Response to Referee 2 Comment

Dynamics and Mechanics of Tracer Particles

by Colin B. Phillips, and Douglas J. Jerolmack

We thank the anonymous reviewer for a thorough and thoughtful review. This supplementary file documents in detail the response to Referee 2's comments for the discussion paper titled: Dynamics and Mechanics of Tracer Particles. Throughout the following text reviewer comments are in plain text, and our responses are in *italics* following a '>' symbol.

Start Referee 2's comments:

The paper deals with the mobility and displacement length of marked particles in gravel and boulder bed channels, which is a very relevant topic of high interest for geomorphologists, engineers and ecologists. The field data are hardly won, the analysis are well performed, and the paper is well written. I think that the paper will be of interest for the readers of ESURFD, and I suggest to accept it after minor revision. Specific comments are as follows:

- I have a certain concern on the use of data gathered from Bisley 3, on which only 50 RFID were installed. It is definitely useful to include these data on some of the analysis, but it seems to me that a full description of limitations and potential errors in managing such limited dataset should be given, in order to avoid some interpretations looking too speculative.
> *We have added additional text to the discussion section concerning the use of the Bisley 3 tracers. We remain surprised that what we perceive to be a limited number of tracers in the Bisley stream produced results consistent with the larger population of tracers within the main channel. We do not currently know what a representative sample size for sediment tracers is, and it is for this reason that we only include Bisley tracers along side the results of the Mameyes tracers.*

- I think that the introductory chapters (1 and 2) could be substantially shortened. It is indeed important to give credits to previous works and present the theoretical framework on which the data are later analysed. However, it seems to me that crucial information accounts for half of the text more or less. For instance, the introductory text on pg431 (rows 1 to 22) could be shortened to half, as the general description of the tracer techniques and advantages of using of RFID tracers. The text at page 433 seems all relevant to me, but the second half of page 434 could be shortened for example. Most of chapters 2.2 and 2.3 could be shortened as well.
> *We realize that the introductory sections 1 and 2 are on the longer side, however we have written a longer introduction section in order to reduce the need to introduce new information later in the text and to present the material to as broad of an audience as possible. We hoped that by providing a theory section those readers who are previously familiar with the various sub disciplines present will be able to skip section 2 entirely. However, it has been our experience that work with sediment tracers often lacks a strong connection to theory, and we attempt to fill that here for the interested reader. We also feel that each topic (tracer displacements at separate timescales, sorting, and hydrologic forcing) that we touch on has a rich history of literature that we would be remiss to exclude. However, if accepted for publication we will thoroughly edit a revised manuscript with the intention of removing unnecessary material.*

- Pg 433, row 28. As you've presented the formulas for calculating shear stress (pg 434, row 10), I would also write explicitly how the shear velocity was calculated. I would basically move here the

formula that is now at page 441, row 6.
> *Not all the references given on row 28 p 433 use the same methodology to calculate the shear velocity. We have added the relation between shear stress and shear velocity following equation 1.*

- Pg 434, rows 14-17. Why it's important for your field application that particle scale framework holds for laminar flow as well? Do you expect or did you observe laminar flow in your field site?
> *It is not important for our field setting that a framework holds in both laminar and turbulent flow as we did not observe laminar flow in the field. However, we believe that any framework that is consistent for both laminar and turbulent flow is reassuring, because it captures the fundamental physics. We cite Charru et al. (2004) to provide acknowledgment to the article where this framework was developed even though it was under laminar flow.*

- Chapter 3. I think the description of the surveys of longitudinal profiles could be shortened by half, as the quantification of slope is not so critical in the study. Instead, I would say something more on the correlation between long-term gauging data and short-term measurements in the study segments, and on the range of discharges measured during the short-term water stage measurements.
> *We agree that the descriptions of slope are perhaps more detailed than necessary. However, it is one of the more common questions asked during presentations of these data. Therefore we include it here for completeness. Slope does represents a crucial piece of our quantification of shear velocity through the depth slope product.*

- Pg439, row 7: How did you calculate flow resistance?
> *Flow resistance is calculated using a modified version of the Keulegan equation (equation 5 in this manuscript). We have added additional text to note the equation used and where it appears in the manuscript.*

- Pg 439, row 11: Because a single grain size was used, I can agree that 150 tracers are enough to describe the movement of sediments. However, if compared with the amount of RFID tags used in previous studies (see for example table 1 of Bradley and Tucker 2012) 150 tracers appears to be quite ~~a~~ few, and this is especially true for the 50 tags used in Bisley 3. Could you better justify that the number of pit tags are enough for the objectives of the study? Or otherwise discuss a little on how a larger population of tags could have changed the results?
> *It is difficult to justify a priori the number of particles needed to determine relevant results. It is because of this that we installed a second population of tracer particles in the Mameyes river. The close agreement of the results between both populations of tracer particles justifies that the results are repeatable. We do not expect that more tracer particles would result in different results, however it is likely that more tracer particles would result in less variability and stronger statistical results (i.e. particle distributions might not suffer from undersampling in Figure 7a). It is for this reason that we have excluded single flood results from near-threshold floods where only a few tracers have moved. We have added additional text in the discussion section on the subject of the results susceptibility to number of tracers.*

- Pg 439, row 15. In general, I understand that working with a single grain size equal to D50 of the bed is easier, but how representative is that in a poorly sorted bed? Could you better justify this experimental choice?
> *Our aim with this research was to understand bed load transport dynamics for tracer particles at length and timescales not allowed by laboratory experiments. We chose to center our tracers around the D50 of the stream in order to facilitate equal mobility among the tracers. The grain size distribution of the bed is a unimodal log-normal distribution for which the median is an adequate*

descriptor. Since our aim is to understand bed load transport dynamics we chose to exclude the largest immobile particles and our drill fractured all cobbles with median diameters below 5 cm. We do not attempt to claim that the tracer particles are representative of the entire stream bed. Thus we have restricted our analysis to the tracer particles only, except in the analysis of downstream particle sorting. The sorting results suggest that the particles behave dynamically similarly to the stream bed, however we do not suggest that even this dynamical similarity means that the stream is currently sorting at this rate.

- Pg 439, rows 26-29. Recovery rates are very high, whereas in literature smaller percentages are reported (e.g. Lamarre and Roy 2008, Liebault et al 2012). It would be interesting to have a little discussion about it. Is it due to the reduced transport distance and relatively low magnitude of floods surveyed?

> *We worked diligently to insure high recovery rates as previous analysis in Phillips et al. (2013) required them. We cannot rigorously comment on why recovery rates in this study are higher than other studies (though ours are comparable to Bradley and Tucker, 2012). Our final transport distance is ~1/2 of the final distance reported by Liebault et al. and about 10 times that of Lamarre and Roy. Liebault et al had to cover greater distance and about twice the area, however they report that they could complete a full survey in 4-5 days. Therefore survey time was not necessarily an issue in their river (the Mameyes takes a full 7-8 days for the final survey). Neither study reports the range of floods in terms of shear velocity or shields stress thus making direct comparisons of flow magnitude impossible. Comparisons of discharge are not quantitative as we lack a stage discharge relation for their rivers (though ratios of peak discharge to the critical discharge are under 3 for Liebault et al., and greater than 10 for the Mameyes). The durations of competent flow are comparable for both the Mameyes and the rivers in Liebault et al., and Lamarre and Roy. Liebault et al. report that they believe their missing tracers to have left their stream and joined a larger river where detection was not possible. From the published descriptions and data we cannot determine why the tracers in Liebault et al., traveled as far as they did (or vice versa why the Mameyes tracers didn't travel further).*

- Pg 440, row 13. It would be interesting to know how many tags were recovered on the bed surface and how many were buried (if the sediments were coloured, the first would be seen in the bed, whereas the latter would be detected by the antenna but not visible on the bed). Being able to demonstrate that most of the tracers were on the bed surface would reinforce the hypothesis that they moved under partial-transport conditions.

> *Unfortunately we were not allowed to paint the particles. However, we used two wands with different detection depths (10-20 cm and 50 cm) which can give us an estimate at the number of tracers near or on the surface as we took detailed field notes on which wand detected the tracer. In that tracers that could only be found with the larger wand could possibly be assumed to be buried beyond the detection limit of the smaller wand. If we make this assumption then the largest percentage of tracers buried beyond 20 cm (1.6 x D50) was 6%. This would indicate that the large majority of tracers were either on or near the surface. We have added text to denote this in the manuscript.*

- Pg 441, row 7. Why not test~~ing~~ the specific stream power as well? As pointed out by Ferguson (2005, Geomorphology), critical stream power is in fact unaffected by form resistance (as it is instead the shear stress), thus I guess you could more easily compare data provided by the two study sites.

> *We did not test specific stream power as it does not allow us to quantitatively compare our results to laboratory results and theory. Critical stream power as derived by Ferguson (2005) is indeed not affected by flow resistance, however it is affected by the critical shields stress as noted by C. Parker et al., (2011, Geomorphology) and re-derived by Ferguson (2012, WRR). The original derivation of the flow resistant invariant formulation of critical stream power assumed that the critical shields stress*

was constant throughout a stream profile, which has been shown to be incorrect. Flow resistance in the form of an inverse relative submergence (D84/depth at threshold) at the threshold of motion needs to be incorporated (Ferguson, 2012 WRR). Interestingly these two approaches may be quite similar in that both the critical stream power and the flow resistance corrected I require measurements of both the threshold and the flow resistance.*

An additional reason to avoid stream power in its simplest form (ω = pgQS/W) is that it implicitly assumes a rectangular cross section. When flow spills over bank discharge will continue to increase unabated, while stress derived from depth will increase at a substantially slower rate. In order to retain a width averaged stream power a separate relation for width needs to be incorporated.

- Pg 441, row 26. The method also implies that U*c is the same at the beginning and end of each flood event, which may not be the case (see for example fig 3 in Rickenmann 1997, ESPL). Long tails on falling limb of hydrographs can, in fact, affect very much the values of I*.

> *We have added additional text to this paragraph to clarify that the assumption of a constant threshold shear velocity is not likely to be true, and could potentially greatly bias the calculations of I*. We acknowledge this problem on line 15 p. 441 "...U*c, a parameter that is known to vary both temporally and spatially...", however there is very little we can do as we did not possess a way to measure the threshold for each individual flood nor is there currently any theoretical or empirical methodology to account for this effect.*

- Pg 442, rows 14-20. I think this could be deleted or at least shortened.

> *We include these lines as a reminder to the reader and colleagues engaging in tracer studies, or landscape evolution models for which these data may form a basis, that discharge is not an adequate variable for which to calculate sediment transport dynamics. In many cases the distributional form of discharge and stress are not the same (see Phillips, 2014 – chapter 4).*

- Pg 442, row 21. It is not so straightforward to me that the intercept on figure 5 should necessarily identify the critical shear stress. It seems to me that figure 5 shows the degree of partial transport experienced by tracers. According to Wilcock and McArdell (1997, WRR), for a certain grain class in an heterogeneous bed, partial transport correspond to a condition in which some grains are transported, and some are immobile. Looking at figure 5, if all grains are immobile f = 0, if they all move (full mobility) then f = 1. The trend showed by Figure 5 could thus be associated to a certain line of figure 3a in Wilcock and McArdell 97 for example. In the same paper, Wilcock and McArdell 97 associated the degree of partial transport to incipient motion. They report that incipient motion is related to certain percentage of sediment entrainment (that would be your f I guess) depending on grain size. I would suggest trying to apply their approach for better supporting the identification of the critical shear stress from data showed on Figure 5.

> *We originally attempted a range of methodologies to determine the threshold stress from the fraction mobile data, however we found no reason to discard a linear relation based on the laboratory results of Lajeunesse et al. (2010). Applying the approach of Wilcock and McArdell (1997) requires the determination of a reference transport rate determined from bed load flux measurements. A quantity that is requires several serious assumptions of these data, and a parameter that could not be measured in the field via conventional methods. Further hampering this approach is that at the low shields stress floods there are not nearly enough mobile particles to separate the results by size and achieve any meaningful trends.*

- Pg 443, row 11. I would use magnitude-frequency rather than frequency-magnitude.

> *We have changed "frequency-magnitude" to "magnitude-frequency".*

- Pg 443, pg 28. It would actually be interesting to compare the identified shear stresses for partial transport and full mobility with previous values available in literature. There are not many field evidences, but you could find some interesting values and reference if you go back to Lisle et al (2000, WRR) or Mao and Surian (2010, geomorphology).

> *We observe partial transport conditions over a threefold increase in the stress needed to entrain particles (τ^*/τ^*_c =~3.0). This is close to the same range reported in Mao and Surian (2010) where they find that the upper limit to partial transport is also a three fold increase in the stress needed to entrain the median particle size (τ^*/τ^*_c =~3.2). The results of Lisle et al. (2000) do not provide a threshold stress and therefore are not comparable, though based on observations they conclude that the range of partial mobility coincides with that proposed by Wilcock and McArdell (1997) which is $1< \tau^*/\tau^*_c <2$. The results of Haschenburger and Wilcock (2003, WRR) are given as a ratio of the bankfull discharge over the threshold discharge, which are unfortunately not comparable without a stage discharge curve. We have added additional text and references to the discussion section to place our reported range of shields stresses for partial transport within the context of previously reported values.*

- Pg 444, row 17. I don't fully understand the need of normalizing transport distance by grain size if all tracers were approximately of the same size (as stated at page 439, row14).

> *The range in grain size for the tracers is narrow when compared to the stream, however there is approximately a factor of six difference between the largest (b-axis = 27 cm) and smallest tracers (b-axis = 5 cm). Normalization by the particle diameter also serves to non dimensionalize the results.*

- Pg 445, row 2. Here I would try to better justify why the intercept is meant to identify the threshold stress.

> *Here as else ware we follow the laboratory and momentum framework in Lajeunesse et al., (2010) where the intercept represents the point at which the shear velocity is no longer able to transport sediment.*

- Pg 445, row 17. Data showed on figure 9 could be somehow related with recent mean transport distance plotted versus the excess of cumulative stream energy as recently done by Schneider et al. (2014, JGR)? Could the slope of regression lines be compared for example?

> *We agree that these two data sets could be compared, however not without considerable difficulty. Only a small amount of the data in figure 9 refer to mean displacement from single floods, which could be compared using an excess stream power or cumulative stream power function (though as noted above stream power may not be the best relation). Rough calculations place the mameyes data near the data from the Lainbach from Gintz et al. (1996) (lower left quadrant of figure 6a, in Schneider et al., 2014). Though due to the low recovery rates in Schneider et al. (2014) the mean values reported may not actually represent the true mean of the transport distances. The fate of the missing tracers needs to be determined in order to determine if the reported mean is an under or over estimate. Much of our data is integrated over several floods, which is not comparable to the data in Schneider et al. (2014). Where appropriate we have added the citation for this recent work.*

- Pg 447, row 15. Because the tracers were more or less of the same size, how relevant is this analysis considering that the actual grain size curve of the bed is much wider than the grain size of the tracers? The analysis is definitely of interest, but I think that the interpretation here tends to be speculative. Could you further stress the potential limitations on this interpretation?

> *The analysis is generally only for the tracer particles, and to understand how the sorting profile for*

the tracers develops. Taken by itself this result would be difficult to justify for the entire stream bed (we do not attempt it), however when coupled with the selective deposition theory and sorting results from the stream (fig 11.a-e) these results suggest that the tracer particles are dynamically similar to the stream bed. We have added text to the manuscript to more fully delineate the applicability of the results between tracer specific and general to the stream. This is a challenging question as to our knowledge there is very little information on how a subset of a grain size distribution behaves relative to the whole distribution. The work of and preceding Wilcock and McArdell (1997) suggests that different size fractions within a bed generally behave mechanically the same, only transporting at different rates. The installed tracers cover the range of grain sizes in the stream from D16 to D84 suggesting a decent overlap, however without a study designed explicitly to test this question we are unsure. Though we generally see no reason as to why other coarse particles should not behave in a similar manner.

- Pg 449, row 16. I'm left wondering if, really, bedload movement as single step lengths is necessarily coincident with bedload under partial transport conditions. Could partial transport occur when particles are moving with multiple steps and rests? Could it depend on the duration of overthreshold discharge as well?

> *We do not see a reason as to why partial transport precludes multiple steps, rather single steps could likely only occur under partial transport. We agree that given a long enough flood particles would likely take multiple steps if the stress exceeded their local threshold for entrainment. However, what we have observed is a series of partial transport floods which displayed a tendency for particles to embark on single steps. We have added additional text in the discussion section to clarify under which conditions these results pertain to.*

- Pg 450, row 27. Could the presence of pools explain this as well? Biron et al (2012, RRA) could be a useful reference to be cited here.

> *We did not observe a clustering of the tracer particles within pools. The presence of several pools could have enhanced trapping of particles for a period of time, but unlike the results of Biron et al., (2012) the Bisley particles readily moved through and beyond several deep pools.*

- Pg 451, row 9. If the Bisley 3 is a step-pool, boulder stream, the D84 is probably not a good descriptor for flow resistance, as form resistance could play a crucial role in energy dissipation. I would suggest to use a different formula or approach to do the analysis. Are results obtained using the Rickenmann and Recking (2011) formula's comparable?

> *We tried a variety of resistance equations (Keulegan, Variable Power Equation, Manning-Strickler and variants there of), and most of them provide a similar collapse of the data. We noted this in the text on pg. 451 line 18. The D84 represents a non trivial fraction of the flow depth at 0.55 m in the Bisley reach.*

> *Additional references cited*

Charru, F., H. Mouilleron, and O. Eiff (2004), Erosion and deposition of particles on a bed sheared by a viscous flow, Journal of Fluid Mechanics, 519, 55–80, doi:10.1017/S0022112004001028.

Ferguson, R. I. (2012), River channel slope, flow resistance, and gravel entrainment thresholds, Water Resour. Res., 48(5), W05517, doi:10.1029/2011WR010850.

Gintz, D., M. A. Hassan, and K.–. H. Schmidt (1996), Frequency and magnitude of bedload transport in a mountain river, Earth Surf. Processes Landforms, 21(5), 433–445, doi:10.1002/(sici)1096-9837(199605)21:5<433::aid-esp580>3.0.co;2-p.

Parker, C., N. J. Clifford, and C. R. Thorne (2011), Understanding the influence of slope on the threshold of coarse grain motion: Revisiting critical stream power, Geomorphology, 126(1-2), 51–65, doi:10.1016/j.geomorph.2010.10.027.

Phillips, C. B. The control of grain-scale mechanics on channel form, landscape dynamics, and climatic perturbations in gravel-bedded rivers, 2014, http://repository.upenn.edu/dissertations/AAI3623912

Schneider, J. M., J. M. Turowski, D. Rickenmann, R. Hegglin, S. Arrigo, L. Mao, and J. W. Kirchner (2014), Scaling relationships between bed load volumes, transport distances, and stream power in steep mountain channels, J. Geophys. Res. Earth Surf., 119, 533–549, doi:10.1002/2013JF002874.

Wilcock, P. R., and B. W. McArdell (1997), Partial transport of a sand/gravel sediment, Water Resour. Res., 33(1), 235–245, doi:10.1029/96WR02672.