Interactive comment on "Controls on the hydraulic geometry of alluvial channels: bank stability to gravitational failure, the critical-flow hypothesis, and conservation of mass and energy" *by* Jon D. Pelletier

Jon Pelletier

jdpellet@email.arizona.edu

Received and published: 28 October 2020

I wish to thank the reviewers for their insightful comments, which have led to a significantly improved paper (which I will submit if the editor deems my responses below to be sufficient for the review process to continue). I apologize that it has taken so long to provide responses. Here I summarize the points of each reviewer, respond to each point, and state how the revised manuscript has been or will be modified to address each comment.

Reviewer 1 (Gordon Grant):

Q: "I remain somewhat skeptical of both the approach and the results. My skepticism is rooted in both some of the underlying theory and assumptions used to derive the mathematical relations that underpin this work, and also the wide variance and assumptions in the data used to validate the methods. I cannot put my finger on a single "smoking gun" amidst all of this, but I was quite surprised that given all the assumptions and uncertainties, the predicted relationships were almost spot-on the widely published values (starting with Leopold and Maddock) for the hydraulic geom-etry relationships. Herein lies the intrigue, but I'm not convinced that the right answer was obtained for the right reasons. To his credit, the author acknowledges the fragility of some of the assumptions, and calls for additional work to better understand key relationships. Thus, this paper will stimulate discussion and work and deserves to be shared with the broader community, irrespective of any misgivings."

A: I appreciate that the reviewer sees value in the paper despite his skepticism.

Q: "The postulated relation between bankfull height and clay content builds on previous work, but as the author points out, the assumption in this earlier work was that cohesion limited erosion by increasing the required near-bank shear stresses for fluvial entrainment. Here the focus is on gravitationally-driven stress failures, modeled with a very simple 1-D relationship (Eq. 2), that when parametrized, gives rise to an even simpler linear equation between bankfull depth and maximum gravitationally stable bank height. This relation is tested at the scale of the Mississippi River basin by comparing bankfull height derived from USGS gage station data with predicted height calculated from clay content derived for each station using gNATSGO soil survey data and estimated with a 10km moving window. Not surprisingly, the resulting scatter plot is. . .well, scattered with a weak positive trend (Fig. 2A). The author does synthetically consider what the error in this relationship might look like; this error analysis is not propagated hrough the rest of the analysis, however. At the end of the day, I remain skeptical that broad-scale soil survey data can be used to parametrize clay content for

point data (USGS stations). To be fair, in the discussion the author does consider other sources of potential error beyond data error, including the use of stage data as a proxy for bank height, and other possible channel adjustments to varying clay content (i.e., varying bank angle, widening). But taken together, these uncertainties raise doubts about the validity of the paper's central claim that bankfull height is primarily controlled by gravi-tational slope failure."

A: Thank you for this comment. First, let me state that my discussion paper did not claim that gravitational failure is the dominant process occurring in channel widening and bank migration more generally. Fluvial scour and gravitational failure influence and complement one another to such an extent that it is difficult to clearly identify a dominant control for even a single alluvial channel, let alone for alluvial channels in general. Fluvial scour transports slumped material away from the bank, likely keeping bank angles higher than they would be without fluvial scour. Gravitational failure trans-ports material from higher on the bank to the toe (often following fluvial scour, which is maximized at the toe), likely keeping bank angles lower than they would be without gravitational failure. The relative importance of these two processes may be reflected in the bank angle, i.e., a bank angle persistently much less than vertical may indicate the dominance of gravitational failure while a bank angle persistently at or above 90ËŽ (i.e., a cantilever or overhang) may indicate the dominance of fluvial scour.

Given that both gravitational failure and fluvial scour play important and complemen-tary roles, which process do data suggest is the more dominant control on alluvial channel morphology? This question can be addressed by comparing the channel ge-ometries predicted by models based on gravitational failure and fluvial scour to data. Dunne and Jerolmack (2020) tested the fluvial scour hypothesis using the same allu-vial channel morphology dataset (i.e., Dunne and Jerolmack, 2018) used in my paper. The authors explain why their model underpredicts data by approximately 2 orders of magnitude in channels with D50 < 1 cm (Fig. 2A of their 2020 paper) this way: "For fine-grained rivers with D50 < 1 cm, however, we see rivers peel off of the Shields curve;

C3

the smaller the riverbed grain size, the larger bankfull shear stress deviates from the threshold expectation. . . we infer that this departure. . . represents the point where ïĄ ť c of cohesive banks, which is rarely measured, becomes larger than ïĄ ť c of noncohe-sive bed sediments, on average." The predictions of a model based on gravitational failure, while scattered, are consistent with data for all grain sizes. The only systematic deviation of a model based on gravitational failure occurs for shallow channels (< 2 m deep) with very low clay contents, a result that, as noted in the discussion paper, may be a result of an upward bias in clay content estimates at very low clay contents. I believe such a bias exists because low clay contents can only be constrained within the range from 0 to 10% clay, the average value of which (5%) is often assigned to the soil horizon being measured. If the actual value is very close to 0 (e.g., < 1%), the clay content estimate will be overestimated by much more than a relatively large clay content will be underestimated (e.g., 5% clay is 400% greater than 1% clay but only 50% lower than 10% clay). The averaging of clay contents from different horizons within the same soil and in different soils within a floodplain adjacent to a channel may partly mitigate this problem, but it is difficult to estimate low clay contents in the field in any case. I excluded channels with clay contents close to zero from my analysis due to this potential bias (line 205).

I did not propagate uncertainties more comprehensively in my discussion paper be-cause there are several sources of uncertainty, not all of which are straightforward to quantify: uncertainty in measured clay content, uncertainty in the relationship between cohesion and clay content, and uncertainty in the bank angle and moisture content that determine the stability parameter Ns. For example, as lines 257-259 noted, variations in bank angle can result in Ns values that differ by a factor of 10, so working to find addi-tional constraints on bank angles and how they vary throughout the MRB might quickly make obsolete the estimation and propagation of uncertainties in the current analysis, which implicitly assumes uniform Ns values. That said, I agree with the reviewer that a more comprehensive approach to uncertainty propagation would strengthen the paper. I am still working on how best to quantify and propagate uncertainties in the revised

C4

manuscript, but the steps I am attempting include: 1) estimating and propagating the uncertainty in clay content within gNATSGO to be approximately 10% (absolute value) based on similar datasets (e.g., Ramcharan et al., 2018), 2) quantifying and propagating the uncertainty in the least-squares linear regression between cohesion and clay content using the data from Dafalla (2013), and 3) compiling data on bank angles from the MRB (where available) to estimate uncertainty in bank angle and the related uncertainty of Ns values.

I appreciate the reviewer's skepticism that broad-scale soil survey data can be used to parameterize clay content for point data. However, I note that soil survey data are derived from point-scale measurements (SCS scientists use data from soil pits to define soil properties across areas of similar slope, geology, etc.). Moreover, rates of bank retreat at a gaging station are likely controlled not by the clay content of a single point but rather by clay content over the spatial scale of a meander bend, which for large rivers can be ∼1-10 km.

Q: "While I'm appreciative of the author's invoking the critical flow hypothesis (correct citation is Grant, 1997, not 2000 as it appears in text but not references), I'm somewhat confused by the role it plays in this story. As I understand it, the author argues that this hypothesis suggests that the range of Froude numbers for both sand and gravel-bed channels should be limited to near- or less than 1.0, and that therefore Froude number and discharge should be weakly and inversely related. The logic here is not entirely clear, and the mechanism that restricts Froude numbers is not entirely accurate. In its original form, the hypothesis argues that in steep channels (typically S>0.01), interactions between the free surface (particularly hydraulic jumps) and the bed result in a rough balance between forces that accelerate the flow and forces that extract energy from the flow and thus retard it, thus promoting near-critical flow conditions. This condition applies irrespective of grain size, although the author rightly points out that the actual bed features that set up this interaction are different for sand, gravel, and even boulder bed channels. A recent paper tests this idea in the flume and articulates the

mechanism well (Piton and Recking, 2019). Most of the streams in the Dunne and Jerolmack data set have slopes much less than 0.01, and consequently much lower Fr, as the data shown in the paper point out."

A: I thank the reviewer for pointing out the incorrect citation, which has been corrected in the revised manuscript. The reviewer is correct that I generalized his 1997 critical flow hypothesis. The wording has been corrected in the revised paper. My reason for generalizing his hypothesis was that, even though the critical flow hypothesis was originally proposed for steep channels, the key idea (to my mind) of the critical flow hypothesis is the existence of a self-regulatory feedback in which an increase in velocity is met with an increase in drag that brings the velocity, and hence Fr, back down to a lower value. Similarly, a decrease in velocity tends to be met with a decrease in drag that tends to increase the velocity. It is likely that such a self-regulatory feedback is not limited to steep channels. In steep, gravel-bedded channels, faster flow tends to lead to more wave breaking and other drag-inducing processes, thereby tending to lower velocities back down to a critical value (i.e., Fr = 1). In less-steep, sand-bedded channels, faster flow tends to lead to the development of larger and more well-developed bedforms, which increases relative roughness and hence drag in the range of Fr values associated with well-developed bedforms (Fr ∼ 0.1-1). The mechanisms for self-regulation differ between steep channels and less-steep channels, but self-regulation exists in both. In the revised manuscript, I clarify that the critical flow hypothesis applies to steep channels only and separately discuss evidence for a similar self-regulation of Fr in less-steep channels.

Q: "More to the point, in my view, the primary control on Fr is channel slope, and I suspect that this is what is behind the weak inverse correlation between Fr and discharge (Fig. 3A). The simple theoretical dependency between Fr and S is shown in Grant (1997; ig. 4); a more sophisticated treatment is given in Palucis and Lamb (2017; Fig. 2). Since smaller streams tend to be steeper than larger ones, my hunch is that discharge is more of a proxy for slope than a driver as in Fig. 3A. This slope dependency

is also probably lurking behind the grain-size/Fr relationship in Fig. 3B as well. None of this fundamentally invalidates the argument being made by the author but I think these relations should be acknowledged, as they have bearing on the physical mechanisms underlying the presentation."

A: Thank you for this comment. My work demonstrates how alluvial channel geometric variables interrelate. I used bankfull discharge as an independent variable because it has traditionally been used for this purpose. I did not and would not argue that any variable is more fundamental, significant, or a "driver" than any other variable. Bankfull Froude number depends explicitly on channel depth and velocity, the latter of which depends on slope and bed roughness. As such, bankfull Froude number certainly depends on channel depth, channel slope, and bed roughness (the latter of which depends on grain size and all of the factors that control bedform geometry (if bedforms are present)). I know of no basis for ranking the relative importance of any of these controls. This view is consistent with Palucis and Lamb (2017), who, despite relating specific channel morphologic types to specific slopes, document the importance of channel width and grain size in controlling channel type (e.g., "stepâĂŘpools form in near supercritical flow or when channel width is narrow compared to bed grain size") and conclude that "certain bed slopes have unique channel morphologies because the process variables covary systematically with bed slope."

Reviewer 2 (Roberto Fernandez):

Q: "1. The use of the data by Dafalla (2013) seems misrepresented in this manuscript (Figure S1). Cohesion values reported by Dafalla (2013) correspond to pure sand, sand and clay mixtures with 5 In addition, Dafalla (2013) shows that for the same clay content (15 I really believe that the approach proposed in this manuscript has a lot of potential and others have included geotechnical considerations in models for stream restoration (e.g. CONCEPTS, see Langendoen et al. 2001; RVR Meander, see Motta et al. 2012). I would encourage the author to dig out some more references regarding cohesion estimates for soils that are more relevant for the Mississippi River

Basin. For example, Masada (2009) presents a very extensive report on geotechnical parameters for the state of Ohio and includes different relations between sediment properties (cohesion for example) and soil composition (amount of silt, amount of clay, etc.). Their approach is specific to highway embankments but the results of their tests might be more general in terms of cohesion values in relation to clay contents."

A: I greatly appreciate the reviewer's drawing my attention to these additional sources of data, Masada (2009) in particular. I have scoured the literature for additional sources of data but have not found any. Masada (2009) is not a good source of data on the relationship between cohesion and clay content because he presents cohesion versus clay content for a very narrow range of clay contents (e.g., 27%-34% in Fig. E6). Also, I prefer to use data from direct shear tests rather than the triaxial shear tests of Masada (2009) because direct shear tests are more similar to gravitational shear failures. Dafalla (2013) is also a good source of data because he presents data for cohesion versus clay content at fixed moisture contents. Assuming that there is no systematic variation between short and tall banks in the water content at which channel banks tend to fail, it is the dependence of cohesion on clay content at similar moisture contents (or averaged across all moisture contents) that is most appropriate for use in the model.

The reviewer is correct that the sand and clay mixtures considered by Dafalla (2013) were obtained by mixing sand and clay as opposed to being natural soils. However, I don't think creating such mixtures necessarily undermines the quality of the data or the applicability of the results to natural soils. Masada (2009) considered natural soils but they were almost entirely glacial in origin, hence not necessarily a better analog for alluvial soils than the mixtures studied by Dafalla (2013).

Q: "2. The use of equation (2) might not be appropriate for riverbanks. The use of that equation as presented by Chen (1969) and Terzaghi et al. (1996; p. 271-272) is for soil embankments located above the water table. Several authors have used it in the past as discussed by ASCE (1988) but even there, the authors suggest that critical depth

approaches are not accurate when the most common bank failure mechanisms for riverbanks are due to tension cracks that cause toppling or cantilever failures. Assuming the equation is indeed an appropriate approach for riverbanks, I would en-courage the author to explore the sensitivity of its input variables to other values. Chen (1969) shows a wide range of Ns values that depend on the internal friction angle of the material (which is sensitive to moisture content) and the actual slope of the bank. The smallest stable bank height would be given by the smallest possible safety param-eter Ns so why not explore a range of Ns values. When the channel has low flow, the bank might be quite dry and its maximum stable height would be quite different from that obtained with a saturated bank (e.g. during the falling limb of a hydrograph where the river stage is getting lower but the bank remains saturated). It would be very useful to see these considerations in the analysis. The author discusses the issue briefly but more details regarding bank failure mechanisms and their prevalence might strengthen the manuscript."

A: Thank you for this comment. The results of eqn. (2) are consistent with more complex models if Ns values are defined in such a way as to incorporate the effects of moisture content. A linear relation between the maximum stable bank height and cohesion is robust across multiple bank-stability models with a variety of water table or moisture content scenarios. For example, Eaton (2006) derived a linear relationship between maximum stable bank heights and cohesion for fully saturated banks. Simon and Thomas (2002) demonstrated how the maximum stable bank height decreases with both an increasing bank angle and an increasing relative height of the water table. These results are consistent with Chen et al. (1969) if Ns values incorporate the effects of moisture content and/or water table height.

In the revised manuscript I will provide a sensitivity analysis that demonstrates the effects of variability and uncertainty in Ns values on the model predictions, as the reviewer suggests. I agree with the reviewer that channel banks may be quite stable when dry. However, over long time scales banks are saturated or near-saturated many

times, hence it is the shear strength during these times of relative weakness that are most relevant to determining the long-term stability of a bank.

Q: "3. Sensitivity analysis: Figures 2b and 2c present results for bank heights based on a synthetic dataset. If the author estimated clay contents using averaging windows for a soils dataset, why not extract second order statistics from it and use them directly instead of creating a synthetic dataset?"

A: Thank you for this comment. However, I don't think that second-order statistics (e.g., coefficient of variation) necessarily capture all of the potential uncertainty in a dataset. For example, spatial variations in the values of a dataset that systematically over- or under-predict actual values won't correctly capture the true uncertainty. I also think creating a synthetic dataset has advantages. For example, a key goal was to demonstrate that an upward bias can exist in very low clay contents. I don't think using second-order statistics in the averaging would necessarily have the same potential to demonstrate this phenomenon.

Q: "4. Use of the Mississippi River Basin data: The author clearly states why the MRB data are used. However, not knowing much about the many different locations along the basin, I have a few questions. (1) What percentage of the cross sections analyzed can be considered natural? (2) Did the author discard those locations where the navigable channels are maintained by the US Army Corps of Engineers? (3) Of the many stations used, how many might be influenced by river control structures (dams, wing dams, chevrons, etc.) or road infrastructure (e.g. culverts, bridges)?"

A: Thank you for this comment. Using Google Earth, I have examined the locations of the 354 stations and have found no instances where cross-sections are located close to infrastructure. This lack of overlap between the U.S.G.S. station locations and infrastructure may be partly due to the fact that my filtering criteria (lines 88-92) did an effective job at removing stations with stage-discharge relationships that are affected by infrastructure.

Q: "5. Figures 3 and 4: It is not at all clear why the author includes regression plots of the Dunne and Jerolmack (2018) dataset. Based on the abstract and introduction, it was unexpected that a different dataset appears in the manuscript and becomes the focus of the second half. I understand the use of the dataset for Figure 5, which is new but the content of Figures 3 and 4, is not. I would encourage the author to make it clear to the reader earlier that the DJ dataset is a substantial part of the analysis and to state explicitly the novelty of including figures 3 and 4."

A: Thank you for this comment. The opening sentence of the abstract makes clear that the problem I am tackling is the power-law scaling of bankfull depths, widths, depth-averaged water velocities, and along-channel slopes to bankfull discharge in alluvial channels. The first part of my paper deals with control on channel depth only, so it stands to reason that there must be another part of the paper that extends the work on channel depth to other aspects of channel geometry using additional principles. Figures 3 and 4 (including the reporting of best-fit exponents) have not been published elsewhere and their inclusion is important for meeting the goals of the paper.

Q: "6. Figure 5: I have a few specific questions about the analysis leading to Fig. 5. (1) What is the number (and percentage) of cases that report ripples/dunes over the entire Ohata (2017) dataset? (2) For those reporting ripples/dunes, what is the number and percentage of measurements obtained in the laboratory and in the field? (3) For those in the field, how many are for large rivers? Cisneros et al. (2020) show that traditional dune scaling equations overestimate the size of dunes in large rivers and propose the following relation between dune height (H) and water depth (h) – H 0.056h - 0.12h. 4) Are the only sources of roughness in the DJ data the ripples/dunes or gravel size? What about bars, meandering, vegetation?"

A: Thank you for this comment. 1) 1574 (42%) of the 3790 data points in Ohata et al. (2019) have ripples or dunes (noted in revised manuscript), 2) of that 42%, 19% are in the field and the remaining 81% are in the laboratory (noted in revised manuscript), 3) 3.2% (123 out of 3791) of the data points are from rivers with h > 5 m. 4) I am assum-

ing that the dominant (not only) sources of roughness in the channels of the DJ data are ripples/dunes or gravel clasts. Long-wavelength topographic features such as bars and meanders are not likely to be dominant roughness/drag-inducing elements given that the presence/absence of the flow separation that tends to dominate drag depends sensitively on the maximum slope of bedforms and other obstacles to the flow, with slopes in excess of 0.2 m/m generally needed to trigger the flow separation (though surface curvature also plays an important role in addition to slope; see below). I concede that vegetation can be a dominant source of roughness on the beds of ephemeral channels, and some of the scatter in my analysis may be a result of vegetation-induced bed roughness.

Cisneros et al. (2020) demonstrates that the lee-side angle of many bedforms in large alluvial channels is lower that empirical equations predict and argues that such lower angles means that flow separation and hence drag is less significant than empirical models would suggest in large rivers. It is important to note, however, that flow separation depends sensitively on the surface curvature in the zone of the maximum adverse pressure gradient (Lamballais et al., 2010), not just on the relative height or lee-side angle of bedforms. As such, more research is needed to conclude that flow separation is rare on the lee sides of bedforms in large alluvial channels.

Q: "As a final general comment, I was hoping to see more analysis on the Mississippi River Basin dataset and comparisons between it and the DJ dataset where possible. The manuscript seems to be split between two separate analyses but the abstract and introduction do not suggest that. I recommend the author to modify these initial sections as necessary and compare the MRB data with the DJ data where possible. What kind of relation does the author obtain between bankfull depth and bankfull discharge for the MRB under the geotechnical considerations? On the other hand, could clay contents (and cohesion) be estimated with a revised version of equation (4) for other rivers in the world where soil data is not readily available?"

A: Thank you for this comment. See my response to point 5. The revised manuscript

will include the relation between bankfull depth and bankfull discharge for the MRB under the geotechnical considerations. It should be possible to infer clay content values for some rivers if bank angles and moisture contents were also well constrained. We do not have such data yet for many alluvial channel cross sections across a sizable region.

Q: "1) How do the bankfull estimates found here for the MRB compare to those of Dong et al (2019). This reference appears in the introduction but is not mentioned in the discussion. A: Since Dong et al. (2019) deal with the Selenga River Delta and my paper deals with the MRB, it is difficult to make a direct comparison with their results.

Q: "2. I did not understand the fourth criteria used to keep a USGS gaging station in the analysis of the MRB."

A: I apologize that this was not clearer. The estimation of bankfull stage requires fitting stage to discharge for the smallest and largest values. The bankfull stage is estimated to be where these two lines meet. In order to verify that the low-flow fit is reasonable, I retained only those stations for which the extrapolation of the fit of low-flow values passes "close" to the correct value: zero flow depth at zero discharge. In cases where the extrapolation of the fit does not pass close to zero, the data are likely of insufficient quality or require an adjustment based on data that are not publicly available. What represents "close" should not be based on an absolute error, e.g., 0.5 m, because such a criterion would require that low-flow fits for a deep channel be much more accurate than one for a shallow channel. So, instead, I defined "close" as within 50% of the bankfull stage from zero. That is, if the bankfull stage is 5 m, then the extrapolation of the low-flow fit to zero discharge must yield a stage within 2.5 m of zero. Similarly, if the bankfull stage is 2 m, then the extrapolation of the low-flow fit to zero discharge must be within 1 m of zero. An abbreviated version of this explanation is included in the revised manuscript.

Q: "3. If the analysis discards rivers with depths smaller than 2m, why is the 0.5m to

1.5m soil depth the only section considered for the analysis. What about river sections with different bank layers? The author mentions that soil data below 1.5m is not reliable but how valid is it to assume a uniform soil profile for the entire channel depth? How sensitive is the proposed model to this assumption?"

A: Thank you for this comment. I assume that the texture of the top 1.5 m of the bank (for which data is readily available) is representative of the entire bank. I understand that this assumption may be violated in many cases. However, given that the floodplain deposits that comprise many banks are the depositional products of channels where well-sorted sediments tend to be the norm, it is reasonable to expect that the texture of the uppermost 1.5 m of the bank will correlate strongly with bank material at greater depths in many cases.

Q: "4. Line 150(and other locations) - Ohata et al 2017 (not 2019)." A: That you for pointing out this typo. It has been fixed in the revision.

Q: "5. Line 151 "How did the author "cross-reference" the Ohata et al. dataset with the Dunne and Jerolmack dataset?

A: Thank you for this comment. Cross-referencing refers to the development of a curve in Fr vs. d50 space that separates channels that have ripples and/or dunes from those that do not, and using that curve to infer that the vast majority of sand-bedded channels in the Dunne and Jerolmack dataset have ripples and dunes. This is done by assuming the existence of ripples and dunes in channels of the D&J dataset that have Fr and d50 values that sit above and to the left (Fig. 3B) of the envelope curve separating channels with and without ripples and dunes in the Ohata et al. (2019) dataset. This has been clarified in the revision.

Q: "6. Line 214 "what is the equation of the curve (envelope) used to identify the conditions conducive to dune/ripple development?"

A: There is no equation. This is simply a drawn curve. Noted in revised manuscript.

Q: "7. Line 259 (and other locations) "Chen 1969 (not 1971)""

A: Typo fixed.

Q: 8. Line 286 "Vegetation. What are the predominant types of vegetation along the MRB? How deep are their roots? Root length might set slump block thickness. Vegetation might be the most relevant factor in shallow channels (up to max root length) and geotechnical considerations might be more relevant in deeper channels where roots might not stabilize the full bank.

A: Thank you for this comment. The stability of any bank is primarily a function of its weakest portion or layer, i.e., failure is more likely to occur in a zone of low shear strength compared to a zone of high shear strength, all else being equal. Failure of a weaker zone underlying a stronger zone can create a cantilever, but such cantilevers cannot continue to grow indefinitely, hence rates of long-term retreat in stronger and weaker zones of the same bank will tend to be similar. This, together with the fact that vegetation is likely to strengthen just the uppermost approximately 2 m of channel banks (globally, >99% of roots are found in the top 2 m of the soil (Jackson et al., 1996)), suggests that bank material shear strength is not likely to be controlled primarily by vegetation in channels deeper than 2 m. I concede that some plant roots can exceed (even greatly exceed) 2 m in depth. However, plant roots become quite small in density at > 2 m depth based on Jackson et al. (1996). I understand that this view appears to contradict the dependence of bank retreat rates on vegetation that has been documented in many studies (e.g., Ielpi and Lapôtre, 2020). I am not stating that vegetation has no effect on bank stability. Rather, I am stating that I know of no study that has attributed bank shear strength to vegetation that has accounted for the fact that wetter climates with more vegetation also tend to have more clay-rich soils. As such, much of the apparent control of vegetation could be due to clay content.

Reviewer 3 (Christopher Hackney):

Q: "I find the section between lines 30 and 34 quite confusing. The author states

that channel incision or floodplain deposition may increase bank height (OK so far). This causes banks to collapse once a critical bank height is reached. The subsequent failure results in channel widening which tends to increase water depths back towards the stable bank height. It is this last bit that I find counter to the previous few lines. I follow the argument that a wider channel results in slower flows, but the reduction in flows is a result of the increased channel capacity due to increasing width, and so there is no need for the channel to continue to incise of raise water depths to hold the same volume of water, as the increased width accounts for this. Furthermore, it has been shown recently that over long time frames, channels tend to maintain an equilibrium width (and thus presumably bank height; see Mason and Mohrig, 2019) and that the channel adjustment due to bank collapse also often sees increased deposition on the inner bank (see van de Lageweg et al. 2014). A clearer explaination of how bank failures can result in increased water depths would be welcome."

A: Thank you for this comment. The reviewer is quite correct that the sentence "Bank failure results in channel widening, which may reduce depth-averaged water velocities and therefore tend to increase water depths (to convey similar water discharges) back towards the maximum stable bank height" was confusing and did not accurately convey my conceptual model for how alluvial channels self-regulate to a maximum depth comparable to the maximum stable bank height. I have replaced it with: "Bank failure results in channel widening, which introduces new sediment into the channel bed. As this sediment is redistributed on the bed, the channel depth tends to decrease back towards its critical value."

Q: "The author also notes a few potential limitations to the data exploration in the discus-sion, but does so briefly and in passing. I would like to see a more developed discus-sion around the role of vegetation induced cohesion, and also the role of failed material (particularly in clay rich soils which are more likely to fail and persist in blocks at the base of the river bank), as these are likely to be key local controls on any variation in the relationships the author has presented. Another potential source of

variaiton that isn't raised but may also be important is the role of floodplain topography in defining bankful depths and bank heights. If a bank is eroding through a scroll-bar then it is likely that following a failure the local bank height may decrease as a result of variable local topography (i.e. on a floodplain where elevation is sloping away from the bank). Following the model presented here, will that new bank remain stable until its critical height either is reached through channel incision or build up from floodplain deposition? Following on from this, on Ln 272 the author states that an increase in bank height caused by floodplain deposition may trigger bank failure. Presumably, to deposit ma-terial on the bank the flow needs to be over-bank. Therefore, is it the deposition of the material on the bank during these flows, for the increased water velocity and bank shear stresses that will induce this erosion?"

A: Thank you for this comment. Discussion items are, by nature, relatively brief (in that they are limited to subsections of the Discussion section only). To the extent that my discussion items are relatively brief, I think this is appropriate given that there is no indication that vegetation-induced cohesion causes the scaling of alluvial channel geometry that is the subject of this paper. Also please see my response to reviewer 2's comment 8 on the issue of vegetation.

The revised manuscript mentions the role of failed bank material in the context of the interplay between fluvial scour and gravitational failure in contributing to channel width adjustment.

Regarding the final question, either incision or floodplain deposition can trigger bank failure: it is not one or the other. Whether a channel incises depends on changes in base level and sediment supply that depend on tectonic and climatic processes that operate at scales much larger than a channel reach. A channel that is relatively shallow relative to its bankfull depth will tend to flood overbank more often than a deeper channel, thereby promoting deposition until the channel deepens to the point where channel banks fail. The model of this paper invokes a long-term dynamic equilibrium between channel deepening processes (channel incision and/or floodplain deposition)

and channel shallowing processes (bank retreat, which introduces sediment into that channel that, once redistributed laterally, causes channel aggradation). My response to the reviewer's first concern clarifies this aspect of the conceptual model in the revised manuscript.

Q: "Overall, I think a lot of the issues raised above come down to the temporal scale being examoined here and there is a need for some discussion the manuscript around the ime-scales over which these processes may become the dominant factor and how time-averaging of the other processes involved in river bank erosion occurs. Could the author examine any historical rates of bank erosion for sites analyses in the manuscript to see whether the theoretical model holds for different time periods?"

A: I appreciate this suggestion. However, while rates of bank erosion are available over multiple time scales, what matters to this paper is channel widening. Channel widening likely occurs most abruptly during channel incision periods (following climatic changes, for example), after which bank retreat continues but channel widening does not be-cause cut bank migration is approximately balanced by point-bar progradation (e.g., Mason and Mohrig, 2019). I am not aware of any available data on channel widening data over multiple time scales. I also wish to emphasize, as I did in my responses to reviewer 1, that I am not arguing that gravitational failure is the dominant processes setting channel depth and width. Fluvial scour and gravitational failure influence and complement one another so intimately that it is difficult to clearly identify a dominant control.

References not cited in the discussion paper or reviewer comments:

Dunne, K.B.J., and D.J. Jerolmack (2020). What sets river width? Science Advances, 6(41), eabc1505, doi:10.1126/sciadv/abc1505

Eaton, B.C. (2006). Bank stability analysis for regime models of vegetated gravel bed rivers. Earth Surface Processes and Landforms. 31, 1438-1444.

Ielpi, A., and M. Lapotre (2020). A tenfold slowdown in river meander migration driven by plant life. Nature Geoscience, 13, 82-86.

Jackson, R.B., J. Canadell, J.R. Ehleringer, H.A., Mooney, O.E. Sala, and E.D. Schulze (1996). A global analysis of root distributions for terrestrial biomes. Oecologia, 108, 389-411.

Lamballais, E., J. Silvestrini, and S. Laizet (2010). Direct numerical simulation of flow separation behind a rounded leading edge: Study of curvature effects. International Journal of Heat and Fluid Flow, 31, 295-306.

Simon, A., and R.E. Thomas (2002). Processes and forms of an unstable alluvial stream with resistant, cohesive streambeds. Earth Surface Processes and Landforms 27, 699–718.