We thank Dr. Simon Mudd (Associate Editor) for his constructive review and for investing time and effort in thoroughly reviewing our manuscript. Below we address the comments in detail.

Comments to the author:

I have now read the revised version of this manuscript and the response to reviewers. I am happy with the response to reviewers: the authors have not exactly addressed all the comments because in some instances addressing these comments would result in a totally new paper. The authors have explained why they have made the choices that are present in this paper and those choices make sense to me.

Thank you for your understanding and support.

I attach an annotated pdf with a number of comments. Most are minor but there are a few that will require a bit more work. It would be nice to see if the analytical solution match those presented by Mitchell and Yanites (2019), which are just dimensional versions of the Royden and Perron (2013) solutions.

Thank you for the comment. We have revised the manuscript and answered the comments line by line. Here, we present how our derivations (equations 14 and 21) resemble equations 5c and 6c of Mitchell and Yanites (2019), respectively. (Comment on Line 140)

Our equation (14) gives an estimate on the response time of the knickpoint:

$$\tau(x_p) = \frac{k_{s\_1}(1-\gamma_{0\_1})}{k_{s\_1}^n(1-\gamma_{0\_1}^n)} \cdot \frac{1}{KA_0^{m/n}} \cdot \chi(x_p), \text{ with } \gamma_{0\_1} = k_{s\_0}/k_{s\_1}$$

Using the relation, $U_0 = K \cdot k_{s\_0}^n$ and $U_1 = K \cdot k_{s\_1}^n$, we can get:

$$\tau(x_p) = \frac{k_{s\_1} - k_{s\_1}}{k_{s\_1}^n - k_{s\_0}^n} \cdot \frac{1}{KA_0^{m/n}} \cdot \chi(x_p) = \frac{(U_1/K)^{1/n} - (U_0/K)^{1/n}}{U_1 - U_0} \cdot \frac{1}{A_0^{m/n}} \cdot \chi(x_p)$$

Rearranging the equation, we can derive:

$$\chi(x_p) = \frac{\tau(x_p) \cdot (U_1 - U_0)}{(U_1/K)^{1/n} - (U_0/K)^{1/n}} A_0^{m/n}$$

Which is identical to equation 5c of Mitchell and Yanites (2019).

Adopting the assumption of Mitchell and Yanites (2019) of a single step increase in the uplift rate from U_0 to U_1, equation (21) of our study can be simplified to be:

$$z\left(t, x_p(t)\right) = U_1 t + \left[\frac{(1-\gamma_{0\_1}^n)}{(1-\gamma_{0\_1})} - 1\right] \cdot U_1 \cdot t = \frac{1-\gamma_{0\_1}^n}{1-\gamma_{0\_1}} \cdot U_1 \cdot t$$

where $\gamma_{0\_1} = k_{s\_0}/k_{s\_1} = (U_0/K)^{1/n}/(U_1/K)^{1/n} = (U_0/U_1)^{1/n}$. Thus, we can get:

$$z\left(t, x_p(t)\right) = \frac{1-U_0/U_1}{1-(U_0/U_1)^{1/n}} \cdot U_1 \cdot t = t \cdot (U_1 - U_0) \cdot \frac{{U_1}^{1/n}}{{U_1}^{1/n} - {U_0}^{1/n}}$$

This equation is similar to equation 6c of Mitchell and Yanites (2019). Importantly, our equation (21) is more general, as it does not assume only a single step change in U.

Following this comment, we added to the manuscript proper referencing to Mitchell and Yanites (2019) in lines 239-241 and Text S3.

Notably, Mitchell and Yanites (2019) showed a different form (equations 5a and 6a) for $n < 1$, which is the coordinate (χ and z) of the intersection point between the channel segment that is under equilibrium with previous uplift rate U_0 and the stretch zone that is caused by increasing U and $n < 1$.

What I do not have in the annotations is this issue:

1) **Only a tectonic history in which knickpoints have not merged can be reconstructed. There are then a wide range of "hidden" histories, not inverted, that are still possible due to knickpoint mergers. But in fact those histories are not completely unknown. The modern channel profiles can still be used to eliminate scenarios that contain a merged knickpoint. This information appears as though it could be extracted from some simple analyses such as figure 2. So, for example, if there were 1 consumed/erased segment, what is the limit of uplift? If there were two consumed segments? And so on.**
2) **I think the authors should at least have something to say about further constraints on the hidden histories. This is a big issue for me. Many papers have now been published applying n = 1 inversions all over the world. Should we believe these at all? How much of the record could be missing? Addressing that question might be the topic of a subsequent paper but at the bare minimum the authors should outline a strategy, since I think that would constitute a major contribution to the field.**

   Thank you for these comments. When slope exponent $n > 1$, step-increases in tectonic uplift rates have the potential to fully erase channel sections. This means that portions of the tectonic uplift history have been lost. In the revised version, Lines (473-480), we present a simple analysis that places restriction on the hidden lost history. We consider a case of one fully consumed channel segment (i.e., merging two knickpoints) and using the relation, $\chi(kp_2) > \chi(kp_1)$, we derive the relation between U_1, T, and T_1 that all possible lost histories should obey (equation 30).

   More importantly, we stress that a Bayesian approach (mentioned around Line 485) or other non-linear approaches that rely on the forward model we developed could be used to find many histories consistent with the observed river profile after full sections have been consumed and with other independent observations (such as dated uplift rate or incision rate markers).

Comments line by line:

1. **Line 13: "Tectonic rates" is quite ambiguous. This could be lateral motion along a strike slip fault. You are really just referring to the uplift rate. All the models you use specifically simulate vertical uplift and erosion. So I think here and throughout you can use more precise terms.**

   Thank you for this comment. We have revised it as "tectonic uplift rates" (Line 13 and throughout the manuscript).

2. **Line 27: I would say "which can eventually lead to"**

   Thank you for this comment. We have revised it (Line 27).

3. **Line 28: see comment about "tectonic rates" in the abstract. You mean here a change in uplift rate, do you not?**

   Yes, we have revised it (Line 28-29).

4. **Line 37: I would say "is widely used to".**

   Thank you for this comment. We have revised it (Line 37).

5. **Line 76: Does this not depend on knowing erodibility? One can get a steadily uplifting landscape with different segments just by having different underlying rocks. So a full**

constraint is only possible if the values of K are known. Is that not the case?

Yes, a full tectonic uplift rate history can be retrieved only under a well-constrained erodibility *K*. We have revised it (Line 76).

6. **Line 95: Morisawa was the first to identify the power relationship and should be cited here. In addition, the Hack, Flint, and Morisawa papers all identified equation (3) through data, not through any incision model. I note this because this section is written as though equations 3, 4 and 5 are based on SPIM but in fact if you substitute k_s into equation 4 these are simply statements of observed topography that do not depend on a model. I think it is useful to state this fact.**

   Thank you for this comment. We have added this important point and the reference Morisawa (1962) in Line 99-100.

7. **Line 117: I would add an earlier paper here since this term was defined before 2010.**

   Thank you for this comment. We have added Wobus et al., 2006 (Line 120).

8. **Line 125: We are all using pdfs now so there is no reason to put the figures at the end of the file. This only made sense when we had paper copies and kept the figures separate. These days it just makes it difficult to scroll between the text and the figures. In the revision could you please embed the figures in the text? Thanks.**

   Thank you for this comment. We have embed the figures in the text.

9. **Line 140: There are a few other equations found in the literature for these terms. The most notable is Royden and Perron (2013). I'm surprised that is not cited in this section. Those solutions are not immediately equivalent to these because they are nondimensionalised, but Mitchell and Yanites (2019) reported dimensional solutions (reported in chi space). I have not gone through the graft of differentiating equations 5c and 6c from that paper and converting from chi space, but I think it would be a useful check to see if your equations agree.**

   Thank you for this comment. Mitchell and Yanites (2019) reported dimensional solutions under the condition of one single knickpoint, e.g. the equations 5c and 6c in their paper. We compared their solutions with our derivations and found a consistence. See details in Line 239-241 and Text S3.

10. **Line 151: "do not merge" is simpler.**

    Thank you for this comment. We have revised it in Line 163.

11. **Line 158: Royden and Perron 2013 should be in this list.**

    Thank you for this comment. We have added the reference in Line 170.

12. **Line 176: It might be useful here or a bit earlier to add some sentences clarifying that when n>1, a steeper segment can consume a gentler segment, but additionally the knickpoints can move at different rates. The reason why "merging" and "consuming" are used is because these are two different phenomena. It says this in the paper but this is an important point and should be highlighted so there is no chance a reader misses this distinction.**

    Thank you for this comment. We have added it in Line 179-181. (“consumption” is reserved for channel segments that are shortened by a fast-migrating knickpoint, and “merging” is reserved for knickpoints to highlight the different dynamics of the merged knickpoint from the two knickpoints that joined to form it.)

13. **Line 178: "to lower values of..."**
    Thank you for this comment. We have revised it in Line 192.

14. **Line 194: I might use a different term than stretch because Royden and Perron used a term "stretch zone" that referred to something completely different. Maybe just "the channel reach between the two knickpoints" because in this sentence "stretch" doesn't add any meaning.**
    Thank you for this comment. We have revised it in Line 220.

15. **Line 198: Say "any evidence that knickpoints have merged"**
    Thank you for this comment. We have revised it in Line 224.

16. **Line 216: Awkward phrase. I'm not quite sure what you mean. Rewrite.**
    Thank you for this comment. We have revised it in Line 244. (generate a piecewise solution for knickpoint elevation before and after knickpoints merging)

17. **Line 229: "of multiple knickpoint merging events"**
    Thank you for this comment. We have revised it in Line 257.

18. **Line 246: "inference of"**
    Thank you for this comment. We have revised it in Line 285.

19. **Line 246: Also assumption of homogenous K, no?**
    Thank you for this comment. We have added "the block has a uniform erodibility" in Line 289.

20. **Line 257: subject to the (somewhat restrictive) assumptions (no merged knickpoint, staircase uplift history)**
    Thank you for this comment. We have revised it as "Consequently, a full uplift rate history, subject to the assumptions of no merged knickpoints and a staircase uplift change, can be derived." in Line 297-298.

21. **Line 284: The Akaike information criterion was designed for exactly this kind of problem (i.e. penalizing overfitting). Why wasn't it used?**
    Thank you for this comment. The Akaike information criterion is surely good for penalizing overfitting. In this study, we mainly focused on both the forward and inverse models, i.e. knickpoint migration, preservation, and merging, and retrieving uplift history. Thus, we plan to prefer leave the application of this method to the future studies. (Line 323-324)

22. **Line 330: Important to state that these are all draining to the same base level (so that it is clear the minimisation of disorder is applicable for determining m/n).**
    Thank you for this comment. We have stated it in Line 390.

23. **Line 332: with different random positions.**
    Thank you for this comment. We have revised it in Line 393.

24. **Line 339: based on what?**
    Thank you for this comment. We have revised it as "based on the best correlation coefficient" in Line 400.

25. **Line 345: In particular,**
    Thank you for this comment. We have revised it in Line 405.

26. **Line 360: A bit clunky. I would say "the analysis of merging knickpoints"**
    Thank you for this comment. We have revised it in Line 421.

27. **Line 362-364: I wonder if this could be stated more clearly. It is very important. If a tectonic history has occurred without merging of knickpoints, our method can reconstruct this history. However, there are many tectonic histories that result in knickpoints merging that cannot be recovered using inversion.**
    Thank you for this comment. We have revised it "The inverse inference, however, has a different property, whereby any particular river long profile could be generated by many tectonic uplift histories (as demonstrated by the evolution depicted in figure 4). If a tectonic uplift history has occurred without merging of knickpoints, our method can reconstruct this history. However, a tectonic history that results in knickpoints merging cannot be recovered using our linear inversion method. More specifically, when our inverse approach is applied to a river long profile, the outcome will be the one history for which all knickpoints are preserved, although this inferred outcome might not be the real history that shaped the profile." (Line 423-429).

28. **Line 376: " (but no certainly)" typo**
    Thank you for this comment. We have erased it.

29. **Line 414: gained**
    Thank you for this comment. We have revised it (Line 486).

30. **Line 726: It would be useful to have another figure that shows the chi profiles of the best fits, to show how they compare with the real chi profiles. I'm slightly surprised, based on the chi profiles, that 2 division points do better than 1, and it would be interesting to see where the knickpoints line up on the best fit inversion.**
    Thank you for this comment. The chi profile of the best fits is in Figure 6b (the gray, thick line). We have added the grey arrows to indicate the position of the knickpoints on the modelled χ-z profile.