

## Jens Turowski

First, both reviewers point out that a large body of relevant literature has been overlooked, which should be acknowledged and discussed. I agree with this point. The reviewers have pointed out a number of relevant papers. I add to these the recent work of Blom et al. (two GRL papers in 2017 and one JGR paper in 2017), who discuss gravel-sand transitions, equilibrium and quasi-equilibrium states and the effects of discharge variability, and of Pfeiffer et al. (PNAS 2017), who investigated the morphology of gravel bed channels in the USA. In addition, the authors may want to look at the literature on steep streams – there has been recent work on the interplay of bed roughness, flow hydraulics and sediment transport that may be informative for some of the discussion. A comprehensive study on flow velocity has been published for example by Rickenmann and Recking (WRR 2011), and on bedload transport for example by Schneider et al. (WRR 2015). There may be other relevant literature that has not been mentioned in the reviews or by me – the newly provided references may be a good starting point for a wider research.

We thank the editor for his comments and suggestions. We have a new and expanded introduction that encapsulates previous work/thinking on equilibrium channel geometry --- especially on rational regime theory, influence of cohesion on channel width, and related approaches. We have a few points of contention related to the references suggested by the editor:

We are aware and appreciative of the work done by Blom et al. Indeed, several theoretical and empirical constraints used by Blom and colleagues to develop their recent analytical models utilize work from our group (Jerolmack and Brzinski, *Geology* 2010; Miller, Retiz and Jerolmack, *GRL* 2014; Phillips and Jerolmack, *Science*, 2016). It is not the intent of this paper, however, to examine discharge variability or other effects. I note that Blom and colleagues' treatment of the gravel-sand transition is similar to our treatment here, so there is no discrepancy. This work is now cited, but a deeper comparison is not really warranted in our opinion.

1. We, and others, have attempted to replicate the results of the Pfeiffer et al. paper with no success. First of all, their statistical treatment of the data is insufficient to demonstrate a real effect of sediment supply *at best*; at worst, their statistical methods are faulty. Moreover, their principle conclusion is dependent upon the biases implicit in the assumption of slope-dependent critical shields stress; for rivers in their data set where critical Shields stress was actually independently constrained, these values indicate rivers only slightly above threshold. Finally, those authors did not actually measure or report hydraulic geometry or discharge (despite the paper title) so one cannot assess whether the hydraulic geometry scaling of their channels is distinct from others. We do not wish to debate the merits of that paper this one. In addition, the main point here is that rivers are *bimodal* in their transport stage, because gravel-bed rivers don't care about bank material, but sand-bed rivers do. Even if the *slightly* higher values for critical Shields stress in the Pfeiffer et al. paper were real, these data would still cluster with the near-threshold gravel. None of that influences our conclusions.
2. While we do agree that the inclusion of steep stream data in the global data set being analyzed would potentially provide interesting insights into the variance within the coarse-grained portion of the data set, the point of this paper is not to provide illumination into this aspect of the science. The objective of this paper, which, upon suggestion from the editor, has been more clearly delineated in the paper, is to test whether or not the bimodality of sediment transport

states survives in light of evidence to the contrary and to provide a hypothesis as to why or why not. We do not believe that the inclusion of steep stream data and further analysis of flow velocity or bed roughness will assist in this endeavor. Note that we have not explicitly excluded steep streams; many in our dataset are up to and even above ten percent. It is likely that flow resistance and other factors correspond to the variance of transport stages seen within gravel-bed river data. But the main point of this paper not the variance within either near- or far-above threshold regimes; rather, it is the separation of data into these two regimes, and its relation to when critical stress for bed material drops below that of the banks.

3. Overall, we want to properly cite the previous literature that is relevant to our study, but we don't think it is useful to exhaustively review all work done on equilibrium channel geometry. This paper is about bi-modal states, and a potential explanation for the two modes; it is not about what happens within each mode.

Second, both reviewers fail to see the novelty in the work. I ask you to clearly delimit where the paper is a mere review of published results and where you go beyond what has been done before, in particular in light of the additional literature mentioned above

It is evident that we have failed to clearly state where our work fits into the wealth of pre-existing work on regime theory. Thus, we add the following clarification to our manuscript:

“Regime theory” is the application of these agreed upon relationships with the addition of one additional threshold channel based-assumption to allow for closure. There are three dominant branches of regime theory, each with their own form of a threshold channel closure assumption that separate regime theory into three distinct schools of thought: 1) assume that river are canals, and thus threshold channels; 2) assume that the transport regime is purely bedload and solve the 2-D flow field to balance fluid shear stress and particle weight at the edge of the channel, while simultaneously allowing for transport at the center; 3) assume that the river undergoes an optimization process that maximizes friction in order to reduce fluid shear stress, ultimately resulting in a threshold channel.

The first school of thought is based upon work done to calculate the shape of a stable canal for which the bed material is at the threshold of motion (Glover and Florey (1951)). This work has been extended to natural rivers by Henderson (1961) and offers an explanation for observations of alluvial river width relating to the water discharge Henderson (1961); Andrews (1984); Metiver et al. (2017)). This line of thinking links well in with the second branch of a regime theory which as was established by Parker (1978) which solved the 2-D stress field to show that, for a pure bedload river, the channel is at the threshold of motion for the material at the banks and slightly above the threshold of motion in the center, allowing for the river to transport sediment, while at the same time maintaining a stable and consistent width. This model is supported by both global compilations of data and case studies of individual rivers that demonstrate that gravel-bedded rivers that translate their sediment load as bedload are slightly offset from a threshold of channel Phillips and Jerolmack (2016); Gaurav et al. (2015); Metivier et al. (2016)). Parallel to this grain size-dependent channel geometry is the concept of optimization which assumes that rivers seek a threshold channel condition by maximizing the flow resistance within the channel to minimize the fluid shear stress (Eaton and Church (2004); Eaton and Church (2007)). The rational regime theory put forward by Eaton attempts to infer the importance of bank strength given deviations away from the threshold condition that is posited by optimality theory

(Eaton and Church (2004); Eaton and Church (2007)), however they are predominantly calibrated on coarse-grained rivers where research has shown that the influence of cohesive mud is a minor control on the erodibility compared to the weight of the gravel (Kothyari and Jain (2008)). What distinguishes our work from this work is that we extend the concept of Parker's threshold channel model into the space occupied by fine-grained rivers by the suggestion that river channel geometries, and their subsequent sediment transport state are either controlled by the erodibility of their beds or their banks. This paper shows the transition from rivers that can be explained entirely by Parker's theory (i.e. channel beds and banks composed of uniform material transported entirely in bedload) to channels that cannot. For natural rivers, this transition most frequently occurs at the transition from a gravel-bedded to a sand-bedded condition. This transition coincides with the point at which bed material becomes small enough such that the cohesion of channel banks should become important. What we show is that sediment transport state is bimodal because grain size is bimodal; the coarser gravel mode is more difficult to entrain than any cohesive bank material, while the finer sand mode is easier to entrain than any cohesive bank material (if present).

[NOTE TO THE EDITOR: NO ONE HAS PROPOSED THIS, THAT WE ARE AWARE OF; THIS IS THE POINT OF OUR PAPER - NOTHING MORE, NOTHING LESS.]

There are a large number of other critical points by the reviewers, which I ask you to address in detail. In particular, I would like to highlight the closing comment of reviewer 2. You have worked with data measured and compiled by many other scientists, and these should be duly acknowledged. The least that should be done is the addition of stream and site names, coordinates, and suitable references in the data tables provided in the supplementary material. Otherwise, researchers will have little chance to scrutinize your data and your results in the future.

We thank the editor and reviewer 2 for this suggestion and have done so. The objective of the original formatting was to provide a "trimmed down" version of the cited data sets to provide only the data that was of immediate relevance to the analysis done in this paper. We have re-included the suggested information.

### **Maartin Kleinhans**

It is not exactly clear what precisely the argument of the authors is. It seems to be that "there are mainly sand-bed, suspension-dominated rivers and gravel-bed, bedload dominated rivers because of the scarcity of the intermediate sizes on this planet and because the typically available floodplain materials are strong enough to withstand the typical bed shear stress in sand-bed rivers but not in gravel-bed rivers." The arguments revolve around the scarcity of sediments between sand and gravel, discussed by the authors, and the scarcity of sediments between sand and floodplain fines, not discussed by the authors. If this is correctly represented then these ideas have already been published in many papers over the past decades and are therefore not novel. If not correct then the argument is not clearly enough presented.

We hope that the newly included paragraph on regime theory addresses what our argument is. We agree with the reviewer that the statement he makes has been done before. We disagree that this is the main conclusion of our paper. We agree (of course, since one of us has worked a lot on this problem!) that river-bed grain sizes are bi-modal; that forms one important piece of this puzzle. Bi-modal sediments are necessary *but not sufficient* to produce bi-modal transport states. What we show here is there is a second constraint; cohesive bank materials. When channel beds are composed of gravel they don't care about the banks, which are always easier to entrain; here we revert to a Parker-like condition of the bed-load river (and any variants of it). Only when beds are composed of sand do the banks matter, because sand is the easiest to entrain material; it is only when sand beds are combined with cohesive banks that rivers depart from Parker-type bedload river predictions. We are not aware of any paper that advocated for this position.

An elegant way, in principle, to look at the relation between bed mobility and channel geometry is the dataset of Singer (2010). This shows a gradual downstream reduction in bed shear stress but a zone with a bimodal reduction in median particle size and in excess mobility due to patchiness of bed material (also found in Paola and Seal 1995). However, the authors here invoke this observation to argue for bimodal behaviour in the entire dataset and this argument is flawed. Here, and on P3 L14-16, important references are missing, such as Yatsu (1955), Paola and Seal (1995) and work by Frings (review in 2008). The large body of literature on downstream fining bracketed by these references show three things. First, many fluvial sediments are bimodal because of the nature of sediment and the manner in which it wears down. The effect is that there are very few rivers with dominantly fine gravel in their bed. Second, in a bimodal sediment distribution where the two modes gradually change in height, as in downstream fining where the finer mode increases in abundance as the coarser mode reduces, the median grain size flops suddenly from one mode to the other where the mean size would not. This means that the representation of downstream fining by the median size is misleading the authors to believe in the bimodal behaviour. Third, many gravel-sand transitions are rather gradual, except in many relatively small rivers (Frings 2011). This is partly due to patchiness, bend sorting, and the transition from clast-supported to matrix-supported sediments. Elsewhere it has also been argued that many rivers show a third mode of sediment: the silt and clay that form the floodplains. Usually the material between that mode and the sand mode is scarce too (see Kleinhans 2010 for review)

It is unclear to us what the problem is here. There is no point of disagreement between the reviewer's statement above and our position. First, the idea that "important references are missing". We can always add more references, and so we have. However, that misses the point; we are not trying to conduct an exhaustive review of the gravel-sand transition; a topic that one of us is well aware of having worked on it for quite some time. Yes, we are aware that sediment is often bimodal because of how it wears down. Indeed, we are presuming bimodality - and we demonstrate it with the data not as a "new" topic but just to verify that point. As for the second point of the reviewer, it is unclear what he is advocating for. He refers us to Paola and Seal --- which is of course all about patchiness in the gravel-sand regime, basically showing that gravel and sand segregate into patches. But then the reviewer makes the point that the sudden switching of grain size is an artifact of choosing the median value --- and by extension that river sediments are not patchy. Besides, this issue does not contradict the premise of the paper; whether the transition itself is patchy, flickering and rapid or more gradual and smooth, the main result is that within a river we see the same behavior that we see with the whole data set. *That*

is the point of the Singer data; not to rehash the entire gravel-sand transition problem, just to show that one river does the same thing as the global dataset, showing that patterns in the global data are not merely some artifact of data treatment. Finally, for the third point; it doesn't really matter how gradual or abrupt the gravel-sand transition is. That said, the reviewer fails to point out a paper that one of us co-authored (Miller, Reitz and Jerolmack) where we showed a collapse of sorting profiles across the gravel-sand transition for rivers big and small. The upshot is that the length of the gravel-sand transition is roughly ten percent of the length of the upstream gravel reach. Interesting? Maybe. Universal? Maybe not. Important for our analysis here? No.

The systematic trend of increasing bankfull Shields number with decreasing grain size is disputed on the basis of two pieces of evidence related to published experiments. The first is that seepage channels formed in experiments with sand cluster at the threshold for motion, and the second is that other sand-bed channels in the laboratory likewise are at the threshold for motion. But this misses the point of scaling in experiments entirely. The Shields number is the relevant scale for sediment mobility, and in experiments that scale down the size of systems in nature by orders of magnitude, of course the sand is near the threshold for motion because it simulates gravel and even at sand size needs steeper slopes than in reality. This is why other experiments with self-formed seepage channels were conducted with low-density sediment (Marra et al. 2014), and consequently the mobility ranged from 0.044-0.68.

We appreciate the reviewer pointing out some additional references that can be useful for us. We have incorporated them into the revised manuscript.

We dispute the contention, however, that threshold sand channels in the lab are simply an issue of scaling. We also dispute that using plastic sediment allows one to make suspension rivers in the lab with non-cohesive sediment.

A first important point we want to make is reporting Shields stress alone is not adequate to estimate transport stage for laboratory experiments. Lab experiments of channelized flows often span the laminar to turbulent regimes, and the hydraulically-smooth to rough regimes; in addition, they have varying aspect ratios that greatly affect flow resistance and roughness. In short, there is no way to accurately know what the critical shields stress is in a laboratory setting without a) explicitly measuring it, or b) allowing a channel to reach an equilibrium *and static geometry, without sediment feed*, and then inferring critical through a threshold channel assumption. The only effective way, in our opinion, to estimate critical given all of these issues is the second option; for the prescribed flow and grain size, allow a stable channel to evolve to a point of no transport by providing no sediment feed. This is what the IPGP group in France does (Seizelles et al., 2014-2014), and in transitional Reynolds number flows they get quite high critical Shields numbers, like 0.2 - 0.4. Note also that threshold Shields determined in laminar experiments can be quite high, up to 0.5 (Charru et al., JFM 2004; Lobkovsky et al., FGM 2008; many others).

In the Marra et al. experiments, the high stress runs were indeed noted for the runs with lower density sediment, however the Reynolds numbers were also relatively low (72-541), potentially resulting in a higher than expected critical entrainment stress. Moreover, the river channels studied by Devauchelle et al. (2011) are not laboratory rivers; they are natural rivers that have both their beds and their banks comprised of non-cohesive sand. This system is indeed small and atypical, however in this case, the data

show that this is a case of the exception proving the rule. In the rare case of a sand-bedded river with banks that are equally as erodible as the bed, we see that the system behaves identically to a gravel-bedded river with banks that are, at most, equally as difficult to erode as the bed. Finally, parallel experiments at UPenn and at Saint Anthony Falls Lab (U. Minnesota) attempted and failed to make self-formed suspension rivers by using low-density sediment and a range of discharges. These results haven't been published so we can't refer to them; but over multiple orders of magnitude, channels evolved so that the banks were at the threshold of motion.

Furthermore, there are a large number of other experiments with self-formed channels (compiled in Kleinhans et al. 2014 Fig. 9 and Table 1 and some more in 2015) that show a range of mobilities in the context of a large river dataset that also show a gradual transition from bedload to suspended load dominated behaviour. This includes high-mobility experiments with sand and floodplain-forming low-density sediment (Paola group, Minneapolis) and with materials that strengthen banks (famous experiments of Hoyal, ExxonMobil) (see Kleinhans et al. 2015 for references). The main point is that the nature of the bank material determines the channel geometry. This is essential for bar and river patterns and was incorporated in the hydraulic geometry in Kleinhans et al. (2015) because correct scaling of the bank material in experiments leads to desired channel aspect ratios. The qualitative understanding behind this work goes back at least to Ferguson (1987), and shows that the experimentally inferred relation of the straight-meandering-gravel transition as a function of slope by Schumm and Kahn (1972) is wrong and the bank stability is key. Furthermore, the natural and experimental channels adhere well enough to the hydraulic geometry relations of Parker et al. (2007) when accounted for the strength of the banks (Kleinhans et al. 2015), which allowed application of the Parker relation to bar theory. It is already well-known that the absolute dimensional shear stress in gravel-bed rivers is much larger than that in sand-bed rivers despite the lower sediment mobility, so that bank material unstable in the gravel reach can be stable in the sand-bed reach. The hypothesis on P10 L14- 16 has already been voiced in a large body of bank erosion literature underlying the BSTEM model by Simon et al. (2000 and later references). This argument has not become more precise by the scatter plot in Fig. 7. Also the authors do not present new data or new insights; for example P10 L27-32 states that information about bank material is usually lacking and this is precisely why Kleinhans and van den Berg (2011) had to introduce that counterintuitive streampower measure and why Eaton and Church (2007) developed a rational regime theory that can be used to invert to bank stability measures from observed channel dimensions. These papers and many others already make that point; however, these references are lacking here.

We hope that the addition of our section summarizing regime theory, and describing where our work fits in, clarifies the goal of our hypothesis. We have added even more references that we hope will rectify the perception that we are ignoring previous work (although we note that we were aware of most of this work already; it was left out only because we wanted to construct a simple and concise point without rehashing all of the work done on equilibrium channel geometry, gravel-sand transitions, etc.) Nothing we say runs contrary to the work mentioned here; however, we strongly disagree that there is nothing new or more precise here. The issues raised above about bar theory are not relevant to the issue at hand. As for the other points; we do not dispute that these previous studies have addressed controls (or potential controls) of cohesion on bank geometry. But none of these papers have shown this: the Parker-type bedload model works to first-order for predicting channel geometry, for all rivers in which particles are large enough that cohesion doesn't matter. Data break away from this prediction

only when bed material becomes small enough for cohesive bank materials (if present) to matter. We are not attempting here to invert each data point in the scatter to infer bank strength, or to introduce new parameters to model or explain things. The simple and novel point of this paper is that we can mark the transition from rivers not caring (to first order) about bank material, to caring about bank material; and that this happens because river sediments are bimodal AND that the smaller mode is the easiest-to-entrain material. We actually believe that this result is entirely consistent with all of the references raised by the reviewer; however, we stand firm that none of these previous references has explicitly made this point.

The final remark of the discussion is somewhat simplistic: "Some of the scatter in hydraulic geometry scaling plots may be due to stochastic fluctuations around the mean behavior." Here, stochastic is a rather dirty term for a number of well-known sources of variability. Apart from the fact that vegetation creates its own patterns with its own scales on the banks, leading to variations in channel dimensions, the balance between bank erosion and floodplain formation is not always in equilibrium because of the specific length and time scales of bank failure and floodplain / levee formation. The assumption of the authors is that measured channel dimensions in their compiled dataset are due to such variations, but it is quite likely that the original authors of the datasets underlying this compilation already averaged out such variation. In fact, this is certainly the case for the Kleinhans and van den Berg (2011) dataset and in the experimental datasets from our laboratory, that were also taken up in the compilation of Metivier, and are therefore possibly part of the present compilation

This is a fair criticism. We have changed the text accordingly. Note that the main point here is not to explain variation within the two clouds of data (near-threshold and far above threshold), but rather to make a direct point about why there are two clouds. Reviewers are right to be interested in the origin and nature of the deviations from the first-order trend of the near-threshold channel. But this paper does not address that, and it cannot given the limited information from the datasets drawn upon. We have chosen to utilize a large dataset with minimal parameter information rather than a small dataset with more information --- because we want to focus on the scaling and bi-modality.

### Reviewer 3

On page 5, line 10, you state that you assume the Chezy friction factor is constant. I can see no possible justification for this assumption. Looking at Ferguson 2007 and Eaton and Church 2011, it is clear that  $C_f$  varies dramatically with relative roughness, and that changes in relative roughness also produce order of magnitude changes in the bedload sediment concentration generated at the same critical dimensionless shear stress. These scale variations are fundamental, and need to be included in the numerical analysis. This is one of the reasons that Millar took the numerical approximation approach that he did, and then generated power law equations using a Monte Carlo modelling approach (see Millar, 2005). This assumption appears to me to fundamentally undercut the entire following analysis.

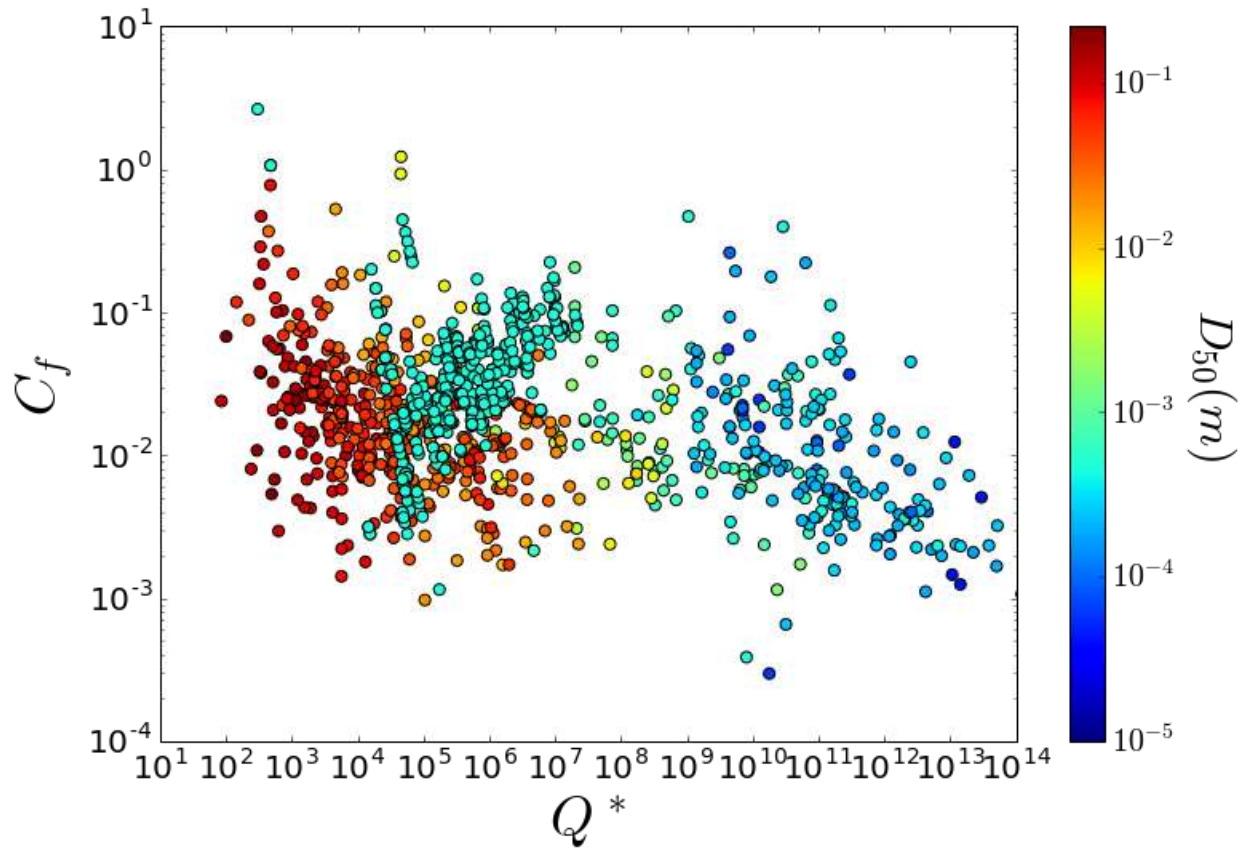
Chezy friction factor can indeed vary dramatically with relative roughness, however it does not vary strongly with dimensionless discharge and has no discernible influence on sediment transport regime

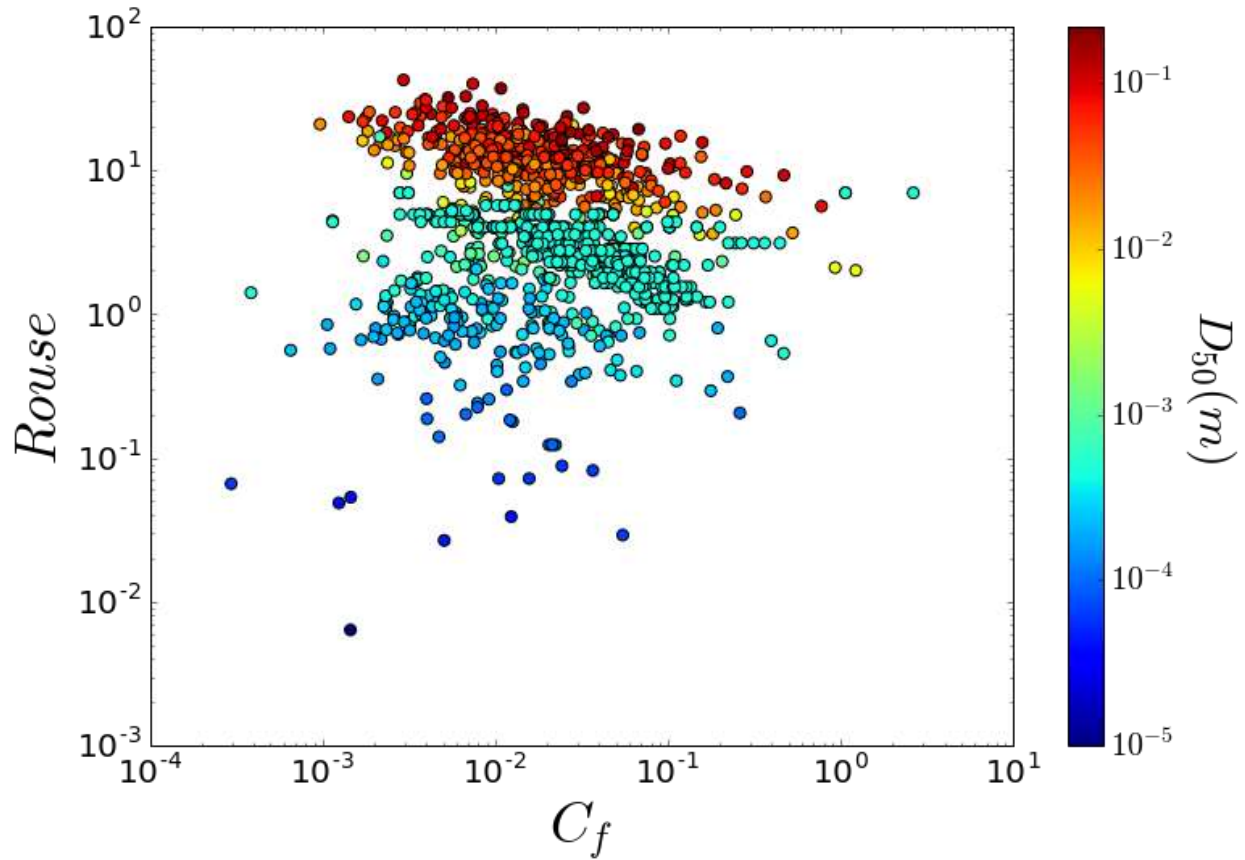
(see attached Figure-1 and Figure-2). For these reasons, we do not believe that the assumption of a constant friction factor is damaging, or particularly relevant, to our analysis.

More broadly, transport rate changes by orders of magnitude with small changes in fluid stress when one is near the threshold of motion. But, the difference between zero and a very small number can be many orders of magnitude, of course. What the reviewer is raising is related to the challenge of estimating critical. Flow resistance is one challenging piece to estimating critical, but it is not the only one. Indeed, the question of predicting critical remains a central challenge to geomorphology. But we do not believe that it will "fundamentally undercut the entire following analysis". First, related to the item above; while flow resistance varies quite a bit in our data, it is NOT bi-modal; so flow resistance cannot explain the apparent bi-modal transport states of alluvial rivers. Second --- related to points made to the previous reviewer --- this paper does not seek to explain the variance within clouds of data for sand and gravel rivers, but rather to explain the origin for two clouds instead of one.

Finally, *we do not actually assume anything in the treatment of our hydraulic scaling data related to the friction factor*. Therefore, none of our analysis is undercut. Rather, we simply plot the predictions from the threshold "canal" theory as a reference point, following Metivier et al.'s work. It would be impossible to generate a threshold prediction for the global dataset using a site-specific and varying friction coefficient. You can view the straight line in the hydraulic geometry scaling plots as the simplest, dumbest model; constant Shields stress (which of course is also not true), constant friction, no cohesion, and no transport! We know that this is not a complete model for rivers; the deviations from this model tell us something. Friction varies among rivers but not systematically with discharge. Therefore, friction factor cannot explain the persistent offset in scaling between the threshold channel prediction and the data.



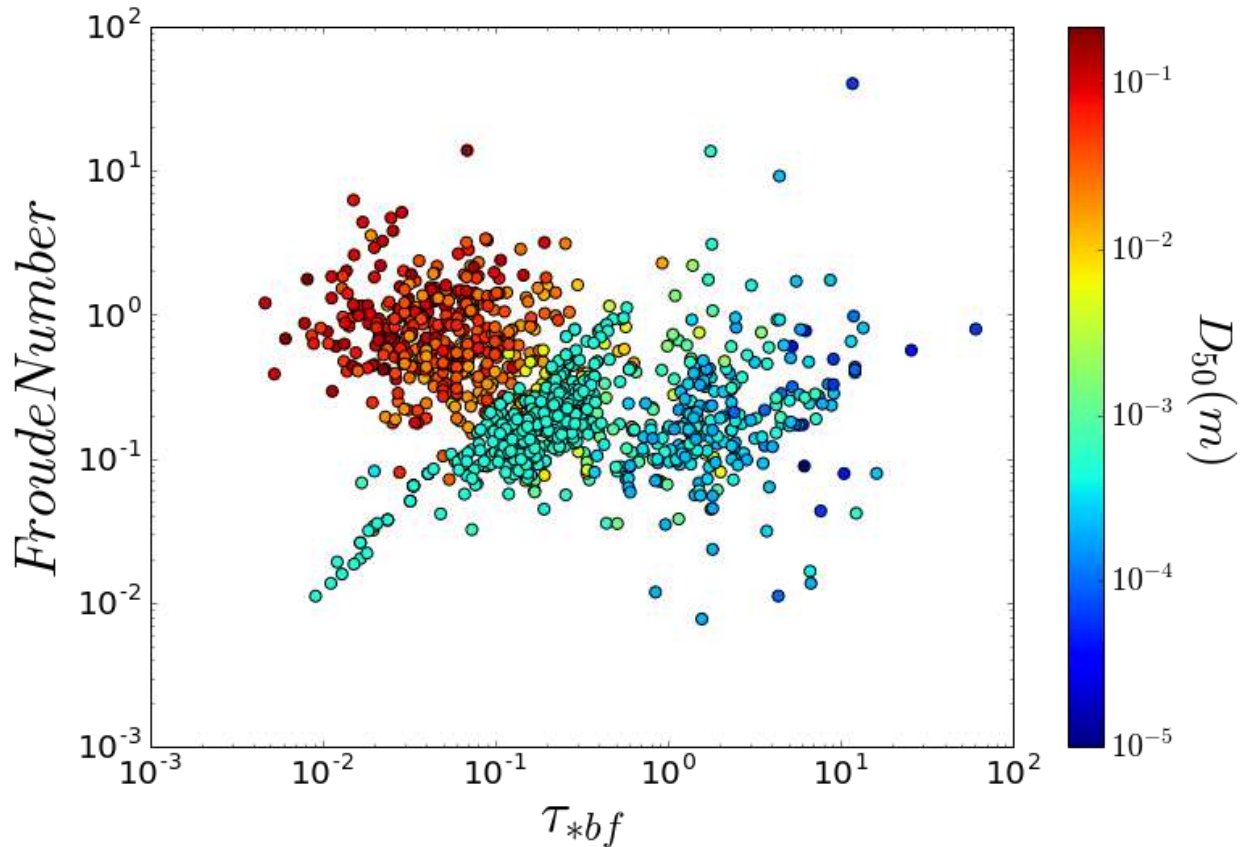




Page 1, Line 15: you claim that the empirical hydraulic geometry equations are robust, and show remarkably little variation in the exponents: this is a stretch. log log plots hide the true scale of differences between datasets; the lumped data includes systems where  $W/d$  ratio declines downstream, and where it increases downstream; it includes systems where depth is nearly constant, and ones where it changes as expected. You also fail to observe that simple Froude scaling nearly explains the trends that are found in nature, which indicates that most of the variance can be explained simply by the size of the system, not any particular organizing principle. This is well described by Eaton in the Hydraulic Geometry chapter of the Treatise on Geomorphology. In fact, there are significant variations in the exponents, when they are compared against the Froude scaling exponents, and I believe this is well described by Eaton and Church (2007) in JGR

We do not deny that scatter exists for empirical hydraulic geometry equations; however we believe that given that the data show approximately one order of magnitude of scatter over approximately fourteen orders of magnitude(!) of range, the relationships can be described as robust. As for Froude scaling, we have two points. First, Froude scaling does not explain the bimodal nature of sediment transport states (see Figure-3). Second, Hydraulic geometry relations (non-dimensionalized width, depth and slope as a function of discharge) are much tighter than any scaling with Froude number. This was already shown in Gary Parker's 2004 E-book, and can be shown with our data or any other; the most we can say about Froude number is that it increases with slope, doesn't exceed 1 by much, and shows a lot of scatter. Froude number and friction factor are similar in this regard; they show weak but detectable trends with some variables like slope; however, the variance around these trends is at least as large as the

magnitude of the trends themselves. Finally, in the Treatise on Geomorphology, it is explicitly said that "An important drawback is that it is not possible in this [Froude scaling] framework to explicitly consider (or vary) the effect of bank strength on channel geometry; thus, one of the important independent variables is omitted." It is not clear what we would add by considering Froude scaling in this framework.



Page 1, Line 20: you attribute the concept of the dimensionless discharge to Metivier (2016). This is certainly not the primary source for a dimensionless discharge. Andrews used it in his 1984 paper (though there is an unfortunate typo in the final version, he used exactly this definition of a dimensionless discharge).

We thank the reviewer for this clarification and will cite Andrews as well. We also note of course that discharge has been non-dimensionalized by many including Parker's 1978 formulation. We are not trying to attribute the "concept" of dimensionless discharge to anyone in particular, since it's not important; rather, we are simply choosing to follow one kind of dimensionless discharge for comparison to recent work.

Page 2, Line 1 It is remarkable that you do not refer to Ferguson's classic work on the basis for regime theory. I believe that this is still one of the best summaries of the problem, and provides a much more satisfactory statement of the problem than the authors provide here

We thank the reviewer for this suggestion and have cited Ferguson. As mentioned in response to the editor's comments, our goal has never been to provide an exhaustive review of all of the work that has been done on equilibrium channel geometry. This paper is about bimodal transport states and a potential explanation for the two modes.

Page 2, line 10: The "ground state" approach by Metivier sound exactly like the threshold channel approach used by many previous authors, including Lane, 1955; Henderson, 1966; Stevens, 1989.

We do cite Lane and Henderson. The term "ground state" was suggested as nomenclature to avoid the misconception that "equilibrium" implies a lack of dynamics. In other words, all channels wiggle around, and many colleagues we talked to were challenged by the formulation of the problem with threshold banks because they thought it was not consistent with dynamics. This is an attempt to clarify that point; that we are making a description of the river as if it is a static thing; in comparison to a natural river, you can think of this static description as describing the averaged behavior without reference to the scales of fluctuations around it.

Page 2, line 2: the discussion of the stable channel paradox seems to be to miss the point. The key thing to realize is that the shear stress acting on channel banks is lower than the average boundary shear stress, and that the shear stress acting on the bed is higher than average. This particular issue was very well addressed by Rob Millar's implementation of a regime theory that considered bank strength, using the shear stress partitioning approach by Knight, Flinham and Carling. I recommend reading Millar 2005 and the numerous relevant papers cited therein. I believe that Millar's contribution has adequately asked and answered the questions that this paper attempts to answer, and that the data analysis presented herein does not provide any additional insight to the problem. In any case, I do not see the justification for publishing this analysis without acknowledging the previous work!

Millar 2005 was focused on gravel-bedded rivers, not the transition from a bed-controlled regime to a bank-controlled regime. We hope that the inclusion of our paragraphs outlining regime theory and where are work fits in addresses this. Again, in the context of our paper, Millar's work is about addressing the variation of data within the gravel regime. Our paper is about why there are two clouds of data --- and why one follows "regime theory" of any kind (Parker, Millar, etc.) and one does not. We're happy to cite Millar's work in our revised version as part of our new summary on regime theory. We just wish to re-emphasize that our paper is distinct from this; neither Millar or other workers showed that (near-)threshold theory predicts first order trends for all gravel rivers, but that rivers with smaller grains depart from this trend; and that the hinge point is related to where bed material becomes weaker than bank material. While we don't disagree that there are deviations from the threshold model for the gravel regime --- which could be addressed by the work of people like Millar --- we point out that the threshold model explains the first-order trend of all gravel rivers, while it definitely does not for sand bedded rivers.

page 3, line 11: you state that there is no accepted model for the equilibrium geometry of rivers far above threshold. I think is is an unfair and inaccurate representation of the current state of the science.

There are models (like Millar and Quick, 1998) that successfully predict the geometry of such streams. They are published, and have been successfully tested, yet you appear to disregard all of these models without even bothering to mention them.

The data presented in Millar and Quick (1998) show gravel-bedded channels that have shields stresses up to approximately 3 times the threshold of motion. When we refer to rivers that are far above threshold, we are referring to the fine-grained rivers that are approximately 2 orders of magnitude in excess of the threshold of motion, for which there has yet to be established a satisfactory understanding.

Analysis in Fig 2: does this really tell us anything that we do not already know from Church's (2006) exploration of the plotting positions of various rivers? While the plots are somewhat different, I do not see what novel insight they provide, particularly given the scale distortions introduced by the inappropriate assumption that  $C_f$  is constant!

We do not understand the reviewer's meaning of "plotting positions". Figure 2 was meant to demonstrate a) the robust nature of hydraulic scaling relationships, and b) that sand-bedded rivers with banks and beds composed of non-cohesive sand plot in the same geometric space occupied by gravel-bedded rivers. This is a context figure, not the fundamental result (which comes later), and we don't see why it shouldn't be included. Note that the plot itself *makes no assumption about  $C_f$  - so there is NO scale distortion!* Only the lines showing the naive threshold model with constant everything would have such a "scale distortion." But note, as we pointed out above, that there is not positive scaling of friction factor with dimensionless discharge; and, therefore, friction cannot explain the deviation in scaling between the naive threshold model and the data.

Page 6, lines 1 to 5: Your text is not an accurate representation of the bank strength issue. With respect to vegetation, the effect on relative bank strength is fundamentally scale dependent (see Eaton and Giles, 2009, Eaton and Millar, 2017), which you fail to mention (and which also introduces significant scale distortions), and the effect on channel width is close to a linear one. As banks become very erosion resistant, then changes in width are reduced, because the channel is able to reach the hydraulically optimal form, beyond which narrowing does not produce any increase in bed shear stress (basically this is the geometry predicted by Wobus 2004 for erosion into bedrock).

We do not dispute that vegetation can have an effect on bank strength. Our hypothesis does not run contrary to this line of thinking at all. The overarching idea of our hypothesis is that alluvial rivers are either bed controlled or bank controlled. To a first order, fine-grained rivers are bank-controlled, however I am certain that in some cases in gravel-bedded rivers, vegetation is strong enough to shift the control on channel geometry from the bed to the bank. Phillips and Jerolmack (2016) demonstrated that, to a first order, the geometry of gravel-bedded rivers can be explained by near-threshold channel geometry. This sort of model, is not design to describe variation around the trend. In this paper, we are demonstrating a systematic departure from that trend when the grain size on the bed gets small enough.

Page 6, line 9: your use of Schumm's M findings is a poor choice, since that analysis is a classic tautology. Clay and silt are never found on the channel bed, so Schumm's index (which is the percent silt and clay averaged over the entire channel boundary) builds in the width depth ratio; therefore it cannot be used to make a meaningful prediction of the width depth ratio. Simons and Albertson (1963) do manage to make some progress on the sedimentological controls for canals, however

We do not use Schumm's findings, we merely mention him as one of the first people to consider the influence of bank composition on channel geometry.

Page 6, line 15: the authors present a hypothesis about what controls bank erosion, but that hypothesis was advanced previously by Nanson and Hickin, and is the basis for the models developed by Millar and Quick (1993) and by Eaton (2006).

We hope that our paragraph summarizing regime theory provides clarification of the uniqueness of our hypothesis.

#### **Andrew Wickert**

p.1,l.17-19: I find Tal and Paola (2007) to be a key reference for meandering vs. braiding when involving vegetation; perhaps it is just not cited due to the many others here, but I suggest that you read it if you haven't!

We do cite Tal and Paola 2007.

Equation 1: Consider splitting onto 3 rows for better readability

Done

p.2,l.1: remove indent

Done

Fig. 1: Why does this schematic have to have different bed and bank material? In particular, your statement that the center of the channel is only slightly in excess of critical Shields stress indicates that it is more likely all gravel that you are trying to illustrate. Do you intend to use this for both the gravel-bed and sand-bed channel examples? If so, just a little bit of clarification/generalization will be needed. (As a graphic, this is an excellent block illustration/generalization of the Parker (1978) line drawing, by the way, and one that I'd like to borrow for lectures.)

This is meant to illustrate a sand-bedded river. we have added clarification.

p.3,l.1: Parker (1978a) predicted... (no need for "Parker's model")

Done

p.3, l.22-24. Note also that Pfeiffer et al. (2017) demonstrate that gravel-bed rivers are significantly above the threshold of motion in rapidly-uplifting settings. They argue that this is due to armoring, though I have my own ideas about this (as-of-yet unpublished, and therefore nothing that you'll have to tangle with in a review).

We, and others, have attempted to replicate the results of the Pfeiffer et al. paper with no success. Their results of their work are dependent upon the biases implicit in the assumption of slope-dependent critical shields stress. As their results cannot be replicated under more rigorous testing, we will not use that paper as a reference. See comments above to the Editor on this.

p.3,l.24: remove comma after "slope"

Done

Fig. 2: Show legend in only one panel to not overlap with data

Done

Fig. 3: It looks in the figure like you have one point from seepage channels and one from experiments. Therefore, I suggest that you reword "River channels ..." to take into account that this is not quite so plural, and is not tested to be generalizable.

The single, larger points are meant to illustrate the mean of multiple measurements taken along a single longitudinal profile or multiple runs in a laboratory setup.

p.8,l5-6: Which site-specific variations are you considering, and how/why are they important?

The term "site-specific" was meant to refer to the grain size and slope. We merely wanted to state that there are multiple ways of estimating critical.

p.8,l5-6: Which site-specific variations are you considering, and how/why are they important?

The term "site-specific" was meant to refer to the grain size and slope. We merely wanted to state that there are multiple ways of estimating critical.

p.11,l.1-3: Could you discuss other reasons for cohesion? Interlocking grains and capillary forces come to mind. Are these significant compared to the surface charge effects

Capillary forces would not be a factor because the bank toe is perennially saturated. We hypothesize that part of the reason for the huge range in bank critical shear stresses from previous work is because previous studies sampled from a variety of locations up the bank, which would then incorporate capillary forces despite those not having an effect at the bank toe. We do not believe that the interlocking of grains is a significant cohesive factor because grain interlocking would happen in both cohesive and non-cohesive settings, however we see that for non-cohesive systems rivers adjust themselves to the threshold of motion. Perhaps grain interlocking is implicitly accounted for in this clustering around the threshold of motion.

p.11,l.3: material; → material:

Done

p.11,l.31-32: Would you like to discuss some of the reasons for the low frequency of channels with 1-10 mm grain size? In particular, do you think that this may have to do with the crystal size / granule break-down problem, or possibly be connected to the transition between cohesion-dominated banks and particle-weight dominated banks that makes these grains either difficult to move or whisked away in a larger-clast gravel-bed river? This is of course ignoring arguments for equal mobility...

While this is indeed an interesting question, we do not believe the grain size gap found in channels is of particular relevance to this paper.