We thank reviewer 2 (Wolfgang Schwanghart) for his thoughtful review of our manuscript. Below we list reviewer comments in italic font and our responses in regular font.

Gailleton et al. present a method that automatically extracts knickpoints from longitudinal river profiles. The algorithms developed by the authors are well described and are implemented in LSD TopoTools, a terrain analysis software written and maintained by the authors. The algorithms are tested against hand-picked knickpoints and those derived with other software, and the code is publicly available. Overall, the manuscript is very well written and nicely illustrated. I have no concern about this paper being appropriate for the journal ESURF. To this end, I only have a few questions and some specific comments.

Thanks. We are glad to hear the manuscript is clear.

Would it make a difference, if you first smooth the elevation values using the TVD-approach and then calculate ksn? The smoothness-parameter would then be independent of theta.

Thanks for this suggestion. We were not keen to do this in the first version of the manuscript as we didn't want to smooth elevation before searching for differences in χ gradient, since it would add an extra layer of complexity to the method. However, we have now attempted the smoothing of elevation using some different techniques to test if it makes a difference to the results.

Firstly, the TVD algorithm cannot be applied on the raw profile since it is designed to flatten signals and cannot be used on monotonically increasing data. Our approach has been to apply the TVD on a detrended elevation for each tributary (*i.e.,* applying the filter on Δ elevation rather than elevation itself. As shown on Figure 1, different values of λ will generate different level of smoothing by flattening Δz with different intensity. The denoised Δz is then applied from the base level to the channel head to produce the denoised profiles. However, the denoising still depends on the λ coefficient as the intensity of denoising might depend on the DEM quality and the user need (*e.g.,* focusing on large-scale gradient changes or small scale). Although the θ dependency cannot be avoided, we have added an elevation denoising option in the algorithm and also a description of this in the manuscript. In addition, we have added figures to the supplementary materials showing the performance of elevation denoising. However, we suggest to be cautious with adding denoising to the method as it involves data loss. We suggest reading relevant literature (e.g., [1]) that discusses this specific issue before considering filtering initial dataset. The following figure illustrates the denoising results: a) shows the effect of lambda on the denoising intensity and b) the resulting k_{sn} for denoising with a voluntary high regulation parameter $\lambda = 25$ to show that another denoising still is required.

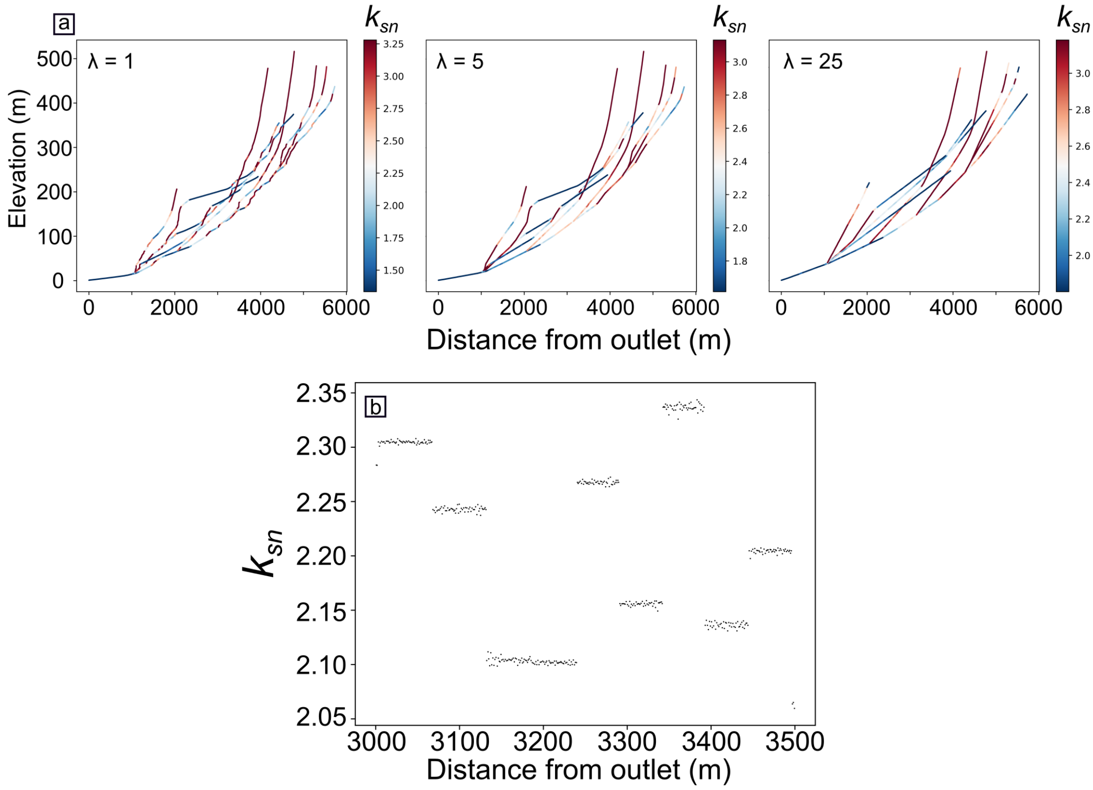

Figure 1: Effect of applying a TVD denoising filter on the elevation prior to the rest of the method as described in the manuscript. a) Long profile representation for values of λ. b)k_{sn} calculated with [2] from a higly denoised profile ($\lambda = 25$). Even significantly denoised, it still requires a run of TVD to clean the signal. The noise magnitude is still dependent on θ in the same way described in the manuscript.

Detecting knickpoints by identifying gradient-changes of ksn could also be achieved by calculating the profile curvature of the elevation data in chi-space. Similarly to M_χ, this could be C_χ (or something similar). Of course, mathematically, this is the same. In addition, curvature is strongly affected by noise in the river long profile. However, using curvature instead of gradients of gradients is slightly more elegant and smoothing curvature might directly yield the peaks and troughs that you are looking for.

Again, thanks for the suggestion. We tested several scenarios of curvature fitting to see if it improved our method. Tests suggested some potential for using curvature to detect and quantify knickpoints, however there were several serious limitations. We first experimented $C_\chi = d(k_{sn})/d(\chi)$, where k_{sn} is calculated with the segmentation algorithm [2] and filtered with the TVD. However as the discrete changes in χ between each node are quite variable, the resulting profile is significantly noisier than using k_{sn} directly. The magnitude of each of the knickpoints detected with curvature becomes more sensitive to χ spacing and therefore θ compared to the method using k_{sn}. This is illustrated in Figure 2a where some of the C_χ differ from Δk_{sn}. θ in our case is relatively low, and therefore the discrete changes in χ happen to be in the same order of magnitude as the discrete changes in elevation. However, for higher values of θ, $\Delta\chi$ can be several orders of magnitude lower than Δz and therefore generate unnecessary high values. A similar issue is discussed in section 5.2) in the manuscript. Moreover, we find the Δk_{sn} quantity more appealing as it can directly be translated into a drop/increase of channel slope (normalised to the concavity). We then explored the possibility of using a direct calculation of χ- elevation profiles to detect knickpoints (*ie.,* $d^2z/d\chi^2$). We applied a moving-average window on the C_χ and on $|C_\chi|$ to smooth and isolate peaks in curvature as suggested in the review. This method fails at identifying single outliers. Figure 2 shows the results of the three methods tested using curvature-based methods.

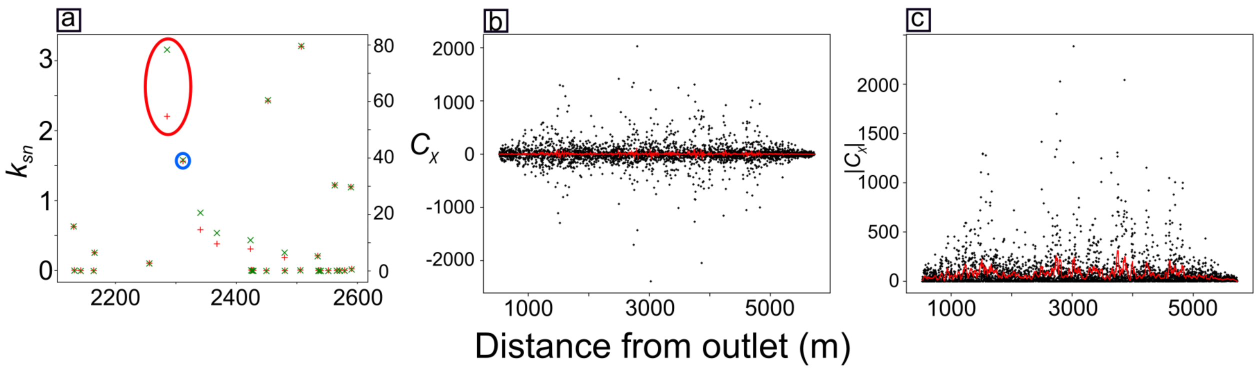

Figure 2: Methods based on chi curvature (C_χ). a) The C_χ values (green) of a river reach derived from $C_\chi = d(k_{sn})/d(\chi)$ and compared with Δk_{sn} values (in red) used in the manuscript to identify slope-break knickpoints. Their magnitude is similar but is very sensitive to $d\chi$, which is a function of θ. The two circles show cases where C_χ and Δk_{sn} show different and similar behaviour from the same original data. b) $C_\chi = d(k_{sn})/d(\chi)$ across the main river in the Smugglers Catchment. The red line represents a moving-average window of 20 nodes across the signal. The high and low peak values suggest some show potential to isolate knickpoints, but would require strong signal processing to be isolated: this method does not do better than our current method. c) Absolute value version of method b).

Detecting change points in noisy data is a common topic in signal processing and statistics (see e.g. Truong et al., 2018). I wonder whether some of the techniques of knickpoint identification could actually be applied in a more formal statistical framework.

Thanks for pointing out this reference; we have included it in the text. Alongside with this addition, we are adapting the vocabulary describing the method to fit with the statistical framework. Moreover, the reference offers (i) a review of the different statistical method to detect point changes and (ii) a python implementation of the main algorithm "*rupture*". The TVD suits our needs, but one might want a different method to adapt to a specific case study that would fall in the limitation of our method. We therefore adapted the code to generate raw files containing the output of the algorithm at different stages of our method. User can now, if needed, fit another method to ours using for example the "*rupture*" package.

Specific comments

6, 25: Filling might cause problems, because it can generate some large steps. Carving might be a better alternative.

Yes, thanks for raising this. Our test examples were in locations where there were few roads and bridges. These features can generate steps after filling. The sites also had relatively little topographic noise and so we did not find valley filling to be a problem in our analysis. However, we recognise that many DEMs will have steps introduced by the filling algorithm. We therefore added a depression breaching algorithm in our software suite [3], as well as an option to directly feed the algorithm with a preprocessed raster from an alternative source (e.g., TopoToolbox, RichDEM). We have also added reference to the carving algorithm in the text.

8, 12: How much does "combining knickpoints" (2.3.2) actually affect the objective to identify the precise location of transitions between segments? It seems to me that knickpoint merging will let you pick knickzones, rather than knickpoints.

We addressed this point by running a sensitivity analysis, which was available in the supplementary materials of the discussion materials. The segments are made of a large number of nodes and results show that, except for a large combining window (i.e., >100 nodes), the combining algorithm only cleans composite transitions between segments and does not combine large knickzones. We agree with the point that in the specific case of a close succession of knickpoints (e.g., a succession of waterfalls) and if the DEM precision is high enough to show them, then the algorithm might combine this succession of knickpoints as a single entity. We add this point to the main manuscript.

Eq. 7: Denoising: The TVD algorithm (Eq. 7) is similar to the smoothing approach by [1], with the difference being the applied smoothness penalty. It would be interesting to know why you chose a gradient penalty instead of a curvature penalty. Wouldn't the gradient penalty require the horizontal distance in the denominator as the node-to-node distance may change depending on whether the node is a cardinal or diagonal neighbor?

We developed a statistical approach in [2] to identify the best fit segments in chi-elevation space and the gradient is calculated on the basis of these segments: the spacing of the nodes is taken into account withing the segmentation routine. The TVD is then applied to the segmented data in order to minimise small variations in the already calculated k_{sn} enabling extraction of knickpoint locations. We have added a reference to the study suggested above and described how it is different from our approach in the revised manuscript.

12, 20: I was wondering about this error radius when reading through section 2.4. Consider to mention the radius also there. Did you use the same radius in the Brazilian test case?

The radius has been chosen from [4] based on their published parameters. We applied the same radius in the Brazilian case study (made with new field-derived dataset of knickpoint) for consistency. This has been clarified in the manuscript.

References

[1] W. Schwanghart and D. Scherler, "Bumps in river profiles: Uncertainty assessment and smoothing using quantile regression techniques," *Earth Surface Dynamics*, vol. 5, no. 4, pp. 821–839, 2017.

[2] S. M. Mudd, M. Attal, D. T. Milodowski, S. W. Grieve, and D. A. Valters, "A statistical framework to quantify spatial variation in channel gradients using the integral method of channel profile analysis," *Journal of Geophysical Research: Earth Surface*, vol. 119, pp. 138–152, feb 2014.

[3] J. B. Lindsay, "Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models," *Hydrological Processes*, vol. 30, pp. 846–857, mar 2016.

[4] A. B. Neely, B. Bookhagen, and D. W. Burbank, "An automated knickzone selection algorithm (KZ-Picker) to analyze transient landscapes: Calibration and validation," jun 2017.