation of the highest protruding grains into nearby available pockets. Such a reduction in the number of highly protruding grains eventually leads to a reduction of sediment transport rate and a more stable bed surface. Masteller and Finnegan (2017) also reported that this reorganization of bed surface, however, does not lead to an evident change of bed surface GSD or surface topography standard deviation.

Therefore, we do not expect that such reorientation would be reflected in the variation of D_{90} or D_{50}, as mentioned by the reviewer. We revise the text in the paper to make this more clear. See Lines 528-530 of the manuscript with Track Changes.

44. 405 either 'e.g.' or 'etc.', not both

We delete "etc.". Thanks for pointing this out!

45. 416 …the conditioning flow was…

Done.

46. 416 In the present paper…

Done.

47. 427 consisted

Done.

48. 430 consisted

Done.

Referee #2

1. This paper presents a series of flume experiments detailing the effect of conditioning flow duration on the sediment flux experienced during a subsequent flood wave. The paper in itself is interesting and would be of interest to the readers of ESurf but at the moment the language is not tight enough and the reader is left to try and fathom out the main take home messaged from each set of data analysis. For example, throughout the results there are phrases like 'increases notably', 'significant degradation' etc. but it is not clear whether or not the authors have undertaken statistical analysis to support their results. If they have then the outputs of those statistical analysis need to be reported. It is also unclear how reproducible the results are – again this needs to be reported. Therefore, this leaves the reader wondering how important the reported trends are, especially given the data is relatively noisy. I would also argue that the majority of the data analysis is relatively basic, which in itself is not a problem, but when coupled with the subjective language used, as discussed above, the main 'story' of the manuscript is lost. That said there are places where I think the analysis could be taken a lot further e.g. analysis of bed surface topography is not really mentioned despite DEMs being collected and mentioned in the methodology. Finally, I have a significant issue with section 4.1 and the back calculation of tau*c using MPM and Wong and Parker since this regression is designed for a small range of known tau*c. Irrespective of this I am not sure of the worth of scaling tau*c from that derived right at the beginning of the experiment – the bed state there is not representative of a true fluvial system, rather is representative of artificial conditions caused by screeding and high sediment transport rates caused by initial scour. So basing analysis off this seems odd and slightly misleading. In this section the authors also say that 'only the slope effect cannot explain the observed range of tau*c' – it is not clear what they mean and given this paper is about steep slope environments this needs careful clarification and expansion.

We appreciate the reviewer for the constructive comments. In the revision of the manuscript, we undertake statistical analysis to support our ideas. The uncertainties of the measurements are estimated. The variations and the correlations of the data are analyzed. Methodology is explained in detail in the text. See Lines 186-197 of the manuscript with Track Changes. Results are presented in reply to the relevant questions of both reviewers, as well as in the manuscript.

As for the reproducibility of the results, the reviewer raised an important issue. In a recently published paper, Church et al. (2020) drew attention to similar issue. Moreover, they distinguished three levels of "reproducibility": (1) "repetition" which repeats the program of observations in the same exercise; (2) "replication" which is duplication of observations using similar resources but in an independent program; (3) "reproduction" which is confirmation of scientific principles using different resources in an independent program. In our paper, the repetition of the experimental results is tested by repeating the conditioning phase with the longest duration (REF6 (15) and REF2 (15)). The two experiments show similar results in terms of standard deviation of bed elevation, GSD of bed surface, sediment transport rate, and GSD of sediment load. However, the reproduction of the experimental results, which requires independent tests undertaken using different materials and/or different conditions of measurement, and is more significant for advancing of the science according to Church et al. (2020), has not been tested in this paper. In this regard, more efforts are needed in future studies to test the reproducibility of the conclusions given in this paper. We discuss this issue in the manuscript. See Lines 562-570 of the manuscript with Track Changes.

The reviewer mentioned that the data reported in the paper was relatively noisy, and that the analysis were relatively basic. We add statistical analysis in the revised manuscript. Details of the statistical analysis are provided throughout this response file (see our reply to related comments of both reviewers). As for the analysis of the bed surface topography, we now calculate both the mean and the standard deviation of the DEM. We explain this in more detail in our answers to Questions 17 and 19 of Reviewer #2.

In the paper, when implementing the MPM relation modified by Wong and Parker (2005) for the estimation of τ_c^*, we assumed that the MPM type relation holds under the condition of our experiments. It is worth mentioning that in a newly published paper, Hassan et al. (2020) applied three different methods to estimate the threshold of sediment motion in a gravel-bed river, including (1) back calculation with the Wong and Parker (2005) relation of MPM; (2) back calculation with the Wilcock and Crowe (2003) relation; and (3) the relation of Church et al. (1998). Estimation with the three different methods shows very similar temporal trend and variability (0.035~0.075 with Wong-Parker relation in their case), which implies that the specific function that is applied does not matter that much. We cite the work of Hassan et al. (2020) in the manuscript. See Lines 416-417 and 431-433 of the manuscript with Track Changes.

We would like to clarify that in the paper, τ_c^* is scaled against the value at t = 15 minute, rather than at the very beginning of the conditioning phase. According to the experimental results (Figures 3, 4, and 5 in the manuscript), adjustments of the bed topography and surface texture have been accomplished within the first 15 minutes (and become rather insignificant after that). Therefore, we do not agree that "the bed state (as the base for scaling) is not representative of a true fluvial system, rather is representative of artificial conditions caused by screeding". Besides, the main purpose of the scaling is to facilitate the comparison of the temporal variation of τ_c^* among different experiments.

As for the slope effect, Reviewer #1 raised similar question. Please refer to our reply to Question 2 of Reviewer #1. More specifically, we say the following in the paper (Lines 462-464 of the manuscript with Track Changes). "Such discrepancies could be ascribed to the fact the relation of Lamb et al. (2008) considers only the influence of bed slope, but without considering the effects of other mechanisms like organization of surface texture, infiltration of fine particles, etc. These potential effects are discussed in more detail in Section 4.2."

2. I really like it when authors include an implications section in their papers and so it is good to see this included in the presented paper however some of the text in this section feels much more like discussion and framing of the authors results within the wider literature of which there was relatively little of in the paper up until that point. So, I think the discussion and implications sections of the paper need re-framing slightly so the discussion section properly frames the results within the wider literature and the implications section talks about the bigger picture and importance of the findings more broadly. If the authors do this, I think it will be much clearer to the reader about the new findings which this paper had generated – to make this even easier for the reader there are places where paragraphs could be re-structured such as to lead with the findings from this paper before framing within the previous literature. This will help make it crystal clear where the additional knowledge is. For these reasons I would suggest a rejection of the paper with a strong encouragement to resubmit once the issues have been addressed. Line by line changes are also suggested below to help the authors revise their manuscript.

The Discussion Section has been rephrased according to the related comments from both reviewers. We really appreciate this!

3. Line 24 – arguably not just mountain streams.

We delete the word “mountain”.

4. Line 26 – I would consider adding Masteller 2019 JGR paper in here as it seems relevant.

We add this reference here.

5. Line 28 – what do you mean by average flow regime?

When we say “average flow regime”, we mean that most sediment transport relations for mountain streams are based on constant flow, which may be regarded as an average of the high unsteady flow regime of mountain streams.

6. Line 44 – remove etc in the citations.

Done.

7. Line 52 – should be Haynes

Thanks for pointing this out! We have corrected the typo.

8. Line 82 – from reading your methodology I would argue that you don’t run experiments which consist of ‘extended cycles’ – for me that reads as you cycled hydrographs but this is not what you did – instead you changed the length of a period of conditioning flow before exposing that bed to the rising limb of the hydrograph.

We agree with the reviewer. We revise the text to avoid misunderstanding. The text now reads “In this paper, flume experiments consisting of high and low flow are conducted to study this problem”. See Line 89 of the manuscript with Track Changes.

9. Line 88 – can you be more explicit with what you mean by guided by?

Here we mean that the East Creek is the prototype for the experimental arrangements in this paper. This was explained in more detail in the subsequent text, including the design of flow discharge (Lines 116-119 of the manuscript with Track Changes) and sediment grain size distribution (Lines 132-135 of the manuscript with Track Changes).

10. Line 89 – to investigate the study objective……

Done.

11. Line 104-106 – so was the flow rate directly scaled?

Yes, the flow rate was directly scaled as written in the paper.

12. Line 105 – I don't think that you ran a hydrograph – you ran a rising limb of a hydrograph but not a full hydrograph. The results you would have come up with by running a full hydrograph would have been very different to those which you report here

The reviewer is correct. To avoid misunderstanding, we add the following sentence in the manuscript. "It should be noted that in the experiments, we only implemented the rising limb of the hydrograph/sedimentograph, rather than a full hydrograph/sedimentograph with both rising and falling limbs. Rather than studying river adjustment during a flow hydrograph, we aimed at determining the influence of conditioning time on bedload and bed surface arrangements as flow rates increased." See Lines 102-105 of the manuscript with Track Changes. The reason that we implemented only the rising limb is that we would like to focus on how the stress history effect is influenced with the increase of flow intensity.

13. Line 108- 111 – more justification and reasoning is needed in this section – what are the details of the trail experiments? How exactly were the feed rates chosen?

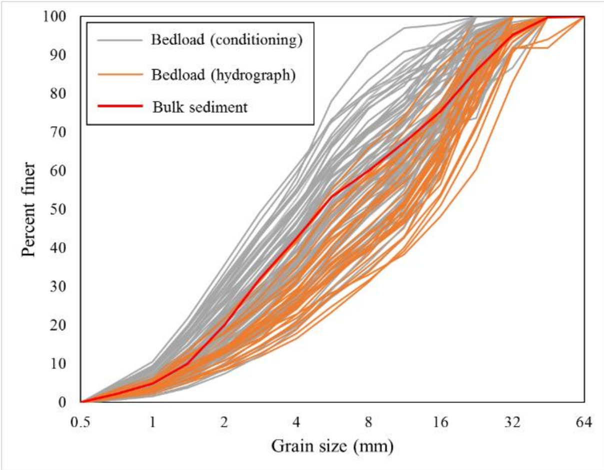
We rephrase the text as "For each step of the hydrograph, the feed rate of sediment was specified to be close to the transport capacity of the flow. Determination of the sediment supply rates was facilitated by a numerical model which was calibrated for similar experimental conditions (Ferrer-Boix and Hassan, 2014)." See Lines 120-122 of the manuscript with Track Changes.

14. Line 117 – why was this sediment scaling chosen?

The scaling factor used in our experiments (1:4) is similar to that used by Chartrand et al. (2018) (which is 1:5). Also, the finest fractions (that less than 0.5 mm) were excluded to avoid sediment transported in suspension.

15. Line 123 – what impact does feeding the bulk sediment rather than feeding a GSD which is representative of the transported sediment have on the surface development?

We compared the GSD of the bulk sediment against the GSDs of bedload measured with the trap. Results are shown in the subsequent figure. It can be seen that the GSDs of bedload vary over a relatively wide range according to the flow intensity, sediment supply rate, bed surface texture, etc. The GSD of bulk sediment falls in about the middle of this range. Therefore, we think that the bulk sediment is a good representative of the bedload in average. Moreover, for gravel-bed rivers, the bulk material is often regarded as a representative of the sediment supply texture in the long term.



16. Line 127 – 129 – how often did you measure the bed surface profile (It is difficult to tell from fig 1) – was this just down the channel centre line?

The measurement of bed surface profile was denoted as "DEM" in Figure 1. Information related to the measurement frequency was given in the text. See Lines 177-180 of the manuscript with Track Changes. The DEM of the bed surface was measured for the whole channel bed, rather than just the channel centerline.

17. Line 131- 136 and Line 156 - What impact did draining the flume have on the transport dynamics – did you assess this? How were the DEMs detrended? You need a sentence or more in here justifying why you are using std of elevations as a measure of surface topography.

To minimize the influence of draining the flume, the flow was lowered slowly with the time interval between the stop of the pump and the stop of the flow being about 3 to 4 minutes, and the subsequent measurements (point gauge, light table, etc.) were taken after the flow was back to stable. We revised the text accordingly to present more details. See Lines 180-183 of the manuscript with Track Changes.

The DEMs were detrended based on linear regression. We add the information in the text. See Line 151 of the manuscript with Track Changes.

We used the standard deviation of elevation as a measure of surface topography, since our recently published paper (Chen et al., 2020) showed that the standard deviation of bed elevation is a good descriptor for bed roughness of gravel-bed rivers. Reviewer #1 asked a similar question. We explain this in the manuscript (Lines 231-233 of the manuscript with Track Changes).

18. Line 148 – was the shear stress corrected for side walls? If so how?

A side wall correction is not implemented in this paper. In our experiments, the flow depth varies in the range of 0.046-0.086 m, which leads to a width/depth ratio of 6.4-12.0. According to Julien (2010), side wall effects are not significant when the width/depth ration is larger than 5.

We apply a side wall correction for the two cases with largest and smallest depth, using the method of Chiew and Parker (1994). Results show that, for the case of the smallest depth, the shear stress without side wall correction is 15.7 Pa, and the shear stress with side wall correction is 15.2 Pa (a reduction of 3.0%). Whereas for the case of the largest depth, shear stress without a side wall correction is 25.3 Pa, and shear stress with side wall correction is 24.7 Pa (a reduction of 2.7%). Therefore, we think it is reasonable to neglect side wall effects in the analysis.

19. Line 166 – I would have thought it would have been useful to present the statistical moment analysis of the bed scans so you can properly link the development of the bed surface with some of the sediment dynamics. I think it would also be useful to plot the surface DEM evolution to allow readers to better appreciate and understand how the surface evolves.

We have calculated the standard deviation (second order moment) of the DEM, as shown in Figure 4. We chose the standard deviation because our recent research (Chen et al., 2020) found it to be a good descriptor for bed roughness of gravel-bed rivers. We also studied the mean (first order moment) of the DEM. This is analyzed in terms of the mean difference of bed elevation during each flow stage, which represents the overall channel aggradation/degradation in each stage. Results are presented in Table 1 (Δz_b) as well as in the text associated with Figure 3.

The DEMs of bed surface at different times during the experiment are now plotted in the Supporting Information (Section S1).

20. Line 168 – I am not sure exactly what you mean by longitudinal DEM – why average it over the cross section? A lot of the previous literature on stress history which has undertaken DEM analysis has shown that the surface develops significant spatial complexity which will be lost by the averaging you have undertaken

Here we averaged the DEMs over the cross section with the purpose to study the overall aggradation/degradation in the longitudinal profile during different stages of the experiments. We add the following sentence in the paper. "The DEM over the cross section is used here to study the overall aggradation/degradation of the channel." See Lines 207-208 of the manuscript with Track Changes.

21. Line 180 – what do you mean by further analysis of the DEM? Where is this analysis?

Our determination of the bedform is visually based on the DEM as well as the direct observation of the channel bed. We realized that the statement was not accurate without further quantification. Therefore, we have removed the sentence in the text to avoid misunderstanding. See Lines 223-224 of the manuscript with Track Changes. We also suggest the reviewer to refer to the DEM that we add in the Supporting Information (Section S1).

22. Line 205 – should be noting not noted

Done.

23. Line 205 – what do you mean by keeps relatively flat?

We mean that the characteristic grain sizes of bed surface do not change much with time. To support our idea, we add statistical analysis by calculating the coefficient of variation (cv = sigma/mean). Results are given in the text. See Lines 256-262 of the manuscript with Track Changes.

24. Line 212 – how accurate is the light table?

Reviewer #1 asked a similar question. We have answered in detail and thus do not repeat here. Please see our reply to Question 14 of Reviewer #1. Thanks!

25. Line 216 – 217 – again be specific – what do you mean by very large? Gradually dropping? Small and relatively constant?

We add some analysis. The following sentences are added in the text.

"In the first 15 minutes, the sediment transport rates drop from more than 500 kg/h to less than 100 kg/h. Afterwards, it takes about another 2 hours for the sediment transport rates to drop to close to 1 kg/h."

"For REF2 (15) and REF6 (15) which have the longest conditioning phase, the sediment transport rates between t = 8 hour and the end of conditioning phase (t = 15 hour) show mean values of 0.35 kg/h (standard deviation = 0.22 kg/h) and 0.37 kg/h (standard deviation = 0.24 kg/h), respectively."

See Lines 279-285 of the manuscript with Track Changes.

26. Line 223 - why is this analysis in the supplemental information? The analysis you have undertaken up to this point in the paper would really seem to benefit from this further analysis

We put this material in the Supporting Information as we think it does not belong to the main conclusions of this paper. Besides, the current manuscript is already very long.

27. Line 268 and 275 – again be specific and give the statistical outputs if you have undertaken statistical analysis.

We add a statistical analysis in the text. See Lines 344-346 and Lines 354-357 of the manuscript with Track Changes.

28. Line 279 – I am not sure I agree that the five experiments show similar sediment outputs – there may not be a systematic trend related to condition flow duration but there are certainly differences between them.

We rephrase the text as follows. "In the last step of the hydrograph, with the flow discharge and sediment supply approaching their peaks, the difference in sediment output among the five experiments again becomes small, with the values ranging between 72.1 kg in REF4 (2) and 119.6 kg in REF6 (15). This demonstrates that little influence of stress history remains in this step." See Lines 361-364 of the manuscript with Track Changes.

29. Line 303 – 308 – again why is this data in the supplemental information – I would have thought it would have been a really important addition to your paper and provided a lot of useful context from which to hang your discussion

We thank the reviewer for the positive comment. We put this material in the Supporting Information as we think it does not belong in the main conclusions of this paper. Besides, the current manuscript is already very long.

30. Line 334 – sediment transport rate

Done.

31. Line 337 – what do you mean by basically show an increasing trend – be specific

We mean that the value of dimensionless sediment transport rate q^* increases with the increase of Shields number τ^*_{s50}. Our calculation shows a rather good correlation between $\tau_{s50}{}^*$ and $\log(q_s{}^*)$ (consistent with the semi-log scale of Figure 9(a)), with the correlation coefficient being 0.58. We add this result in the text to support our conclusion. See Lines 440-441 of the manuscript with Track Changes.

32. Line 376 – remove etc. from the citations

Done.

33. Line 391 – what implications – can you be specific?

We remove this sentence.

34. Line 405–413 – I am afraid I don't see the relevance of this paragraph to the paper.

We would like to keep this paragraph. The idea of this paragraph is that, in our flume experiments sediment was fed at the upstream with low supply during conditioning flow and high supply during flood. However, natural mountain streams are more complicated and not always like this. Sediment supply during low flow could be abundant (e.g. in the form of landslides), especially at places where the hillslopes are active. Such effects are not included in our experiments, but are discussed in this section. We think this would also merit future research.

35. Line 415 – more should have been made in the discussion of the comparison between the work presented in the current paper and the results of Haynes and Pender (2007) since this is a very relevant study which would have provided really useful comparators.

We add more discussion in the text. See Lines 544-557 of the manuscript with Track Changes.

36. Line 419 – again I am not sure I agree with you here – the data presented in this paper has shown

that the effects of stress history are effectively cancelled out under higher flows in any subsequent flood. However, in line 419 you say 'might be more lasting during subsequent flood' – this seems counter to the rest of the message in the paper.

We rephrase the text to avoid misunderstanding. See Lines 544-557 of the manuscript with Track Changes. We appreciate the reviewer for the comment.

References

Chartrand, S. M., Jellinek, A. M., Hassan, M. A., and Ferrer-Boix, C.: Morphodynamics of a width-variable gravel bed stream: New insights on pool-riffle formation from physical experiments, Journal of Geophysical Research-Earth Surface, 123(11), 2735-2766, https://doi.org/10.1029/2017JF004533, 2018.

Chen, X., Hassan, M. A., An, C., and Fu, X.: Rough correlations: Meta-analysis of roughness measures in gravel bed rivers, Water Resources Research, 56, e2020WR027079, https://doi.org/10.1029/2020WR027079, 2020.

Chiew, Y.-M., and Parker, G.: Incipient sediment motion on non-horizontal slopes, Journal of Hydraulic Research, 32(5), 649-660, https://doi.org/10.1080/00221689409498706, 1994.

Church, M., Dudill, A., Venditti, J. G., and Frey, P.: Are Results in Geomorphology Reproducible? Journal of Geophysical Research-Earth Surface, 125(8), e2020JF005553, https://doi.org/10.1029/2020JF005553, 2020.

Church, M., Hassan, M. A., and Wolcott, J. F.: Stabilizing self-organized structures in gravel-bed stream channels: field and experimental observations, Water Resources Research, 34, 3169-3179, 1998.

Curran, J. C., and Wilcock, P. R.: Effect of sand supply on transport rates in a gravel-bed channel, Journal of Hydraulic Engineering, 131(11), 961–967, https://doi.org/10.1061/(ASCE)0733-9429(2005)131:11(961), 2005.

Hassan, M. A., Saletti, M., Johnson, J. P. L., Ferrer-Boix, C., Venditti, J. G., and Church, M.: Experimental insights into the threshold of motion in alluvial channels: sediment supply and streambed state, Journal of Geophysical Research-Earth Surface, e2020JF005736, https://doi.org/10.1029/2020JF005736, 2020.

Julien, P. Y.: Erosion and sedimentation, Cambridge University Press, New York (USA), 390 p, 2010.

Lamb, M. P., Dietrich, W. E., and Venditti, J. G.: Is the critical Shields stress for incipient sediment motion dependent on channel-bed slope? Journal of Geophysical Research-Earth Surface, 113, F02008, doi:10.1029/2007JF000831, 2008.

Masteller, C. C., and Finnegan, N. J.: Interplay between grain protrusion and sediment entrainment in an experimental flume, Journal of Geophysical Research-Earth Surface, 122, 274–289, https://doi.org/10.1002/2017GL076747, 2017.

Venditti, J. G., Dietrich, W. E., Nelson, P. A., Wydzga, M. A., Fadde, J., and Sklar, L.: Mobilization of coarse surface layers in gravel-bedded rivers by finer gravel bed load, Water Resources Research, 46, W07506, https://doi.org/10.1029/2009WR008329, 2010.

Wilcock, P. R., Kenworthy, S. T., and Crowe, J. C.: Experimental study of the transport of mixed sand and gravel, Water Resources Research, 37(12), 3349–3358, 2001.

Wong, M., and Parker, G.: Reanalysis and correction of bed-load relation of Meyer-Peter and Müller using their own database, Journal of Hydraulic Engineering-ASCE, 132, 1159–1168, doi:10.1061/(ASCE)0733-9429(2006)132:11(1159), 2006.

Zimmermann, A., Church, M., and Hassan, M. A.: Video-based gravel transport measurements with a flume mounted light table, Earth Surface Processes and Landforms, 33(14), 2285-2296, https://doi.org/10.1002/esp.1675, 2008.

Effect of stress history on sediment transport and channel adjustment in graded gravel-bed rivers

Chenge An[1,2], Marwan A. Hassan[2], Carles Ferrer-Boix[3], Xudong Fu[1]

[1] Department of Hydraulic Engineering, State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing, China.
[2] Department of Geography, The University of British Columbia, Vancouver, BC, Canada.
[3] Serra Húnter Fellow, Department of Graphic and Design Engineering~~Civil and Environmental Engineering~~, Technical University of Catalonia, Barcelona, Spain.

Correspondence to: Xudong Fu (xdfu@tsinghua.edu.cn)

Abstract. With the increasing attention on environmental flow management for the maintenance of habitat diversity and ecosystem health of mountain gravel-bed rivers, much interest has been paid to how inter-flood low flow can affect gravel-bed river morphodynamics during subsequent flood events. Previous research has found that antecedent conditioning flow can lead to an increase in the critical shear stress and a reduction in sediment transport rate during a subsequent flood. But how long this effect can last during the flood event has not been fully discussed. In this paper, a series of flume experiments with various durations of conditioning flow are presented to study this problem. Results show that channel morphology adjusts significantly within the first 15 minutes of the conditioning flow, but becomes rather stable during the remainder of the conditioning flow. The implementation of conditioning flow can indeed lead to a reduction of sediment transport rate during the subsequent hydrograph, but such effect is limited only within a relatively short time at the beginning of the hydrograph. This indicates that bed reorganization during the conditioning phase, which induce the stress history effect, is likely to be erased with increasing intensity of flow and sediment transport during the subsequent flood event.

1 Introduction

Prediction of sediment transport is of vital importance because it is related to many aspects of river dynamics and management, including river morphodynamics modeling (Parker, 2004), river restoration (Chin et al., 2009), aquatic habitats (Montgomery et al., 1996), natural hazard planning (Marston, 2008), bedrock erosion (Sklar and Dietrich, 2004), and landscape evolution (Howard, 1994). In ~~mountain~~ gravel-bed rivers, sediment transport is controlled by flow magnitude and flashiness, sediment supply, bed surface structures, channel morphology and the grain size distribution (GSD) of sediment (Montgomery and Buffington, 1997; Masteller et al., 2019). Therefore, prediction of sediment transport in mountain rivers still remains difficult despite the large body of existing theories. This is due to the fact that these theories were mostly developed for lowland streams with continuous sediment supply and an average flow regime, which do not apply to mountain streams (Gomez and Church, 1989; Rickenmann, 2001; Schneider et al., 2015).

For example, the hydrograph of mountain gravel-bed rivers is often characterized by large fluctuations of flow discharge, including both short-term flash flood and long-term inter-flood low flow (Powell et al., 1999). However, research on the morphodynamics of mountain rivers often focuses on the effects of floods (or constant high flow) and neglects the role of inter-flood low flow, with the consideration that most sediment transport and morphological adjustments of mountain rivers occur during relatively high flows (Klingeman and Emmett, 1982; Paola et al., 1992).

Reid and colleagues (Reid and Frostick, 1984; Reid et al., 1985) studied the effects of inter-flood low flow on subsequent sediment transport in Turkey Brook, England. They found that bedload transport rates were reduced during relatively isolated flood events (e.g., events separated by long time intervals) compared to those that were closely spaced, with the entrainment threshold up to as large as three times higher. They linked this with sediment reorganization during prolonged periods of antecedent flow, which can make the river bed more armored and more resistant to entrainment, thus delaying the onset of sediment mobility in the following flood event. Carling et al. (1992) also reported differences in the initial motion criteria between flood events due to changes in~~the~~ packing and orientation of sediment particles.

To further study such "memory" effects of antecedent flow on the sediment transport during a subsequent flood, a number of flume experiments as well as field surveys have been conducted in the past decade, and different terms have been proposed, including "stress history effect" (Monteith and Pender, 2005; Paphitis and Collins, 2005; Haynes and Pender, 2007; Ockelford and Haynes, 2013), "flood history effect" (Mao, 2018), "flow history" (Masteller et al., 2019)~~, etc~~. The difference in the terminology could be partly due to the available data and the chosen approach in different research works. ~~Given that all these terms are similar, h~~Here we adopt the term "stress history" in this paper. It should also be noted that the approach based on shear stress (and therefore terminology), even though widely applied for laboratory experiments, is much less reliable for field measurements.

Paphitis and Collins (2005) conducted flume experiments to study the entrainment threshold of uniform sediment subjected to antecedent flow durations of up to 120 minutes. They found that with a longer and higher antecedent flow, the critical bed shear stress increases and the total bedload flux decreases. The work of Paphitis and Collins (2005) was extended by Monteith and Pender (2005) and Haynes and Pender (2007) to consider bimodal sand-gravel mixtures. They found that for a graded bed, longer periods of antecedent flow increase bed stability due to local particle rearrangement, in agreement with Paphitis and Collins (2005); whereas higher magnitudes of antecedent flow reduce bed stability due to selective entrainment of the fine matrix on bed surface, counter to Paphitis and Collins' (2005) conclusion based on uniform sediment. Haynes and Pender (2007) further analyzed the two competing effects and concluded that particle rearrangement may be of greater relative importance than the winnowing of the fine sediment as it affects subsequent sediment transport. By using high resolution laser scanning and statistical analysis of the bed topography, Ockelford and Haynes (2013) also demonstrated that the response of bed topography to stress history is grade specific: bed roughness decreased in uniform beds but increased in graded bed with an increase length of an antecedent flow period. Performing a series of flume experiments, Masteller and Finnegan (2017) studied the evolution of the river bed on particle scale during low flow. They linked reduction of bedload flux to the re-organization of the highest protruding grains (1%-5% of the entire bed) on bed surface.

Because of the above-mentioned research, existing sediment transport formulae for gravel-bed rivers (e.g. Meyer-Peter and Müller, 1948; Parker, 1990; Wilcock and Crowe, 2003; Wong and Parker, 2006) are regarded to be inaccurate because they do not take the effect of stress history into account. To this end, Paphitis and Collins (2005) proposed an empirical formula for the exposure correction factor in the critical shear velocity for a uniform sand-size bed based on their experimental data. Johnson (2016) developed a state function for the critical shear stress in terms of transport disequilibrium, which incorporates the effects of stress history and hydrograph variability. Ockelford et al. (2019) proposed two forms of functions to link the antecedent duration and the critical shear stress. The two alternatives proposed by Ockelford et al. (2019) correct the function proposed by Paphitis and Collins (2005), whose exposure correction uses a logarithmic function which implicitly assumes an unbound growth as antecedent time tends towards infinity.

Research to date has shown that antecedent flow can stabilize the river bed, thus influencing the threshold of sediment motion as well as bedload flux. However, most of the previous research about stress history is either under conditions with relatively low sediment transport or with relatively short durations of sediment transport in order to capture the threshold of sediment motion (Monteith and Pender, 2005; Paphitis and Collins, 2005; Haynes and Pender, 2007; Ockelford and Haynes, 2013; Masteller and Finnegan, 2017; Ockelford et al., 2019). On the other hand, other researchers have found that exceptionally high discharge events can reduce critical shear stress by disrupting particle interlocking and breaking of bed structure (Lenzi, 2001; Turowski et al., 2009; Turowski et al., 2011; Yager et al., 2012; Ferrer-Boix and Hassan 2015; Masteller et al., 2019). Flume experiments by Masteller and Finnegan (2017) also indicate an increase in the number of highly mobile, highly protruding grains in response to sediment transporting flows. Therefore, the effect of high discharge events in reducing the critical shear stress likely counterbalances the stress history effect of antecedent flow to increase the critical shear stress. Besides, the supply of fine sediment (during high discharge events) is also widely observed to enhance the mobilization of coarse sediment (Wilcock et al., 2001; Curran and Wilcock, 2005; Venditti et al., 2010). In consideration of these opposing mechanisms, how long can the stress history effect last during a subsequent flood event is not well understood. Such a question is important especially in light of the fact that most sediment transport and channel adjustment of mountain gravel-bed rivers occurs during high discharge events, when the flow shear stress is high.

In this paper, flume experiments consisting of ~~extended cycles of~~ high and low flow are~~is~~ conducted to study this problem. The experimental arrangement is described in Sect. 2. In Sect. 3, we present the experimental results showing how channel morphology and sediment transport during a subsequent hydrograph respond to various durations of antecedent conditioning flow. The threshold of motion is analyzed in Sect. 4 based on the experimental data. Implications and limitations of this study are also discussed in Sect. 4. Finally, conclusions are summarized in Sect. 5.

2 Experimental arrangements

The experimental arrangements were guided by conditions observed in East Creek, a small mountain creek in Malcom Knob Forest, University of British Columbia (for details on the study site see Papangelakis and Hassan, 2016). To investigate the study objectives, we conducted flume experiments in the Mountain Channel Hydraulic Experimental Laboratory at the University of British Columbia. The experiments were conducted in a tilting flume with a length of 5 m, a width of 0.55 m and a depth of 0.80 m. The initial slope was 0.04 m/m. Water, but not sediment was recirculated by an axial pump. A set of six experiments (REF2 – REF7) was conducted; the experimental conditions are briefly summarized in Table 1. For experiments REF3 – REF7, the same hydrograph and sedimentograph were conducted, but with different durations of constant conditioning flow prior to the hydrograph/sedimentograph. It should be noted that in the experiments, we only implemented the rising limb of the hydrograph/sedimentograph, rather than a full hydrograph/sedimentograph with both rising and falling limbs. Rather than studying river adjustment during a flow hydrograph~~s~~, we aimed at determining the influence of conditioning time on~~in~~ bedload and bed surface arrangements as flow rates increased. We denote these as REF3 (10), REF4 (2), REF5 (5), REF6 (15) and REF7 (0.25), with the numbers in the brackets denoting the duration of the conditioning flow in hours. Experiment REF2 (15) consists of a 15-hour conditioning period without a subsequent hydrograph/sedimentograph, to test the reproducibility of our experimental results during the conditioning flow.

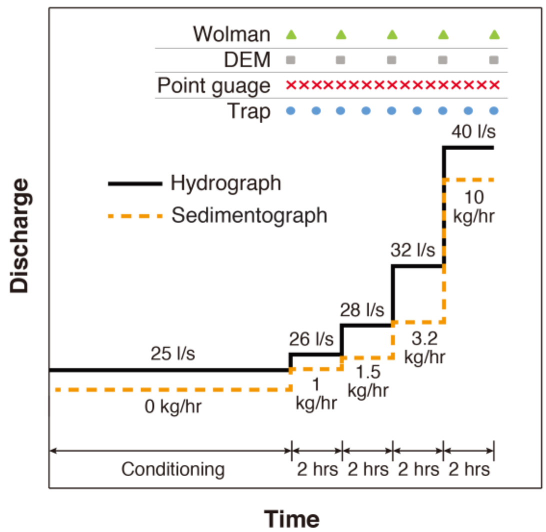
Table 1. Summary of the experimental conditions and measurements. The experiments are listed in the table in order of decreasing duration of conditioning flow.

No.	Phase	Duration (h)	Flow discharge (l/s)	Water surface slope (%)	Flow depth (cm)	Froude number (-)	τ_b (Pa)	Δz_b (mm)	Sediment feed (kg/h)	D_{s50} (mm)	D_{s90} (mm)	D_{l50} (mm)	D_{l90} (mm)	τ^*_{s50}	Q_s (kg/h)
REF2 (15)	Conditioning	15	25	2.62	6.33	0.91	16.27	-30.2	0	~~15.2~~15.5	~~29.6~~29.7	1.07	5.43	~~0.069~~0.065	0.27
REF6 (15)	Conditioning	15	25	3.27	6.47	0.88	20.76	-16.6	0	~~15.87~~15.7	~~30.69~~30.8	35.18	42.84	~~0.089~~0.082	0.89
	Step 1	2	26	3.34	6.39	0.94	20.93	0.3	1	~~15.66~~14.4	~~29.98~~30.0	12.51	39.38	~~0.083~~0.090	0.68
	Step 2	2	28	3.10	6.29	1.03	19.13	0.0	1.5	~~17.18~~17.3	~~30.40~~29.4	7.28	27.59	~~0.069~~0.068	0.76
	Step 3	2	32	3.06	6.80	1.05	20.41	-1.9	3.2	~~15.34~~16.2	~~30.85~~31.8	12.39	36.54	~~0.082~~0.078	6.73
	Step 4	2	40	2.81	7.78	1.07	21.45	-16.1	10	~~15.95~~15.9	~~30.34~~31.6	11.48	36.03	~~0.083~~0.083	13.39
REF3 (10)	Conditioning	10	25	2.73	6.02	0.98	16.12	-25.8	0	~~14.9~~14.8	~~29.5~~29.2	2.17	9.98	~~0.071~~0.067	0.28
	Step 1	2	26	2.75	5.93	1.04	16.00	0.1	1	~~15.0~~15.6	~~29.3~~29.5	2.55	19.94	~~0.066~~0.063	1.71
	Step 2	2	28	2.69	6.35	1.01	16.77	0.3	1.5	~~15.5~~15.8	~~29.7~~30.2	4.06	26.99	~~0.067~~0.065	2.19
	Step 3	2	32	2.88	6.81	1.04	19.25	-1.7	3.2	~~15.9~~15.9	~~29.7~~30.1	6.18	24.26	~~0.075~~0.075	2.44
	Step 4	2	40	2.48	8.34	0.96	20.28	-8.0	10	~~15.6~~14.2	~~32.8~~32.8	14.45	39.13	~~0.080~~0.088	12.45
REF5 (5)	Conditioning	5	25	3.26	5.51	1.12	17.63	-16.8	0	~~16.35~~15.3	~~31.14~~32.0	8.23	25.34	~~0.066~~0.071	0.49
	Step 1	2	26	3.24	6.19	0.98	19.68	-0.6	1	~~16.30~~15.4	~~30.90~~31.5	6.57	23.63	~~0.075~~0.079	2.24
	Step 2	2	28	3.09	6.21	1.05	18.82	-0.3	1.5	~~16.87~~17.2	~~31.27~~31.4	9.38	28.44	~~0.069~~0.067	3.30
	Step 3	2	32	3.05	6.65	1.08	19.91	-1.2	3.2	~~16.04~~16.8	~~31.04~~31.9	11.90	47.91	~~0.077~~0.073	5.72
	Step 4	2	40	2.78	7.82	1.06	21.33	-13.4	10	~~14.72~~15.1	~~31.44~~34.5	15.09	38.56	~~0.090~~0.087	40.03

REF4 (2)	Conditioning	2	25	2.82	5.55	1.11	15.34	-17.8	0	~~13.58~~12.3	~~28.78~~27.8	3.10	15.79	~~0.070~~0.077	1.50
	Step 1	2	26	2.73	5.55	1.16	14.85	-0.5	1	~~14.61~~14.8	~~28.83~~28.9	3.90	20.31	~~0.063~~0.062	0.96
	Step 2	2	28	2.71	6.19	1.06	16.46	-0.1	1.5	~~15.44~~15.6	~~29.09~~29.2	6.28	46.76	~~0.066~~0.065	2.41
	Step 3	2	32	3.15	6.85	1.04	21.15	-6.4	3.2	~~14.44~~14.5	~~28.65~~28.8	17.34	37.76	~~0.091~~0.090	26.73
	Step 4	2	40	2.76	8.01	1.02	21.69	-7.7	10	~~14.06~~13.7	~~30.22~~29.7	10.88	35.45	~~0.095~~0.098	5.23
REF7 (0.25)	Conditioning	0.25	25	3.46	6.20	0.94	21.06	-14.9	0	~~14.67~~14.0	~~30.10~~29.5	10.54	28.03	~~0.089~~0.093	19.44
	Step 1	2	26	3.20	6.54	0.90	20.53	-4.8	1	~~15.52~~15.6	~~30.86~~31.6	7.11	28.91	~~0.082~~0.081	3.48
	Step 2	2	28	3.14	6.58	0.96	20.27	-0.7	1.5	~~16.62~~16.2	~~31.33~~31.2	6.91	30.73	~~0.075~~0.077	2.52
	Step 3	2	32	3.12	7.00	1.00	21.41	-4.5	3.2	~~14.89~~14.3	~~30.78~~30.5	10.09	37.40	~~0.089~~0.092	12.32
	Step 4	2	40	2.73	8.29	0.97	22.19	-9.6	10	~~17.68~~17.3	~~36.20~~33.6	12.13	30.78	~~0.078~~0.079	16.80

a. Q_s: bedload transport rate, Δz_b: mean difference of bed elevation averaged over the whole river channel, τ_b: shear stress, D_{s50} and D_{s90}: D_{50} and D_{90} of bed surface, D_{l50} and D_{l90}: D_{50} and D_{90} of bedload, τ^*_{s50}: Shields number for D_{s50}. Here D_{90} denotes the grain size such that 90% is finer, and D_{50} denotes the grain size such that 50% is finer. All values presented in this table are measured at the end of each stage, except for Δz_b which denotes the mean difference of bed elevation during each stage (i.e., difference between the end of this stage and the end of last stage). A positive value of Δz_b denotes aggradation, and a negative value of Δz_b denotes degradation.

Figure 1 shows the water and sediment supply implemented during~~in~~ the experiments. The water discharge was selected to represent typical flows in East Creek, with the 25 l/s flow during the conditioning period being equivalent to half the bankfull flow, and the peak flow discharge of 40 l/s during the hydrograph being about 1.1 times the bankfull flow in East Creek. Because the purpose of this paper is to study the evolution of bed stability, sediment was not fe~~e~~d during the conditioning flow. For each step of the hydrograph, the feed rate of sediment was specified to be close to the transport capacity of the flow. Determination of the sediment supply rates was facilitated by a numerical model which was~~had been~~ calibrated for~~with~~ similar experimental conditions (Ferrer-Boix and Hassan, 2014).~~we chose a feed rate through numerical simulations following Ferrer-Boix and Hassan (2014) in combination with trial experiments.~~ Sediment was fed into the flume at the upstream end using a conveyor belt feeder at the calculated transport rate capacity. The feed rate of the sedimentograph ranged between 1 ~~kg/hour~~kg/h and 10 ~~kg/hour~~kg/h. Both the hydrograph and the sedimentograph consisted of four steps, with each step lasting for 2 hours.



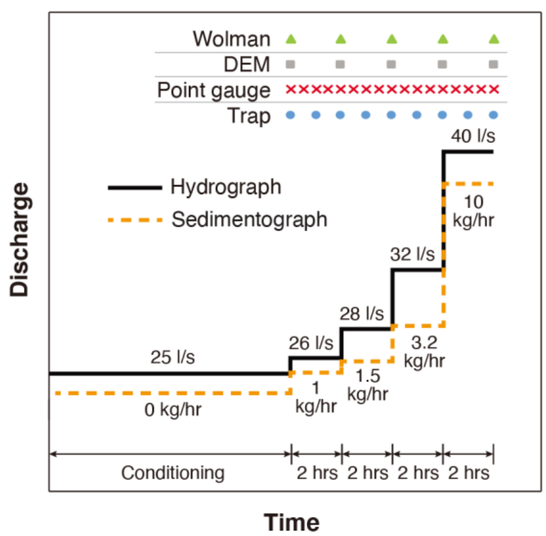


Figure 1. Water and sediment supply implemented in the experiments. Markers in top of the figure denote the time of measurements during the hydrograph phase. Time of measurements during the conditioning phase is not shown in this figure.

Figure 2 shows the GSD of the bulk sediment used in the experiments, with the grain size ranging between 0.5 and 64 mm. The GSD was scaled from East Creek by a ratio of 1:4, except that sediment (after scaling) with a grain size less than 0.5 mm was excluded. This preserved the entire gravel distribution of East Creek with a maximum size of 256 mm (scaled to 64 mm in Fig. 2). The model was "generic" rather than specific. This means ~~in~~ that no attempt was made to reproduce the geometric details of the prototype channel. The bulk sediment was sieved at~~in~~ half φ intervals and each grain size class was painted in different colors for ~~each size class for~~ texture analysis and visual identification. Before the commencement of each experiment, we hand-mixed and ~~screeded~~ leveled the bulk sediment to make a flat and uniform layer of loose material with a depth of 0.15 m. The sediment was then slowly flooded and then drained to aid settlement. The bulk sediment wa~~i~~s also used for the sediment feed in each experiment.

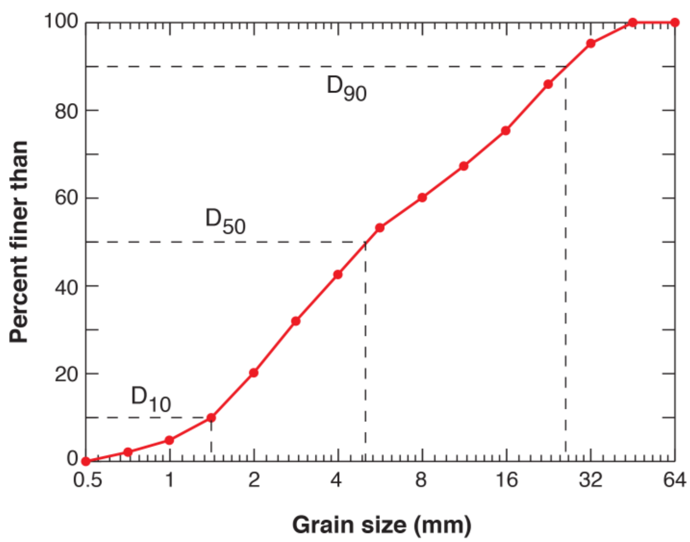


Figure 2. Grain size distribution of the bulk sediment used in the experiments.

The ~~elevations of the~~ bed ~~surface~~ and water surface elevations were measured along the flume every 0.25 m using a mechanical point gauge with a precision of ±0.001 m. Water depth fluctuations due to wave effects at a point were about 5% or less. Water surface slope and bed slope are calculated based on a linear regression of the point gauge data measured between 0.5 m and 4.75 m upstream of the outlet. The most upstream and downstream sections are excluded to avoid boundary effects. A green laser scanner mounted on a motorized cart was also used to measure the bed surface elevation along the flume. Bed laser scans were composed of cross sections spaced 2 mm apart with 1 mm vertical and horizontal accuracy (for details see Elgueta-Astaburuaga and Hassan, 2017). The standard deviation of bed elevation was calculated based on the DEM data from scans. Before the calculation of standard deviation, the DEM was detrended based on linear regression to remove spatial trends with scales larger than the scale of sediment patterns (e.g., bed slope or undulations). To estimate the particle size distribution of the bed surface we used digital cameras mounted on a motorized cart along the entire flume. Images were merged together to visualize the bed and

~~preform~~ perform the particle size analysis (Chartrand et al., 2018). To avoid ~~the~~ distortion effects due to image merging, the width of the image strips that were stitched to get a composite image was specified as just 2 cm. The particle size distribution of the bed surface was estimated using the Wolman (point count) method, by identifying the grain size of particles at the intersections of a 5 cm grid superimposed on the photograph.~~The particle size distribution of the bed surface was estimated using the grid by number (point counts) method, by identifying particle size at the intersection of a 5 cm grid superimposed on each photograph.~~ Individual grains were identified by color. ~~Collected data~~For each experiment, the grain size distribution of the bed surface was calculated at different times to quantify its changes during the experiment.~~ were used to quantify changes in the bed surface particle size distribution throughout each experiment.~~

~~Material evacuated from the flume was trapped in a 0.25 mm mesh screen in the tailbox, and weighted and sieved at half φ intervals to calibrate a light table.~~ The sediment transport rates for various size ranges were measured at the end of the flume using a light table (for details see Zimmerman et al., 2008; Elgueta-Astaburuaga and Hassan 2017) and automated image analysis at a resolution of 1 second~~ (for details see Zimmerman et al., 2008; Elgueta-Astaburuaga and Hassan 2017)~~. Material evacuated from the flume was ~~also~~ trapped in a 0.25 mm mesh screen in the tailbox, ~~and~~ weighted and sieved at half φ intervals, and then used to calibrate the light table data. To avoid random fluctuations in sediment transport, we report the bedload transport rate measured by light table at a 5-minute resolution, and characteristic grain sizes of bedload at 15-minute resolution. A range of methods for the estimation of bed shear stress has been suggested in the literature (reviewed in Whiting and Dietrich, 1990). In this study, the shear stress is estimated using the depth-slope product corresponding to normal (steady and uniform) flow. This method is selected because the focus of this work is on overall (mean) parameters controlling bed evolution; in addition, the water was too shallow to use an ADV. The water surface slope, rather than bed slope, is implemented in the calculation of shear stress, with the consideration that water surface slope is closer to the friction slope and also has less random fluctuations than bed slope.

The frequency of measurements during the hydrograph phase is also plotted in Fig. 1(a), with the point gauge measurements conducted every 30 minutes, the trap weighting/sampling conducted every hour, and the DEM/Wolman measurements by laser scan/photograph conducted every 2 hours (i.e. at the beginning/end of each stage of the hydrograph). For each measurement of DEM/Wolman, we stopped the pump instantaneously and let the flow ~~was~~ slowly lower~~ed~~ and then stop~~ped~~ to allow for the bed to be scanned by a laser and photographed. The time interval between the stop of the pump and the stop of the flow was about 3 to 4 minutes. To avoid the influence of the following rising discharge, all subsequent measurements were taken after the flow became stable. The frequency of measurement during the conditioning phase was adjusted in each experiment in accordance with the duration of the conditioning phase, and is therefore not plotted in Fig. 1~~(a)~~.

The uncertainties ~~of~~ associated with the measurement are also studied. For the uncertainties of the standard deviation of bed elevation, we scanned the floor of the flume twice and calculated the standard deviations of the scanned DEM. The floor of the flume was horizontal and flat, with no sediment on the bed. Theoretically, the standard deviation of the DEM should be zero. Therefore, the calculated standard deviations of the flume floor are regarded as an estimation of the uncertainties of our calculations during experiment. To estimate the uncertainties of the bed surface GSD, for each measurement the Wolman method was implemented ~~for~~ 5 times on the same photograph, with 100 samples/counts ~~for~~ each time. The 5 measured GSDs for each time interval were used to calculate the mean and standard deviation of the bed surface texture (in terms of D_{s10}, D_{s50}, and D_{s90}). To estimate the uncertainties of the light table method, we compare the data measured by the light table with the data measured by the sediment trap, in terms of both sediment transport rate and the characteristic grain sizes of sediment load. To estimate the variations of the measured/calculated data, we calculate their coefficient of variation (cv), ~~which is~~ defined as the ratio of the standard deviation to the mean value.

3 Experimental results

Table 1 presents an overall schematization of the experimental results, including water surface slope, flow depth h, Froude number F_r ($F_r = u / (gh)^{0.5}$), where u is depth-averaged flow velocity), bedload transport rate Q_s, shear stress τ_b, D_{50} and D_{90} of bed surface (D_{s50} and D_{s90}), D_{50} and D_{90} of bedload (D_{l50} and D_{l90}), and Shields number τ^*_{s50} for a given D_{s50}. Here D_{90} denotes the grain size such that 90% is finer, and D_{50} denotes the grain size such that 50% is finer.

3.1 Channel adjustment

In this section, we present the channel adjustments during each experiment. Figure 3 shows the difference of longitudinal DEM averaged over the cross section, which can represent the adjustment of channel topography during different periods of the experiment. The DEM averaged over the cross section is used here to study the overall aggradation/degradation of the channel. For reference, detailed information about~~of~~ the DEM at different times during the experiment is provided in the Supporting Information, with REF6 (15) as an example. From Fig. 3(a) we can see that for each experiment, evident degradation occurs during the first 15 minutes, especially at the upstream end of the flume. This is due to the fact that no sediment supply is implemented during the conditioning period, and also the initial bed material is relatively loose. From 15 minutes until the end of the conditioning phase (as shown in Fig. 3(b)), no evident aggradation/degradation is observed for any experiment, indicating that most of the adjustment of channel topography during the conditioning phase has been accomplished within the first 15 minutes. For Step 1 of the hydrograph (as shown in Fig. 3(c)), no evident aggradation/degradation is observed for any of the experiments (with the mean difference of bed elevation Δz_b less than ±1 mm, as shown in Table 1), except for REF7 (0.25)~~,~~ which has the shortest conditioning phase and experienced a mean degradation of 4.8 mm over the whole bed channel. Similarly, the channel keeps relatively stable during Step 2 of the hydrograph for all experiments (as shown in Fig. 3(d)), with no evident~~trend for~~ aggradation/degradation being observed (the mean difference of bed elevation Δz_b is less than ±1 mm for all experiments). With the increase of flow discharge, some degradation (with a magnitude of about 10 ~ 20 mm) can be observed in Step 3 for all experiments at the upstream end of the channel, as shown in Fig. 3(e). Such degradation becomes more evident over the entire channel in Step 4 of the hydrograph, when flow discharge reaches its peak value. This is in agreement with the values of Δz_b presented in Table 1. ~~Further analysis of the DEM data shows that no bedform were evident during the experiment.~~

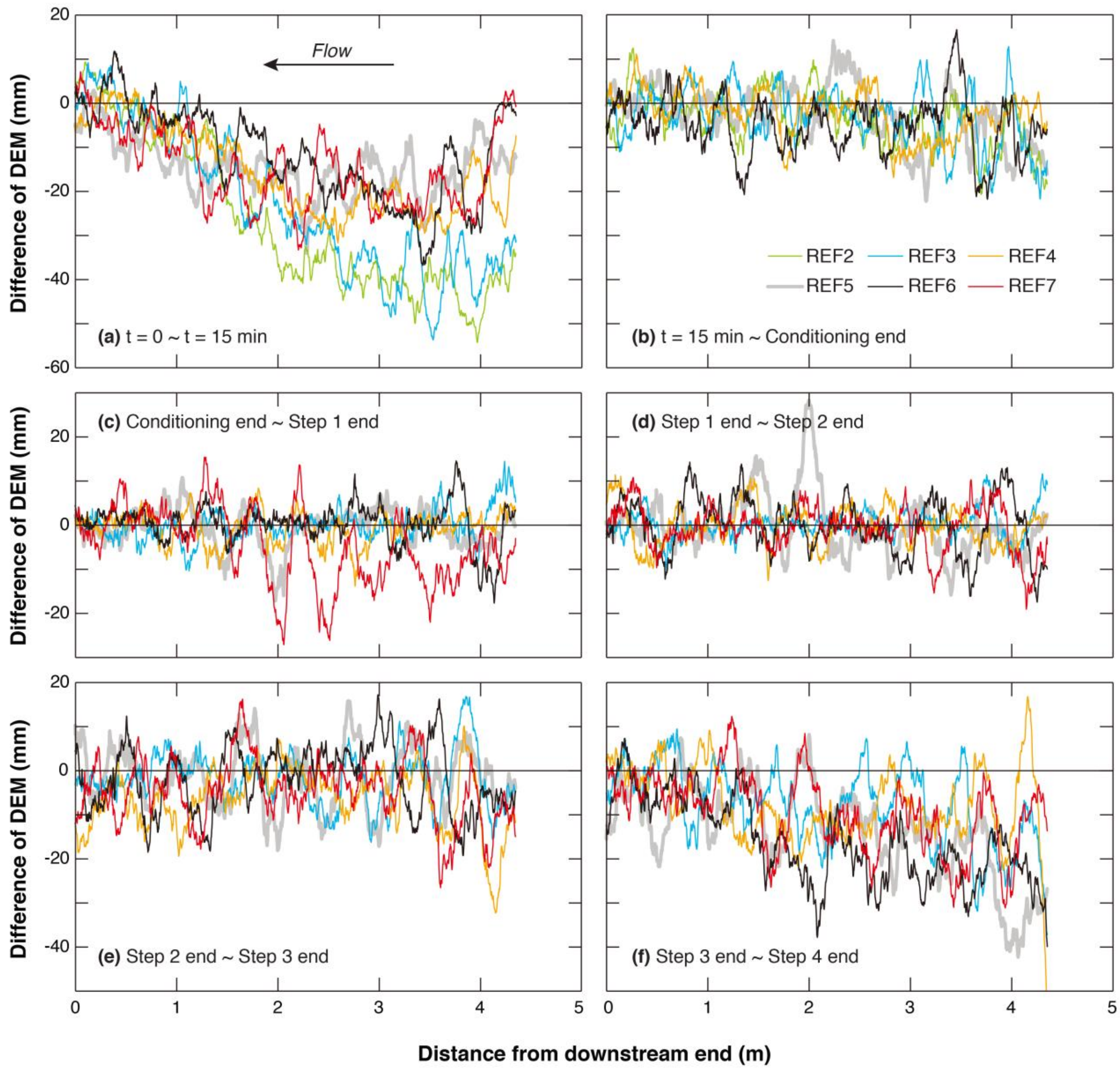
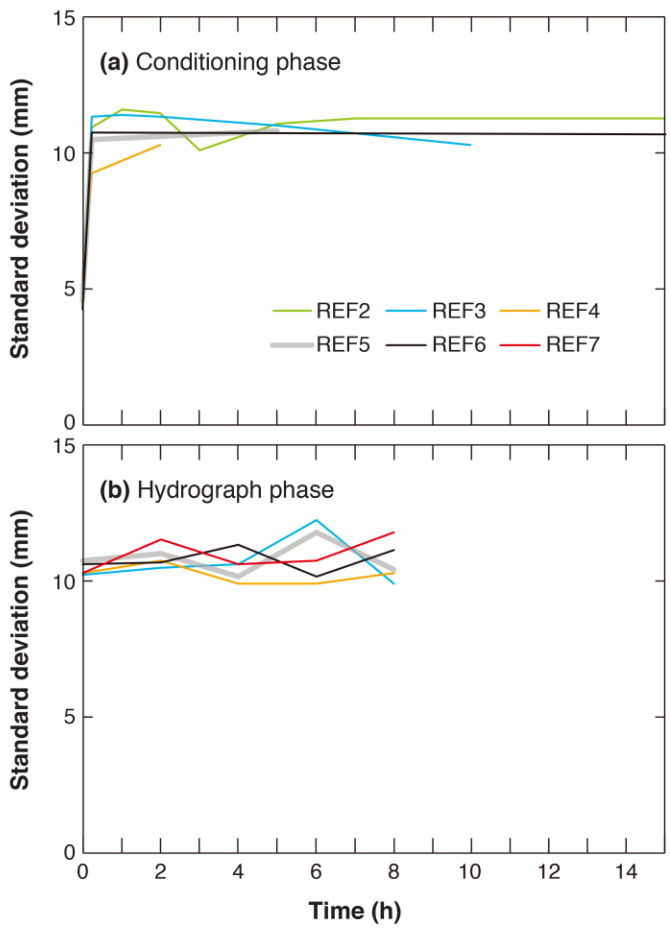


Figure 3. Spatial distribution of elevation difference from cross-sectionally averaged longitudinal DEM during the experiment: (a) from beginning of experiment to t = 15 minutes; (b) from t = 15 minutes to the end of conditioning phase; (c) from the end of conditioning phase to the end of Step 1 of hydrograph phase; (d) from the end of Step 1 to the end of Step 2 of the hydrograph phase; (e) from the end of Step 2 to the end of Step 3 of the hydrograph phase; (f) from the end of Step 3 to the end of Step 4 of the hydrograph phase.

Figure 4 shows the temporal variation of the standard deviation of bed elevation, which is often scaled with the bed roughness for gravel-bed rivers (see Chen et al. (2020) for a detailed discussion on this topic), over the length of the erodible bed during the experiment. Results show that the standard deviation of bed elevation is relatively small at the beginning of the experiments (corresponding to a relatively smooth bed depending on the way we prepared the initial bed), but increases notably within 15 minutes after the start of the conditioning phase. Such an increase of the ~~bed roughness~~standard deviation of bed elevation is accompanied by significant degradation during the first 15 minutes, as shown in Fig. 3(a). The standard deviation of bed elevation ~~remains almost constant~~becomes quite stable during the remaining conditioning phase, as well as during the hydrograph phase, despite the fact that degradation is evident as the flow approaches its peak value. For the standard deviation of bed elevation during the conditioning phase, we calculate the coefficient of variation (cv) for REF2 (15), which has the longest conditioning phase. T~~, and~~ ~~t~~he result shows a value of 0.038 from *t* = 15 minutes to the end of conditioning flow. For the standard deviation of bed elevation during the hydrograph phase, we calculate the cv for all experiments;~~, and~~ the results show~~s~~ that the values of cv vary between 0.031 and 0.075. Besides, the value of standard deviation is almost identical for each experiment, indicating the period of conditioning phase exerts little effect on the standard deviation of bed elevation.



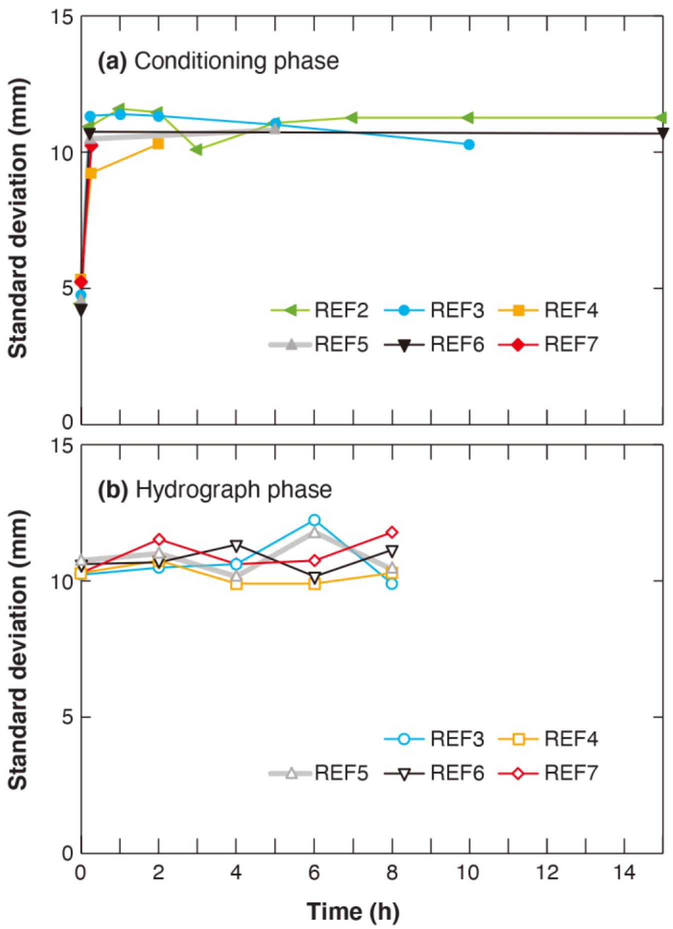
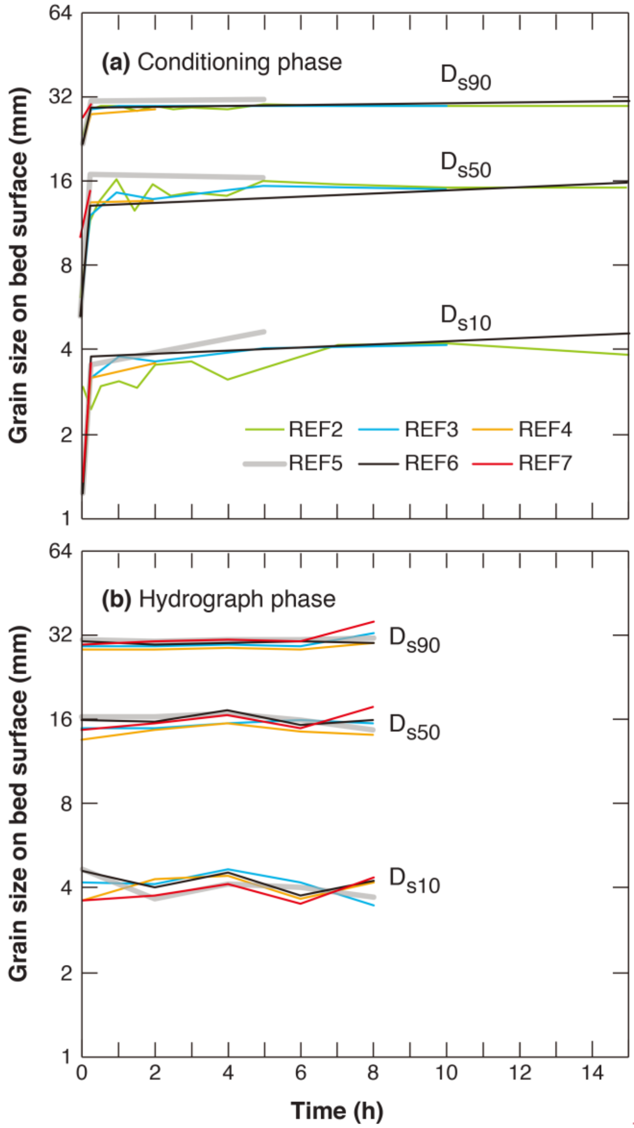


Figure 4. Temporal adjustments of standard deviation of bed elevation calculated over the whole erodible bed: (a) the conditioning phase; (b) the hydrograph phase. The uUncertaintyies of the calculation is in the range of 1.6~2.5 mm, which isare close to the vertical resolution of the laser (1 mm).

Figure 5 shows the temporal variation of the characteristic grain size of bed surface material, as well as an estimation of the uncertaintiesy of associated with measurements of the surface texture. Three parameters are presented here; D_{s10}, D_{s50}, and D_{s90}. The adjustment of bed surface GSD follows similar trends as the adjustment of standard deviation of bed elevation. That is, Ffor all experiments, the bed surface is fine at the beginning, and experiences a fast coarsening period during the first 15 minutes (along with the bed degradation in Fig. 3 and the increase of bed roughness in Fig. 4). The characteristic grain sizes of bed surface remain relatively stable after the first 15 minutes, despite variabilities due to the measurement uncertainty. For REF2 (15) which has the longest conditioning phase, cv (coefficient of variation) values of the mean D_{s10}, D_{s50}, and D_{s90} (over the five repeated measurements) are 0.15, 0.09, and 0.02 respectively from t = 15 minutes to the end of the conditioning flow. It is worth noted noting that the GSD of bed surface keeps relatively constant even during the hydrograph phase, during which a flood event is introduced in the flume and evident bed degradation is observed. For each experiment, the cv values of the mean D_{s10}, D_{s50}, and D_{s90} (over the five repeated measurements) are less than 0.13, 0.08, and 0.04

respectively during the hydrograph phase.This is in agreement with the observation of Ferrer-Boix and Hassan (2015) during successive water pulses.



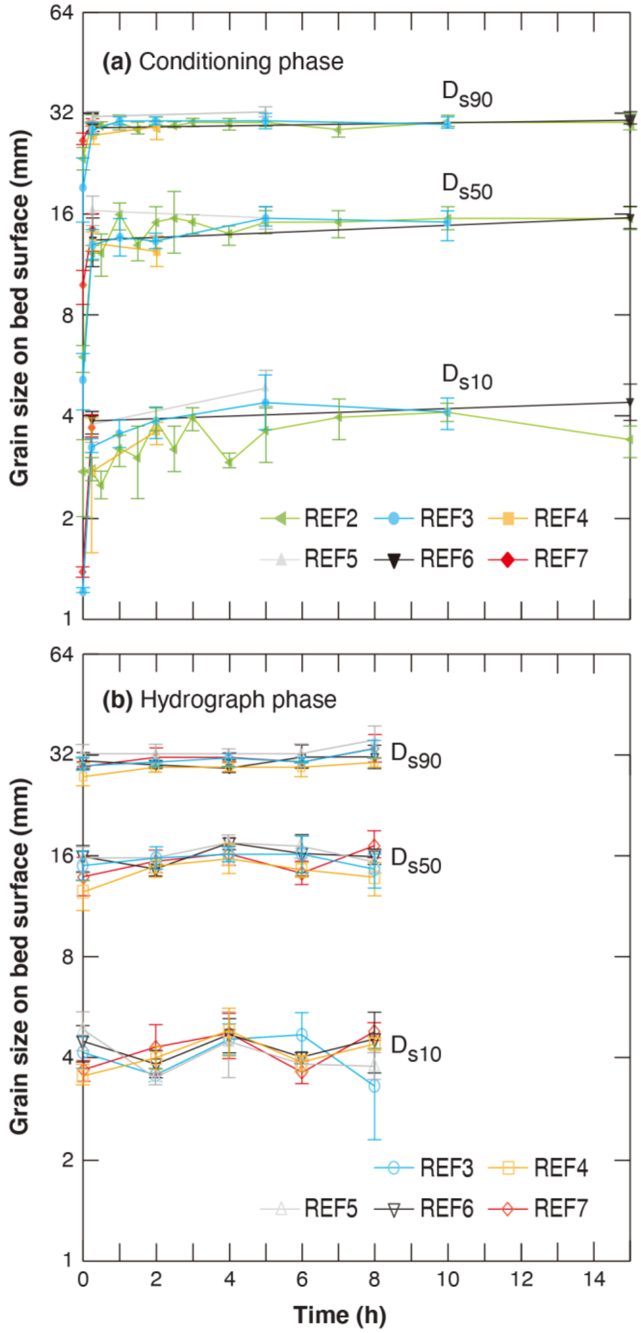


Figure 5. Temporal adjustments of characteristic grain sizes of bed surface material calculated over the whole erodible bed: (a) the conditioning phase; (b) the hydrograph phase. Markers show mean values of five repeated Wolman measurements. Range bars show the mean values ± the standard deviations of the five repeated Wolman measurements.

3.2 Sediment transport

In Fig. 6 we ~~exhibit~~ present the instantaneous sediment transport rate Q_s measured by the light table ~~in~~ during each experiment. Sediment transport is reported every 5 minutes, as described in Sect. 2. Accuracy of the results is estimated by comparing the light table data with the data measured by the trap. Results show that for our experiments, the light table method has good accuracy in terms of the sediment transport rate, with an overestimation by 4% on average (111 samples and a standard deviation of 14.5%). 70 out of 111 samples show an accuracy of ±10%, and 93 out of 111 samples show an accuracy of ±20%. Details of this uncertainty analysis are presented in the Supporting Information.

It can be seen in Fig. 6(a) that the temporal variation of sediment transport rate during the conditioning phase follows the same trend in all six experiments. That is, the sediment transport rate decreases significantly during the conditioning phase, with the decreasing rate being very large at the beginning and then gradually dropping. In the first 15 minutes, the sediment transport rates drop from more than 500 ~~kg/hour~~kg/h to less than 100 ~~kg/hour~~kg/h. Afterwards, it takes about another 2 hours for the sediment transport rates to drop to close to 1 ~~kg/hour~~kg/h. The sediment transport rate eventually approaches a small and relatively constant value after about 8 hours of conditioning flow. For REF2 (15) and REF6 (15) which have the longest conditioning phase, the sediment transport rates between t = 8 hour and the end of conditioning phase (t = 15 hour) show mean values of 0.35 ~~kg/hour~~kg/h (standard deviation = 0.22 kg/h) and 0.37 ~~kg/hour~~kg/h (standard deviation = 0.24 kg/h), respectively. Nevertheless, there are random high points in the sediment transport rate even after 8 hours, despite no sediment feed from the inlet. These spikes imply that partial destruction (or reorganization) of the bed structure occurs even after a long duration of conditioning.

Previous researchers (Haynes and Pender, 2007; Masteller and Finnegan, 2017) have suggested that an exponential function can be implemented to describe such a decrease of sediment transport rate under conditioning flow. Additional analysis is implemented in the Supporting Information to fit REF2 (15) and REF6 (15) (which have the longest duration of conditioning phase) against a two-parameter exponential function. Results show that the exponential function can describe the general decreasing trend of sediment transport rate during the conditioning phase, except at the beginning of the experiment where the decrease of sediment transport rate is much more significant than that predicted by the exponential function. Readers can refer to the Supporting Information for more details.

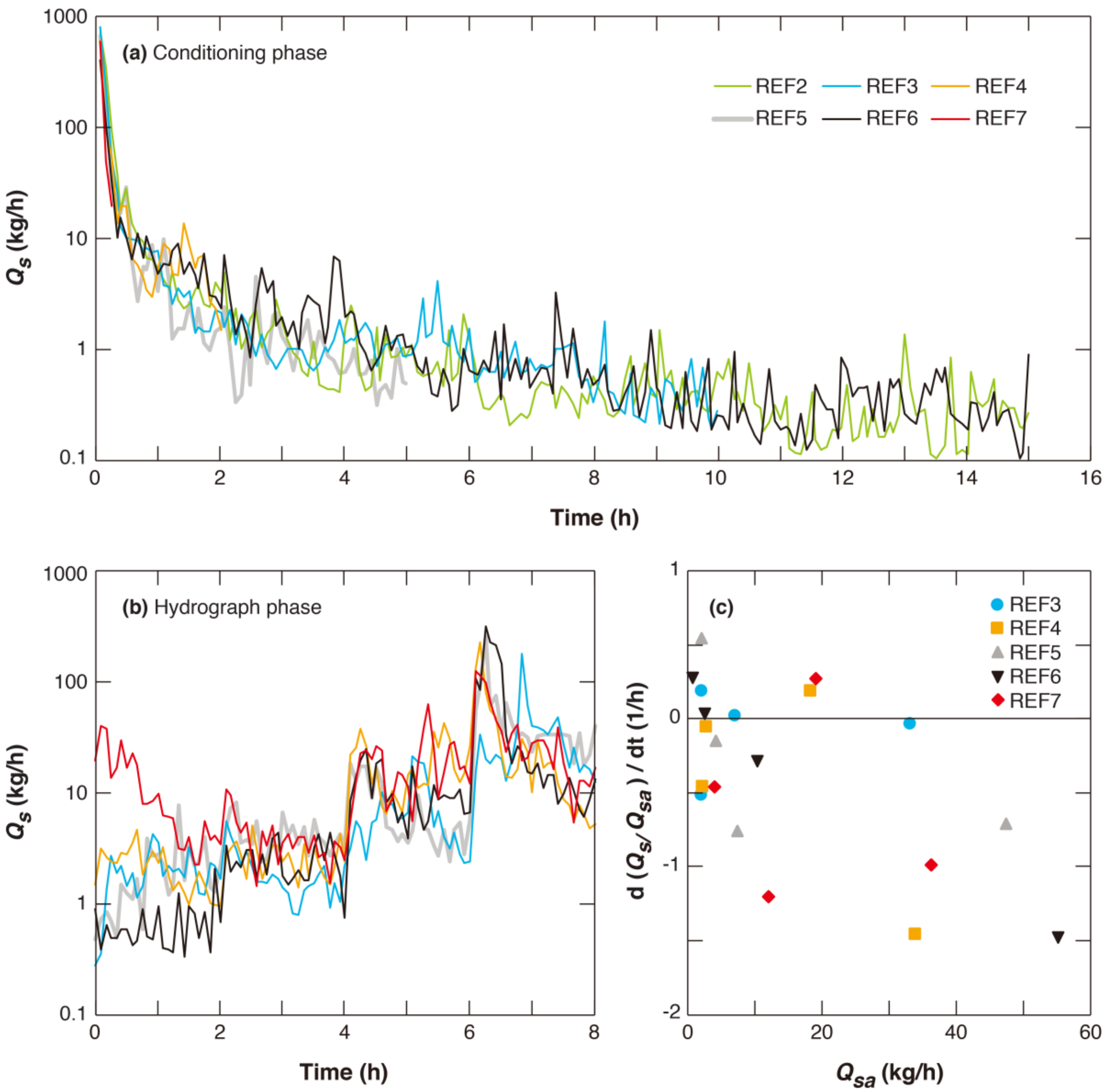


Figure 6. Instantaneous sediment transport rate measured by light table during (a) the conditioning phase; and (b) the hydrograph phase. (c) Intra-step temporal change rate of Q_s normalized against Q_{sa} for each hydrograph step. Q_s is the sediment transport rate, and Q_{sa} is the averaged sediment transport rate of a given hydrograph step.

Figure 6(b) presents the instantaneous sediment transport rate during the hydrograph phase. Results show that variation of sediment transport rate among different experiments prevails in the first step of the hydrograph, with the highest sediment transport rate for the experiment with the shortest conditioning duration (REF7 (0.25)); and the smallest sediment transport rate for the experiment with the longest conditioning duration (REF6 (15)). Such variation among experiments, however, diminishes towards the end of Step 1 and is not observed in the following three steps of the hydrograph, with the line for each experiment collapsing together in the figure. ~~The~~ Such adjustments of sediment transport rate are consistent~~agree~~ with the process of channel deformation shown in Fig. 3~~,~~. That is, for both sediment transport and channel deformation, ~~where the pattern of variation in~~ results of REF7 (0.25) deviate~~s~~ from other experiments in Step 1 (larger sediment transport rate and more degradation in REF7 (0.25)), but collapse~~s~~ with other experiments in the following three steps.

Results in Fig. 6(b) also show large variations of sediment transport rate during each step of the hydrograph. Such intra-step variations of sediment transport rate are investigated in Fig. 6(c), with the x axis being the averaged sediment transport rate of each step Q_{sa} and the y axis being $d(Q_s/Q_{sa})/dt$.~~, which~~ The value of $d(Q_s/Q_{sa})/dt$ is estimated by linear regression. Here the instantaneous sediment transport rate Q_s is scaled against the average sediment transport rate of the corresponding step Q_{sa}, in order to facilitate the comparison among different hydrograph steps.

Results in Fig. 6(c) shows that a large fraction of the data (11 out of 20) exhibits a decreasing trend in time for Q_s (i.e. a negative value in vertical coordinate). Basically, the larger the averaged sediment transport rate Q_{sa}, the larger ~~is~~ the rate of reduction in Q_s. Ferrer-Boix and Hassan (2015) observed similar declines in sediment transport during their water pulses experiments. They attributed this to (1) the presence of bed structures, which could have reduced skin friction up to 20% and (2) streamwise changes in the patterns of bed surface sorting. Out of 20 datasets, 5 exhibit some temporally increasing trend in Q_s (though not as evident as the decreasing trend mentioned before). They are REF5 (5), REF3 (10), REF6 (15) during the first step; and REF7 (0.25), REF4 (2) during the third step. This shows that for the three experiments with long conditioning duration, Q_s is very low at the end of the conditioning phase, and the first step of the hydrograph sees a temporally increasing trend in Q_s. Whereas for the two experiment with short conditioning phase, Q_s is still high at the end of the conditioning, so that the sediment transport rate keeps decreasing during the first step, until in the third step an increasing trend in Q_s is observed, at which the water and sediment supply become evidently higher. The decreasing/increasing trends of Q_s during steps of the hydrograph reflect the transient adjustments of the bed to the changed water and sediment supply before equilibrium is achieved.

Sediment collected in the trap/tailbox at the flume outlet allows us to plot the total amount of sediment output during each step of the hydrograph. ~~To better understand the effect of the conditioning duration on sediment transport, we calculate the cumulative sediment transport during the entire hydrograph phase as well as each step of the hydrograph.~~ Fig. 7(a) shows ~~that~~ the total sediment output during the entire hydrograph. It can be seen that the effect of conditioning duration on the total sediment output during the entire hydrograph phase is not evident: a longer duration of conditioning flow does not necessarily lead to a smaller (or larger) sediment output. The largest sediment output occurs in REF7 (0.25), which is 55% larger than the sediment output in REF3 (10) which has the smallest output, but is about the same as (only 4% larger than) the sediment output in REF6 (15). We further calculate the correlation coefficient between the total sediment output and the duration of conditioning flow, and obtain a value of r = -0.14, indicating that there is almost no correlation between the two parameters.~~ does not show much difference for each experiment, indicating that the duration of conditioning flow does not pose much influence on the total volume of sediment transport during the subsequent flood.~~

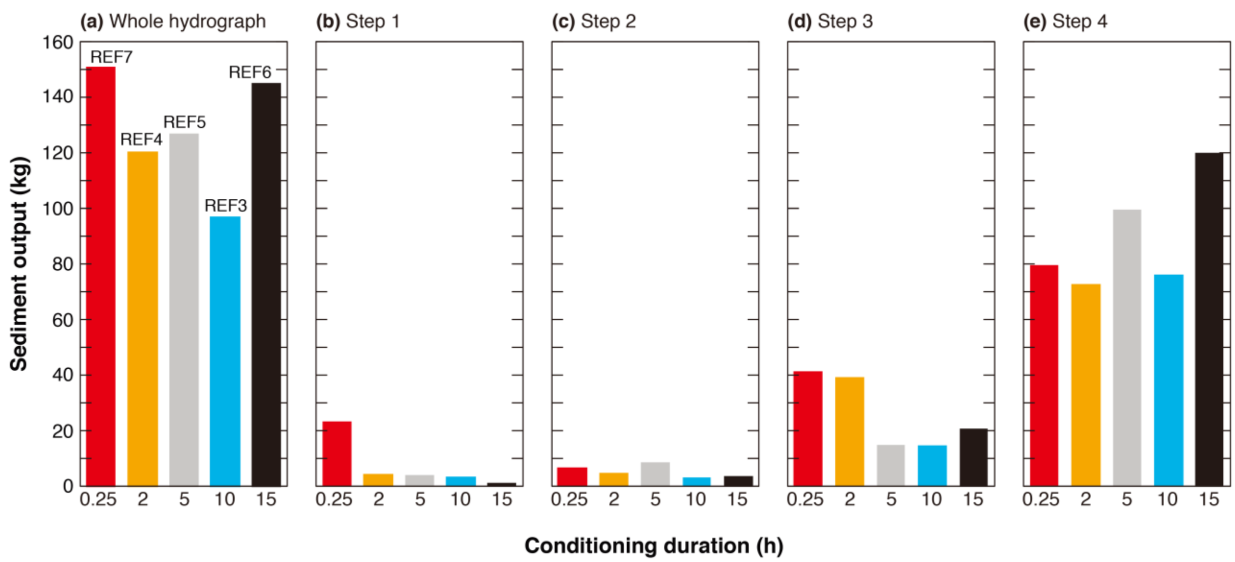


Figure 7. Sediment output measured at a trap during (a) the whole hydrograph; (b) Step 1 of the hydrograph; (c) Step 2 of the hydrograph; (d) Step 3 of the hydrograph; (e) Step 4 of the hydrograph.

However, if we study the sediment transport during each step of the hydrograph, we can find that in Step 1 REF7 (0.25) has much larger sediment output than the other experiments, as shown in Fig. 7(b). For Step 1, the sediment output is 1.1 in REF6 (15), is 3.4~4.4 kg in REF4 (2) REF5 (5) and REF 3(10), and increases sharply to 23.4 kg in REF7 (0.25) (which is more than 20 times of that in REF6 (15)). This agrees with the results for instantaneous sediment transport rate shown in Fig. 6(b), and shows that the duration of conditioning flow can influence the sediment transport at the beginning of the subsequent flood, with a longer conditioning phase leading to less sediment transport. When the duration of conditioning flow is over 2 hours, the subsequent sediment transport rate becomes rather insensitive to further increase of conditioning duration, indicating that the reorganization of the river bed under conditioning flow is mostly finished within 2 hours. The effects of stress history on subsequent sediment transport can hardly be observed during Step 2 of the hydrograph (Fig. 7(c)). Sediment output in REF7 (0.25) reduces significantly to similar magnitude of other experiments, because most of the loose bed material in REF7 (0.25) has been moved by the end of Step 1. More specifically, the volumes of sediment output in this step range~~s~~ between 3.1 kg and 8.6 kg, with the largest output occurring in REF5 (5) and the minimum output occurring in REF3 (10). We further calculate the correlation coefficient between sediment output and conditioning duration and obtain a value of r = -0.61, indicating that a longer conditioning duration can no longer lead to a larger sediment output in this step. In Step 3 of the hydrograph (Fig. 7(d)), sediment output in REF7 (0.25) and REF4 (2) is larger than in other 3 experiments which have longer conditioning phases. But in this step the sediment output in REF7 (0.25) is no more than three times that of the sediment output in REF3 (10), which has the minimum sediment output. ~~this~~ This difference of sediment output among experiments is not as significant as in Step 1. In the last step of the hydrograph, with the flow discharge and sediment supply approaching their peaks, the difference in sediment output among the five experiments again becomes small, with the values ranging between 72.1 kg in REF4 (2) and 119.6 kg in REF6 (15). This demonstrates~~present similar sediment outputs, demonstrating~~ that little influence of stress history remains in this step.

Figure 8 shows the temporal variation of the grain size distribution of the bedload. Here D_{l10}, D_{l50}, and D_{l90} denote grain sizes such that 10%, 50%, and 90% are finer in the bedload, respectively. Accuracy of the measurements is estimated by comparing the light table data with the trap data. Results show that for our experiments, the light table method has good accuracy in terms of the median size of bedload (D_{l50}), with an overestimation by 3% on average (111 samples and a standard deviation of 40.1%). Measurements of D_{l10} and D_{l90} show less accuracy, with an underestimation by 20% on average (111 samples and a standard deviation of 39.0%) for D_{l10} and an overestimation by 30% on average (111 samples and a standard deviation of 26.5%) for D_{l90}. Details concerning~~of~~ this uncertainty analysis are presented in the Supporting Information.

The value of D_{l10} shows a decreasing trend during the conditioning phase (Fig. 8 (a)), with a value of more than 2 mm at the beginning to about 0.6 mm after 15 hours, in spite of the large fluctuations before 8 hours. The decrease of D_{l10} reflects an increase in the fraction of the finest sediment in bedload. In the first two steps of the hydrograph (Fig. 8(b)), the value of D_{l10} is relatively stable for experiments with long conditioning phases (i.e., REF6 (15) and REF3 (10)), but shows a decreasing trend along with fluctuations for experiments with short conditioning phases (i.e., REF7 (0.25), REF4 (2), and REF5 (5)). The last two steps of the hydrograph see an evident increase in the value of D_{l10} compared with the first two steps, due to the increase of flow discharge and sediment supply (Fig. 8(b)). We note that such an increase in the D_{l10} is larger than the standard deviation of measurements, as shown above.

Figures 8(c) and 8(d) show the temporal variation of D_{l50}. Compared with that of D_{l10}, the temporal variation of D_{l50} shows more significant fluctuations during the conditioning phase (especially after t = 10 hour), as well as at the beginning of the hydrograph~~,~~. This can be shown by the coefficient of variation (cv) of the grain size. For the conditioning phase (after t = 10 hour), the cv of D_{l10} show an average value of 0.05 whereas the cv of D_{l50} show an average value of 1.44. For Step 1 of the hydrograph phase, the cv of D_{l10} show an average value of 0.35 whereas the cv of D_{l50} show an average value of 0.66. For Step 2 of the hydrograph phase, the cv of D_{l10} show an average value of 0.12 whereas the cv of D_{l50} show an average value of 0.54. ~~and a decreasing or increasing trend for grain size in the conditioning/hydrograph phase is not as evident.~~ As for the temporal variation of D_{l90} (in Figs. 8(e) and 8(f)), the fluctuations are still significant, with the average cv being 0.61, 0.34, 0.27 for the conditioning phase (after t = 10 hour), Step 1 of hydrograph phase, and Step 2 of hydrograph phase, respectively. ~~and~~ Besides, there is~~ almost~~ no significant increase of decrease of ~~trend for~~ D_{l90} ~~either increasing or decreasing grain size~~ during the experiment. This indicates that the transport of the coarsest sediment is not sensitive to the variation of our experimental conditions. The more significant fluctuations in D_{l50} and D_{l90} might be attributed to the fact that during relatively low flow coarse sediment is more likely to be near the threshold of motion and move intermittently, e.g. as individual grains~~in pulses~~, as opposed to the more continuous movement for fine sediment. These fluctuations gradually diminish with the increase of flow and sediment supply, as the static armor on bed surface transits to mobile armor and the movement of coarse grains become more continuous.

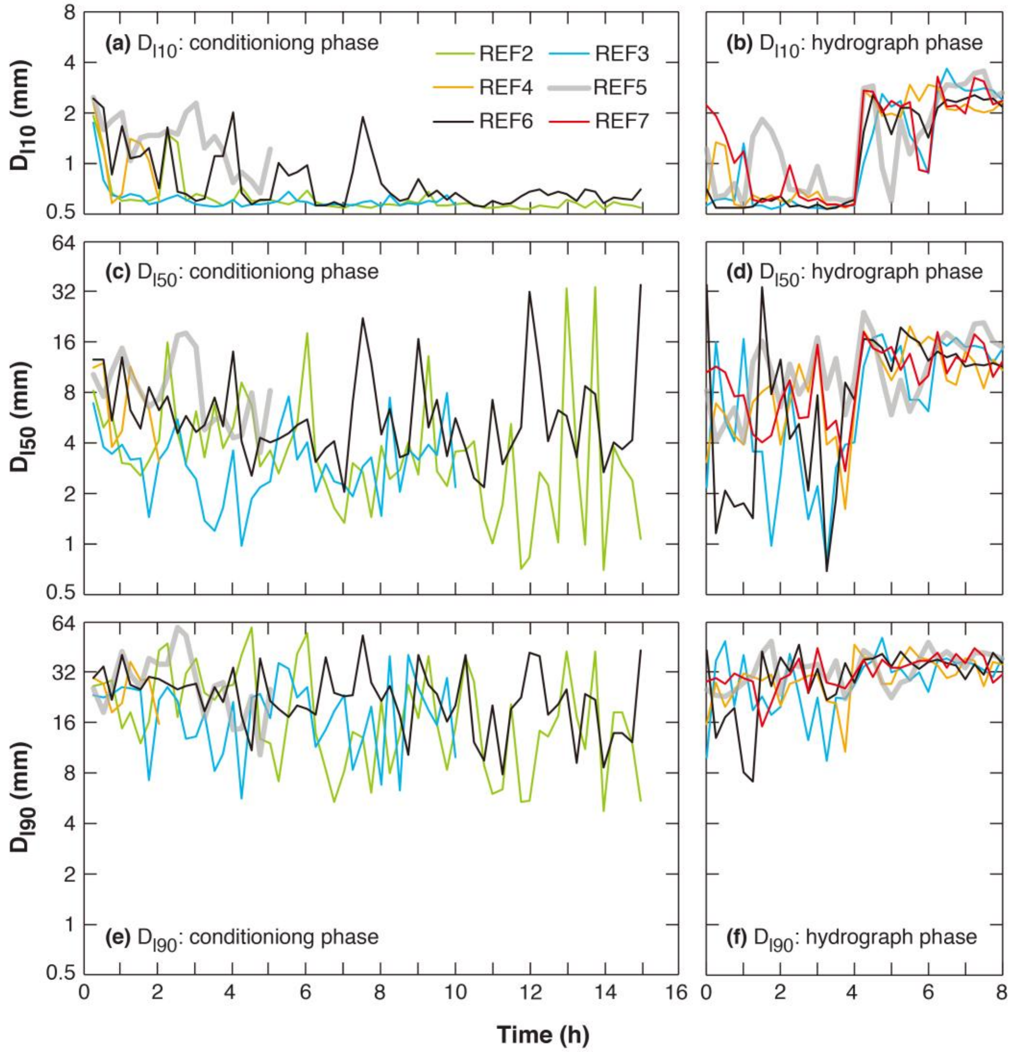


Figure 8. Temporal adjustments of characteristic grain sizes of bedload. (a) D_{l10} during conditioning phase; (b) D_{l10} during hydrograph phase; (c) D_{l50} during conditioning phase; (d) D_{l50} during hydrograph phase; (e) D_{l90} during conditioning phase; (f) D_{l90} during hydrograph phase.

With the fractional sediment transport rate measured by the light table, we also analyze the sediment mobility of each size range during the experiment. Results show that sediment transport rate is characterized by equal mobility (i.e., the GSD of sediment load matches the GSD of sediment on bed surface) at the beginning of the conditioning phase, but moves to partial/selective mobility after a relatively long conditioning phase as well as during the first two steps of the hydrograph. However, with the increase of flow discharge and sediment supply, the sediment transport regime gradually returns to equal mobility during the last two steps of the hydrograph. Details of the analysis are presented in the Supporting Information.

4 Discussion

4.1 Threshold of sediment motion in experiments

The threshold of sediment motion is a key parameter for the prediction of bedload transport. Previous studies on the stress history effect often start with a conditioning flow that is below the threshold of motion, and then gradually increase the flow discharge, so that the threshold of motion can be directly estimated in the experiment (e.g., Monteith and Pender, 2005; Masteller and Finnegan, 2017; Ockelford et al., 2019; etc.). Because our experiments implement a conditioning flow which can mobilize sediment (sediment transport at the beginning of the conditioning phase is especially large), the threshold of motion cannot be observed directly in the experiment. Here we follow the method applied in Hassan et al. (2020), and estimate the threshold of sediment motion ~~by adopting~~with the Wong and Parker (2006) sediment transport relation, which is a revision of the Meyer-Peter and Müller (1948) relation.

We use the Wong and Parker (2006) relation, which maintains the exponent 1.5, of Meyer-Peter and Muller (1948):

$$q_s^* = 3.97\left(\tau_{s50}^* - \tau_c^*\right)^{1.5} \tag{1}$$

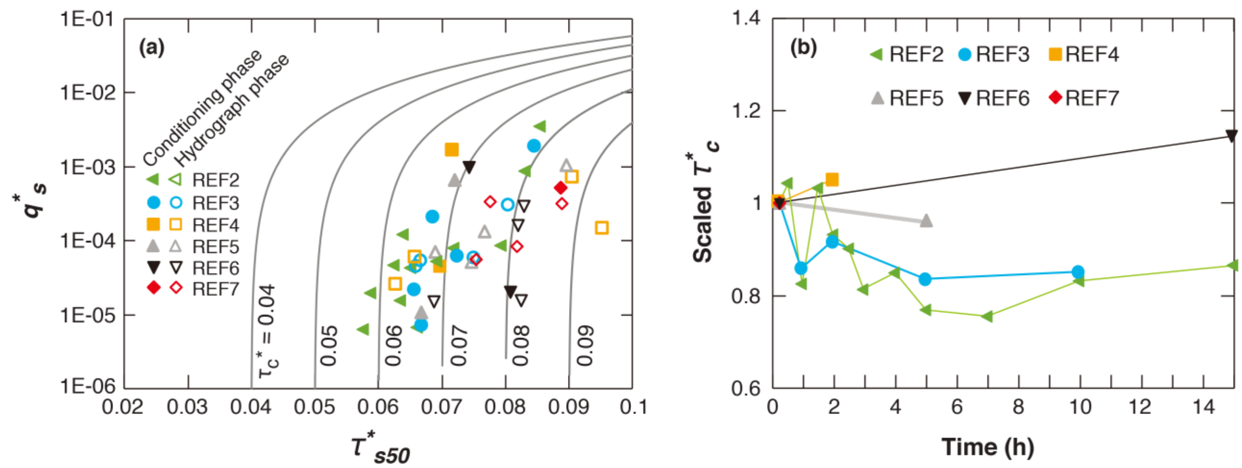
$$q_s^* = \frac{q_s}{\sqrt{RgD_{s50}}D_{s50}} \tag{2}$$

$$\tau_{s50}^* = \frac{\tau_b}{\rho gRD_{s50}} \tag{3}$$

$$\tau_b = \rho ghS_w \tag{4}$$

where q_s^* is the dimensionless bedload transport rate (Einstein number) defined by Eq. (2), τ_{s50}^* is the Shields number for surface median grain size D_{s50} defined by Eq. (3), τ_b is the flow shear stress calculated using the depth-slope product (Eq. (4)), τ_c^* is the critical Shields number for the threshold of sediment motion, q_s is the volumetric sediment transport rate per unit width; h is water depth, S_w is water surface slope, R = 1.65 is the submerged specific gravity of sediment, g = 9.81 m/s^2 is the gravitational acceleration and ρ = 1000 kg/m^3 is the water density. Wong and Parker (2006) proposed a value of 0.0495 for τ_c^* in Eq. (1). Here we obtain q_s^* and τ_{s50}^* from the measured data of the experiments, and back calculate the value of τ_c^* using Eq. (1). It is worth mentioning that in Hassan et al. (2020) three different methods, including the method as described above, are applied to estimate the threshold of sediment motion. Estimation with the three different methods shows very similar temporal trend and variability.

Figure 9(a) shows the values of q_s^* vs. τ_{s50}^* for each experiment, along with the Wong and Parker (2006) type relation (Eq. (1)) with various values for τ_c^* (from 0.04 to 0.09). It can be seen from the figure that the measured sediment transport rate is relatively low, with most points below the dimensionless value of 0.001. This indicates that the Shields number in our experiment is slightly larger than the critical Shields number, a state that is typical for gravel-bed rivers (Parker, 1978). The four points with dimensionless transport rate above 0.001 are all at the beginning of the conditioning flow (t = 15 minutes). The values of q_s^* basically show an increasing trend with the increase of τ_{s50}^*, with the correlation coefficient between τ_{s50}^* and log(q_s^*) (in consistentce with the semi-log scale of Figure 9(a)) being 0.58. Besides,but with the values of critical Shields number τ_c^* shown in Figure 9(a) covers a rather wide range (from less than 0.06 to larger than 0.09).



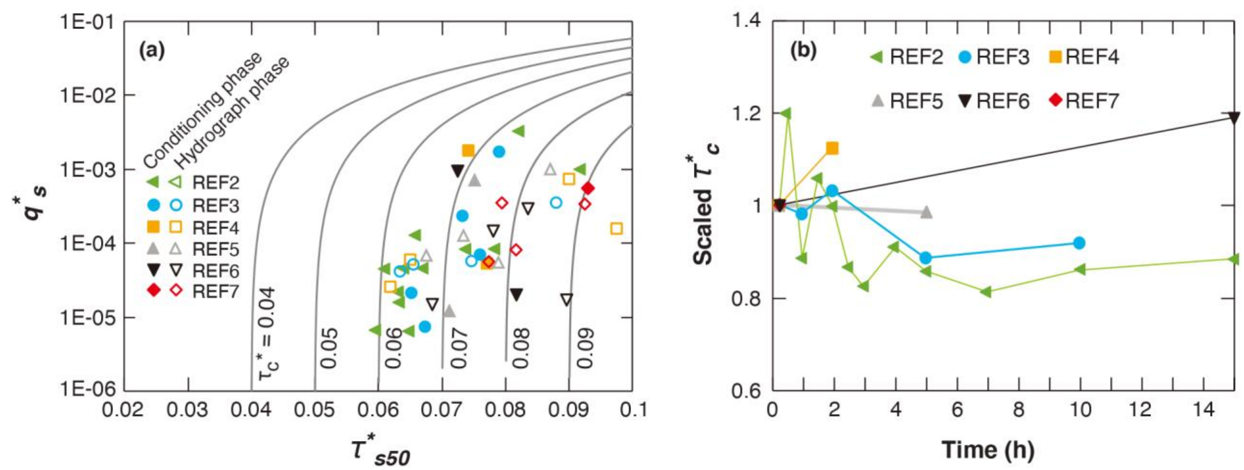


Figure 9. (a) Dimensionless sediment transport rate q_s^* vs. Shields number τ_{s50}^* using surface median grain size for measured transport rates (points). Also shown are lines for the Wong and Parker (2006) type equation (Eq. 1) using different values for τ_c^*. (b) Temporal adjustment of scaled τ_c^* (τ_c^* over τ_c^* at 15 minutes) during the conditioning phase. Here τ_c^* is back-calculated using Eq. (1) (Wong and Parker (2006) type relation).

Table 2 shows the values of τ_c^* back-calculated at the beginning (t = 15 minutes) and the end of the conditioning phase in each experiment. The back-calculated values of τ_c^* vary in the range 0.06665~0.086 090 for the conditioning phase, which is well above the value of 0.0495 as recommended by Wong and Parker (2006). Lamb et al. (2008) demonstrated that critical shear stress can become larger for large bed slope, and they proposed a relation which considers the effect of bed slope,

$$\tau_c^* = 0.15 S_b^{0.25} \tag{5}$$

where S_b is bed slope. For comparison, Table 2 also shows the values of τ_c^* calculated by Eq. (5). Results shows that for the conditioning phase of our experiments, τ_c^* calculated by Eq. (5) is above 0.06, which is much higher than the recommended value of Wong and Parker (2006)~~ and is closer to the values back-calculated by Eq. (1)~~. Besides, the τ_c^* values predicted by the Lamb et al. (2008) relation show little variability among different experiments, compared with the values back~~-~~-calculated with equation (1) based on experimental data. More specifically, the cv values are 0.032 at t = 15 minutes and 0.031 at the end of the conditioning phase for τ_c^* predicted by Lamb et al. (2008) relation, but become 0.10 at t = 15 minutes and 0.12 at the end of the conditioning phase for τ_c^* back~~-~~-calculated with equation (1) using measured data. Such discrepancies could be ascribed to the fact the relation of Lamb et al. (2008) considers only the influence of bed slope, ~~but~~ without considering the effects of other mechanisms like organization of surface texture, infiltration of fine particles, etc. These potential effects are discussed in more detail in Section 4.2.~~ indicating that only the slope effect cannot explain the observed range of τ_c^*.~~

Here we also estimate the uncertainties associated with the calculation of τ_c^*. For ~~the~~ τ_c^* back~~-~~-calculated with equation (1), the~~its~~ global uncertainty is estimated by combining the uncertainties of each parameter ~~as~~ involved in the calculation, i.e. water depth h, water surface slope S_w, sediment transport rate q_s, and surface median grain size D_{s50}. The applied ranges of h and S_w are the measured values plus/minus the errors associated with the gauge point. The applied ranges of q_s and D_{s50} are the measured values plus/minus the standard deviations as reported in Section 3. Results of the uncertainties are presented in the brackets in Table 2. For the τ_c^* values calculated with ~~the~~ Equation (5), the uncertainties are only from the bed slope S_w (which is related with the resolution of point gauge), and is less than ±1% according to our ~~as we~~ estimates~~d~~. Therefore, the uncertaint~~y~~ies of τ_c^* calculated with the Equation (5) is not presented in the table. It can be seen from Table 2 that the values of τ_c^* calculated with the Equation (5) are mostly within the uncertainty range of τ_c^* back~~-~~-calculated with Eq. (1), with the values closer to the lower bound of the uncertainty range.

Table 2. Values of τ_c^* at the beginning (t = 15 minutes) and the end of conditioning phase in each experiment. Here τ_c^* is back-calculated with Eq. (1). Also shown here are values of τ_c^* estimated with the equation of Lamb et al. (2008) for comparison. Values in the brackets denote the range of uncertainty associated with the τ_c^* values back~~-~~-calculated with Eq. (1).

		~~REF2 (15)~~	~~REF6 (15)~~	~~REF3 (10)~~	~~REF5 (5)~~	~~REF4 (2)~~	~~REF7 (0.25)~~
~~t = 15 minutes~~	~~Back calculated by Eq. (1)~~	~~0.076~~	~~0.070~~	~~0.078~~	~~0.069~~	~~0.066~~	~~0.086~~
	~~Lamb et al. (2008)~~	~~0.063~~	~~0.066~~	~~0.061~~	~~0.065~~	~~0.061~~	~~0.066~~
~~End of conditioning~~	~~Back calculated by Eq. (1)~~	~~0.066~~	~~0.081~~	~~0.067~~	~~0.066~~	~~0.069~~	~~0.086~~
	~~Lamb et al. (2008)~~	~~0.061~~	~~0.063~~	~~0.060~~	~~0.063~~	~~0.062~~	~~0.066~~

	t = 15 minutes		End of conditioning	
	Back-calculated by Eq. (1)	Lamb et al. (2008)	Back-calculated by Eq. (1)	Lamb et al. (2008)
REF2 (15)	0.073 (0.064, 0.083)	0.063	0.065 (0.057, 0.074)	0.061
REF6 (15)	0.068 (0.053, 0.089)	0.066	0.081 (0.072, 0.093)	0.063
REF3 (10)	0.073 (0.061, 0.088)	0.061	0.067 (0.058, 0.079)	0.060
REF5 (5)	0.072 (0.061, 0.085)	0.065	0.071 (0.062, 0.081)	0.063
REF4 (2)	0.068 (0.059, 0.079)	0.061	0.077 (0.066, 0.090)	0.062
REF7 (0.25)	0.090 (0.075, 0.109)	0.066	0.090 (0.075, 0.109)	0.066

In Fig. 9(b), we plot the scaled τ_c^* during the conditioning phase of our experiments. For each experiment, the scaled τ_c^* is calculated as the ratio between τ_c^* and the corresponding τ_c^* at t = 15 minutes. τ_c^* implemented here is back-calculated with Eq. (1). The scaled τ_c^* collapses on a value of unity at t = 15 minutes (i.e., the first point of each experiment). It can be seen from the figure that different trends are exhibited for the adjustment of τ_c^* from t = 15 minutes to the end of conditioning phase, with REF2 (15) and REF3 (10) exhibiting a decreasing trend, ~~REF4 (2) and~~ REF5 (5) exhibiting very slight changes, and REF4 (2) and REF6 (15) exhibiting an increasing trend. The decrease of τ_c^* in REF2 (15) an REF3 (10) is accompanied by a reduction of Shields number τ_{s50}^*, mainly due to the increase of surface median grain size D_{s50}. Moreover, the variation of back-calculated τ_c^* is mostly within a range of ±20%, in agreement with our observation that variation of bed topography and bed surface texture become insignificant after 15 minutes. It should be noted that τ_c^* cannot be back-calculated using Eq. (1) within the first 15 minutes of the conditioning phase, since the information for flow depth, water surface slope and bed surface GSD is not available. Nevertheless, we expect the adjustment of τ_c^* could be evident within the first 15 minutes, since the adjustments of both bed topography and bed surface are significant during this period (as shown in Sect. 3.1).

4.2 Implications and limitations

Previous research has shown that antecedent conditioning flow can lead to an increased critical shear stress and reduced sediment transport rate during subsequent flood event (Hassan and Church, 2000; Haynes and Pender, 2007; Ockelford and Haynes, 2013; Masteller and Finnegan, 2017~~; etc.~~). Our flume experiments also show a ~~reduced~~reduction in sediment transport rate, especially at the beginning of the hydrograph, in response to the implementation of antecedent conditioning flow (as shown in Fig. 6(b) and Fig. 7). However, our results are different from previous research in that the influence of antecedent conditioning flow is found to last for a relatively short time at the beginning of the following hydrograph, and then gradually diminish with the increase of flow intensity as well as sediment supply (Figs. 6 and 7). Such results indicate that increasing flow intensity and sediment supply during a flood event can lead to the loss of memory of stress history. A similar phenomenon was observed by Mao (2018) in his experiment, where sediment transport during a high-magnitude flood event was not much affected by the occurrence of lower-magnitude flood event before. Besides, the subsequent hydrograph leads to evident bed degradation (Fig. 3) and increase of sediment transport rate (Figs. 6 and 7), but does not lead to evident change of surface texture or break of the armor layer (Fig. 5). This is in agreement with the observation of Ferrer-Boix and Hassan (2015) during experiments of successive water pulses.

Our results have practical implications for mountain gravel bed rivers. The importance of conditioning flow has long been discussed in the literature, and researchers have suggested that the stress history effect be considered in the modeling and analysis of gravel bed rivers. For example, previous research states that existing sediment transport theory for gravel bed rivers (e.g., Meyer-Peter and Müller, 1948; Wilcock and Crowe, 2003; Wong and Parker, 2006; etc.) might lead to unrealistic predictions if the stress history effect is not taken into account (Masteller and Finnegan, 2017; Mao, 2018; Ockelford et al., 2019). Our results indicate that the stress history effect is important and needs to be considered for low flow as well as the beginning of the flood event, but becomes insignificant as the flow gradually approaches high flow discharge. ~~This could have implications in river engineering such as water and sediment regulation schemes for mountain gravel-bed rivers.~~

To explain the effect of stress history, Ockelford and Haynes (2013) has summarized the following possible mechanisms. (1) Vertical settling during the conditioning flow consolidates the bed into a tighter packing arrangement which is more resistant to entrainment. (2) Local reorientation and rearrangement of surface particles provide a greater degree of imbrication, less resistance to fluid flow, as well as direct sheltering on the bed surface. (3) The infiltration of fines into low-relief pore spaces can further increase the bed compaction. In the experiment of Masteller and Finnegan (2017), it was found that the most drastic changes during conditioning flow are manifest in the extreme tail of the elevation distribution (i.e., the reorientation of the highest protruding grains into nearby available pockets) and go therefore undetected in most bulk measurements (e.g. the mean bed elevation ~~or~~, standard deviation of bed topography, or the bed surface GSD). They demonstrated that such reorganization of the highest protruding grains can indeed lead to noticeable differences in the threshold of sediment transport (Masteller and Finnegan, 2017). This might explain the observation in our experiment that after the first 15 minutes of the conditioning phase, adjustments of the bed topography and the bed surface GSD become insignificant, but the sediment transport rate as well as its GSD keeps adjusting consistently.

In our experiments as well as previous experiments that study the effect conditioning flow (e.g., Monteith and Pender, 2005; Masteller and Finnegan, 2017; Ockelford et al., 2019~~; etc.~~), no sediment supply is implemented during the conditioning flow, and the flow can reorganize the bed surface to a state that is more resistant to sediment entrainment. Therefore, it is straightforward to expect that the conclusions based on our flume experiments to apply for natural rivers where sediment supply is relatively low during low flow conditions. However, some gravel-bed rivers have quite active hillslopes, and sediment input from hillslopes to river channel can occur regularly (Turowski et al., 2011; Reid et al., 2019). Since the sediment material from hillslopes is typically loose and easy to transport, under such circumstances a long inter-event duration (i.e., low-flow duration) might lead to an enhanced sediment transport rate in the subsequent flood (Turowski et al., 2011).

It should also be noted that in previous experiment on the stress history effect, conditioning flow is often set below the threshold of sediment motion. One exception is the experiment of Haynes and Pender (2007) in which the conditioning flow ~~is~~ was above the threshold of motion for D_{50}. By implementing conditioning flow with various durations and magnitudes, they demonstrated that a longer duration of conditioning flow will increase the bed stability whereas a higher magnitude of conditioning flow will reduce the bed stability. However, since the subsequent flow they implement to test the bed stability was constant through time, their results did not show how a subsequent flow event with increasing intensity would affect the stress history. ~~In this the present paper~~ Here we ~~also~~ implement a conditioning flow which can mobilize sediment, especially at the beginning of the conditioning phase during which evident sediment transport occurs. Moreover, by implementing a subsequent (rising limb of) ~~the~~ hydrograph, we find that the stress history can persist during the beginning of the hydrograph but is eventually erased out as the flow intensity increases~~goes large~~. In our experiments, we varied the duration of conditioning flow by fixing the conditioning flow magnitude. In this sense, how the stress history formed under various magnitudes of conditioning flow (both above-and below-threshold) would be affected by a subsequent hydrograph still merits future research.~~Compared with the below-threshold conditioning flow, we consider that the above-threshold conditioning flow can induce more evident reorganization of bed surface, which might be more lasting during subsequent flood. That said, we expect the conclusion of this study can still hold if below-threshold conditioning flow is implemented. Nevertheless, flume experiments with various magnitudes of conditioning flow (both above- and below-threshold of motion) merit future study.~~

Recently, Church et al. (2020) drew attention to the reproducibility of results in geomorphology. They distinguished three levels of "reproducibility", including "repetition", "replication", and "reproduction". In this paper, the repetition of ~~the~~ experimental results is tested by repeating the conditioning phase with the longest duration (REF6 (15) and REF2 (15)). The two experiments show similar results during the conditioning phase in terms of standard deviation of bed elevation, GSD of bed surface, sediment transport rate, and GSD of sediment load. However, the reproduction of the experimental results, which requires independent tests undertaken using different materials and/or different conditions of measurement, and which is more significant, according to Church et al. (2020), for advancing of the science ~~according to Church et al. (2020)~~, has not been tested in this paper. In this regard, more efforts are needed in future study to test the reproducibility of the conclusions given in this paper.~~Besides, considering that the conditions of existing experiments on stress history effect are limited, implementation of numerical simulations under a wider range of conditions also merits future study.~~

5 Conclusions

In this paper, the effect of antecedent conditioning flow (i.e., the effect of stress history) on the morphodynamics of gravel-bed rivers during subsequent floods is studied via flume experimentation. The experiment described here is designed based on the conditions of East Creek, Canada. The experiment consisted~~s~~ of two phases: a conditioning phase with constant water discharge and no sediment supply, followed by a hydrograph phase with hydrograph and sedimentograph. Five runs (REF 3~7) were conducted with identical experimental conditions except different durations of conditioning phase. Another run (REF 2), which consisted~~s~~ of only the conditioning phase, is conducted in order to test the reproducibility of experimental results during the conditioning flow. Experimental results show the following.

- Adjustments of channel morphology (including channel bed longitudinal profile, standard deviation of bed elevation, characteristic grain sizes of bed surface material) are evident during the first 15 minutes of the conditioning phase, but become insignificant during the remainder of the conditioning phase.
- The implementation of conditioning flow can indeed lead to a reduction in sediment transport during the subsequent hydrograph, which agrees with previous research.
- However, the effect of stress history on sediment transport rate is limited to a relatively short time at the beginning of the hydrograph, and gradually diminishes with the increase of flow discharge and sediment supply, indicating a loss of memory of stress history under high flow discharge. Also, the effect of stress history on the GSD of both bed surface and bedload is not evident.
- The threshold of sediment motion is estimated with the form of the Wong and Parker (2006) relation. The estimated critical Shields number varies in the range 0.066~0.086 during the conditioning phase (excluding the first 15 minutes), and is higher than the value recommended by Wong and Parker (2006).

Our study has implications in regard to a wide range of issues for mountain gravel-bed rivers, including sediment budget analysis, river morphodynamic modeling, water and sediment regulation, flood management, and ecological restoration schemes.

Notation

D_{l50}: grain size such that 50 percent in sediment load is finer (similarly D_{l10} is such that 10 percent in sediment load is finer and D_{l90} is such that 90 percent in sediment load is finer).

D_{s50}: grain size such that 50 percent on bed surface is finer (similarly D_{s10} is such that 10 percent on bed surface is finer and D_{s90} is such that 90 percent on bed surface is finer).

F_r: Froude number.

g: gravitational acceleration.

h: water depth.

Q_s: sediment transport rate.

q_s: volumetric sediment transport rate per unit width.

q_s^*: the dimensionless bedload transport rate (Einstein number).

R: submerged specific gravity of sediment.

S_b: bed slope.

S_w: water surface slope.

ρ: water density.

Δz_b: mean difference of bed elevation;

τ_b: bed shear stress.

τ_c^*: critical Shields number for the threshold of sediment motion.

τ_{s50}^*: dimensionless shear stress (Shields number) of the D_{s50}.

Data availability

Data used for the analysis can be found at doi: 10.6084/m9.figshare.12758414 (An, 2020).

Author contribution

Marwan A. Hassan and Xudong Fu designed the research. Carles Ferrer-Boix performed the experiments. Chenge An processed and analyzed the experimental data. Chenge An prepared the manuscript with contributions from all coauthors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

Gary Parker provided constructive comments and helped edit ~~of~~ this paper. Maria A. Elgueta-Astaburuaga helped conduct the experiments. Rick Ketler provided support in equipment and data collections. Eric Leinberger provided support in designing the figures. We thank Jens Turowski and another anonymous reviewer for their constructive comments, which helped us greatly improve the paper. The participation of Chenge An was supported by grant from China Postdoctoral Science Foundation (grant 2018M641368). The participation of Xudong Fu was supported by grants from the National Natural Science Foundation of China (grants 51525901 and 91747207).

References

An, C.: Experimental data on sediment transport and channel adjustment in a gravel-bed river: stress history effect, doi: 10.6084/m9.figshare.12758414, 2020.

Carling, P. A., Kelsey, A., and Glaister, M. S.: Effect of bed roughness, particle shape and orientation on initial motion criteria, in: Dynamics of Gravel-bed Rivers, edited by: Billi, P., Hey, R. D., Thorne, C. R., and Tacconi, P., John Wiley & Sons, Chichester, UK, 23-39, 1992.

Chartrand, S. M., Jellinek, A. M., Hassan, M. A., and Ferrer-Boix, C.: Morphodynamics of a width-variable gravel bed stream: New insights on pool-riffle formation from physical experiments, Journal of Geophysical Research-Earth Surface, 123(11), 2735-2766, https://doi.org/10.1029/2017JF004533, 2018.

Chen, X., Hassan, M. A., An, C., and Fu, X.: Rough correlations: Meta-analysis of roughness measures in gravel bed rivers, Water Resources Research, 56, e2020WR027079, https://doi.org/10.1029/2020WR027079, 2020.

Chin, A., Anderson, S., Collison, A., Ellis-Sugai, B. J., Haltiner, J. P., Hogervorst. J. B., Kondolf, G. M., O'Hirok, L. S., Purcell, A. H., Riley, A. L., and Wohl E.: Linking theory and practice for restoration of steppool streams, Environmental Management, 43, 645–661, doi:10.1007/s00267-008-9171-x, 2009.

Church, M., Dudill, A., Venditti, J. G., and Frey, P.: Are Results in Geomorphology Reproducible? Journal of Geophysical Research-Earth Surface, 125(8), e2020JF005553, https://doi.org/10.1029/2020JF005553, 2020.

Curran, J. C., and Wilcock, P. R.: Effect of sand supply on transport rates in a gravel-bed channel, Journal of Hydraulic Engineering, 131(11), 961–967, https://doi.org/10.1061/(ASCE)0733- 9429(2005)131:11(961), 2005.

Elgueta-Astaburuaga, M. A., and Hassan, M. A.: Experiment on temporal variation of bed load transport in response to changes in sediment supply in streams, Water Resources Research, 53, 763–778, doi:10.1002/2016WR019460, 2017.

Ferrer-Boix, C. and Hassan, M. A.: Influence of the sediment supply texture on morphological adjustments in gravel-bed rivers, Water Resources Research, 50(11), 8868-8890, doi:10.1002/2013WR015117, 2014.

Ferrer-Boix, C., and Hassan, M. A.: Channel adjustments to a succession of water pulses in gravel bed rivers, Water Resources Research, 51, 8773–8790, doi:10.1002/2015WR017664, 2015.

Gomez, B. and Church, M.: An assessment of bed load sediment transport formulae for gravel rivers, Water Resources Research, 25, 1161–1186, doi:10.1029/WR025i006p01161, 1989.

Hassan, M. A., and Church, M.: Experiments on surface structure and partial sediment transport on a gravel bed, Water Resources Research, 36, 1885-1895, 2000.

Haynes, H., and Pender, G.: Stress history effects on graded bed stability, Journal of Hydraulic Engineering, 33, 343–349, 2007.

Howard, A.: A detachment-limited model of drainage basin evolution, Water Resources Research, 30(7), 2261-2285, 1994.

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

Klingeman, P. C., and Emmett, W. W.: Gravel bedload transport processes, in: Gravel-Bed Rivers, edited by: Hey, R. D., Bathurst, J. C., and Thorne, C., John Wiley & Sons, Chichester, UK, 141-180, 1982.

Lamb, M. P., Dietrich, W. E., and Venditti, J. G.: Is the critical Shields stress for incipient sediment motion dependent on channel-bed slope? Journal of Geophysical Research-Earth Surface, 113, F02008, doi:10.1029/2007JF000831, 2008.

Lenzi, M. A.: Step–pool evolution in the Rio Cordon, northeastern Italy, Earth Surface Processes and Landforms, 26, 991-1008, https://doi.org/10.1002/esp.239, 2001.

Mao, L.: The effects of flood history on sediment transport in gravel bed rivers, Geomorphology, 322, 192–205, https://doi.org/10.1016/j.geomorph.2018.08.046, 2018.

Marston, R. A.: Land, life, and environmental change in mountains, Annals of the Association of American Geographers, 98, 507–520, https://doi.org/10.1080/00045600802118491, 2008.

Masteller, C. C., and Finnegan, N. J.: Interplay between grain protrusion and sediment entrainment in an experimental flume, Journal of Geophysical Research-Earth Surface, 122, 274–289, https://doi.org/10.1002/2017GL076747, 2017.

Masteller, C. C., Finnegan, N. J., Turowski, J. M., Yager, E. M., and Rickermann, D.: History dependent threshold for motion revealed by continuous bedload transport measurements in a steep mountain stream, Geophysical Research Letters, 46, 2583–2591, 2019.

Monteith, H., and Pender, G.: Flume investigation into the influence of shear stress history, Water Resources Research, 41, W12401, https://doi.org/10.1029/2005WR004297, 2005.

Montgomery, D. R., and Buffington, J. M.: Channel‐reach morphology in mountain drainage basins, Geological Society of America Bulletin, 109(5), 596–611, https://doi.org/10.1130/0016‐7606(1997)109<0596:CRMIMD>2.3.CO;2, 1997.

Montgomery, D. R., Buffington, J. M., Peterson, N. P., Schuett-Hames, D., and Quinn, T. P.: Stream-bed scour, egg burial depths, and the influence of salmonid spawning on bed surface mobility and embryo survival, Canadian Journal of Fisheries and Aquatic Sciences, 53, 1061–1070, 1996.

Meyer-Peter, E., and Müller, R.: Formulas for bed-load transport, in: Proceedings of the 2nd Congress of International Association for Hydraulic Structures Research, Stockholm, Sweden, 7-9 June 1948, 39-64, 1948.

Ockelford, A., and Haynes, H.: The impact of stress history on bed structure, Earth Surface Processes and Landforms, 38, 717–727, https://doi.org/10.1002/esp.3348, 2013.

Ockelford, A., Woodcock, S., and Haynes, H.: The impart of inter-flood duration on non-cohesive sediment bed stability, Earth Surface Processes and Landforms, 44, 2861-2871, doi:10.1002/esp.4713, 2019.

Paola, C., Heller, P. L., and Angevine, C. L.: The large-scale dynamics of grain-size variation in alluvial basins, I: Theory, Basin Research, 4, 73–90, 1992.

Papangelakis, E., and Hassan, M. A.: The role of channel morphology on the mobility and dispersion of bed sediment in a small gravel-bed stream, Earth Surface Processes and Landforms, 41, 2191–2206, 2016.

Paphitis, D., and Collins, M. B.: Sand grain threshold, in relation to bed stress history: an experimental study, Sedimentology, 52, 827–838, 2005.

Parker, G.: Self-formed straight rivers with equilibrium banks and mobile bed. Part 2. The gravel river, Journal of Fluid Mechanics, 89, 127-146, 1978.

Parker, G.: 1D sediment transport morphodynamics with applications to rivers and turbidity currents, available at: http://hydrolab.illinois.edu/people/parkerg//morphodynamics_e-book.htm, 2004.

Powell, D. M., Reid, I., and Laronne, J. B.: Hydraulic interpretation of crossstream variations in bed-load transport, Journal of Hydraulic Engineering, 125, 1243–1252, 1999.

Reid, D. A., Hassan, M. A., Bird, S., and Hogan D.: Spatial and temporal patterns of sediment storage over 45 years in Carnation Creek, BC, a previously glaciated mountain catchment, Earth Surface Processes and Landforms, 44, 1584-1601, doi: 10.1002/esp.4595, 2019.

Reid, I., and Frostick, L. E.: Particle interaction and its effects on the thresholds of initial and final bedload motion in coarse alluvial channels, in: Sedimentology of Gravels and Conglomerates-Memoir 10, edited by: Koster, E. H., and Steel, R. J., Canadian Society of Petroleum Geologists, Calgary, Canada, 61–68, 1984.

Reid, I., Frostick, L. E., and Layman, J. T.: The incidence and nature of bedload transport during flood flows in coarse-grained alluvial channels, Earth Surface Processes and Landforms, 10, 33–44, 1985.

Rickenmann, D.: Comparison of bed load transport in torrents and gravel bed streams, Water Resources Research, 37, 3295–3305, doi:10.1029/2001WR000319, 2001.

Schneider, J. M., Rickenmann, D., Turowski, J. M., Bunte, K., and Kirchner, J. W.: Applicability of bed load transport models for mixed-size sediments in steep streams considering macro-roughness, Water Resources Research, 51, 5260–5283, doi:10.1002/2014wr016417, 2015.

Sklar, L., and Dietrich W. E.: A mechanistic model for river incision into bedrock by saltating bed load, Water Resources Research, 40, W06301, doi:10.1029/2003WR002496, 2004.

Turowski, J. M., Yager, E. M., Badoux, A., Rickenmann, D., and Molnar, P.: The impact of exceptional events on erosion, bedload transport and channel stability in a step-pool channel, Earth Surface Processes and Landforms, 34, 1661-1673, https://doi.org/10.1002/esp.1855, 2009.

Turowski, J. M., Badoux, A., and Rickenmann, D.: Start and end of bedload transport in gravel-bed streams, Geophysical Research Letters, 38, L04401, https://doi.org/10.1029/2010GL046558, 2011.

Venditti, J. G., Dietrich, W. E., Nelson, P. A., Wydzga, M. A., Fadde, J., and Sklar, L.: Mobilization of coarse surface layers in gravel-bedded rivers by finer gravel bed load, Water Resources Research, 46, W07506, https://doi.org/10.1029/2009WR008329, 2010.

Whiting, P. J., and Dietrich, W. E.: Boundary shear stress and roughness over mobile alluvial beds, Journal of Hydraulic Engineering, 116, 1495–1511, 1990.

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

Wilcock, P. R., Kenworthy, S. T., and Crowe, J. C.: Experimental study of the transport of mixed sand and gravel, Water Resources Research, 37(12), 3349–3358, 2001.

Wong, M., and Parker, G.: Reanalysis and correction of bed-load relation of Meyer-Peter and Müller using their own database, Journal of Hydraulic Engineering-ASCE, 132, 1159–1168, doi:10.1061/(ASCE)0733-9429(2006)132:11(1159), 2006.

Yager, E. M., Turowski, J. M., Rickenmann, D., and McArdell, B. W.: Sediment supply, grain protrusion, and bedload transport in mountain streams, Geophysical Research Letters, 39, L10402, https://doi.org/10.1029/2012GL051654, 2012.

Zimmermann, A., Church, M., and Hassan, M. A.: Video-based gravel transport measurements with a flume mounted light table, Earth Surface Processes and Landforms, 33(14), 2285-2296, https://doi.org/10.1002/esp.1675, 2008.