Reply to comments and queries of Anonymous Referee #1 for the manuscript "Deriving principle channel metrics from bank and long-profile geometry with the R-package cmgo"

Original manuscript title: Deriving principle channel metrics from bank and long-profile geometry
                                with the R-package cmgo
Date of first submission: May 10 2017
Manuscript No: esurf-2017-32
Date of first review: June 04 2017

Please note:

- This document contains the comments of the first anonymous reviewer in green text color and our replies in black text color.
- Line numbers of the Reviewers' comments refer to the original manuscript, while line numbers in our replies refer to the updated manuscript. They are consistently given in brackets, e.g. [76].
- When figures are referenced in our replies they refer to this reply letter if written without brackets and to the manuscript if written with brackets, e.g. [Figure 2]

Reply to comments of Anonymous Referee #1

We thank anonymous Referee #1 for the supportive and very constructive comments on our manuscript, which we used to improve our manuscript. Below, we answer the comments one by one in detail.

On a general note, as this was a major objection by the referee, we would like to elaborate on the type of this manuscript. We submitted this manuscript as a **technical contribution**. Technical articles are a new class of manuscripts at ESurf, and their introduction has been recently decided at the Editorial Board Meeting at the 2017 EGU General Assembly (pers. comm.). The type of this article has been communicated to the Editor in the letter to the Editor but apparently, there has been a poor communication about the type of this article to the referees. Please consider this general fact for the reviews on our manuscript.

- The manuscript submitted by Golly and Turowski presents a very useful tool to extract river metrics, such as channel width. This parameter is required for many studies, as pointed by the authors. Therefore, the topic is of interest and might fit with the journal scopes.

Thank you for this assessment of our contribution.

- There are already other tools available to extract the same metrics, and the authors reviewed some of them, but it could be very interesting to actually test the tool developed by the authors (showing some validation data) and also compare it with the other available tools.

We appreciate the idea and performed a comprehensive comparison of the results using two different products to calculate the channel width: *RivMap* and our product *cmgo*. We included the analyses in section *8. Evaluation of the data quality*.

- As this is a scientific manuscript submitted to a scientific journal I think that the evaluation of the performance of the tool, and the comparison with existing tools would be more interesting for the readers than all the details about the coding.

Although, it is currently marked as a scientific paper, we submitted it as a technical article as accepted by the Editor. At the Editorial Board Meeting at the 2017 EGU General Assembly it has been decided to extent ESURF's scope towards technical contributions (pers. comm.). This decision is recent and a proper item in the selection box on the submission page (i.e. "Type: Technical paper") was not available by the time of submission. We addressed this in detail in the letter to the Editor.

- Technical details are provided in the tool user guide, or accompanying documentation, and can be also provided as a supplementary material, but in my opinion, it should not be the main core of the manuscript. Instead, authors may provide some details about potential applications, etc. I would recommend replacing the Codeboxes by figures illustrating some of the issues presented and discussed in the manuscript.

For the structure of the manuscript we aimed for a clear distinction of the *algorithm description* and the *technical implementation*. The algorithm description is independent of the implementation and allows to transcribe our approach to any platform. We see two main purposes of the technical descriptions. On the one hand we introduce our various parameters that control and tweak the program results. These parameters are relevant to understand, assess and improve on the quality of the results. For example, the centerline is directly dependent on the parameter settings, while the centerline course directly affects the generation of the transects and thus the channel width calculation. We think that these information are important and well placed in the main body of the manuscript. On the other hand the examples in the Codeboxes enable the users to easily make use of the program, even if they are not very familiar with the R working environment. Since an easy and effortless usage of the program is one of our main goals with this tool we would like to keep our main structure. However, we added real-world comparisons and potential applications in [9. Concluding remarks].

- Although readers might be familiar with the term, GIS should be defined before using the acronym (e.g., line 35).

Thanks for this suggestion. Updated in [35].

- Lines 50-51: give some references here with some examples or relevant works illustrating these efforts.

This sentence is an introduction for all following reviewed products. The individual products are referenced when mentioned in the following paragraph. We did not change anything here.

- Line 60: what do the authors mean by "without providing spatially resolved information"? Please explain what type of information is missing.

Agreed, this is a bit cryptic. We meant that the product calculates one average width value (!) for the entire stream. Local information (e.g. the width each 10 meters) is not generated. We added "(single value)" for clarification in [62].

- Line 67: what type of manual post process is required?

Thank you for the advice. The adequate description has been added in [70].

- Line 68: details are poorly documented?

Basically, the algorithm of the "Polygon to Centerline Tool for ArcGIS" is not documented at all. There is only some information text in the user interface when using the tool within ArcMap. I verified this statement by contacting the author of the Polygon to Centerline tool directly (pers. comm.). We do not intend to discredit the authors approach and we acknowledge the value of the product for many users but the statement regarding the documentation holds. We did no changes to this particular statement but smoothed our statements generally, for example [70ff].

- In general lines 67-69 are very strong statements, I would recommend to smooth them a bit and maybe explain better what the problems with this tool are. I personally have used most of the tools mentioned in this manuscript, and I have realized that they usually need some additional verification or manual correction, but I would not say that the assessment of the data quality is limited.

Agreed, our statement appeared very strong. We would like to argue that the tool – although very helpful for many – acts more or less like a black box, since it is not documented and intermediate results are not accessible. We smoothed the statement in line [70ff].

- The verification of the output can only be done with field measurements or comparing the automatic result with manually digitalized data. When this data is available, verification or assessment of data quality is possible. Did the authors actually test their tool? If so, they should provide the validation data, if not, I would recommend to do that.

Agreed. We added a new section [8. Evaluation of the data quality] to the manuscript. We validated the width at 15 randomly selected transects along a 449 m reach of a natural channel with high agreement of our product [Table 3]. Picking more than 15 locations might not be improving the evaluation since arguably the width of an irregular polygon is better – i.e. more objectively – measured by automated techniques. The following example (Figure 5) shows a comparison of a manual tracking approach and the automated solution. We see two main drawbacks with manual methods: a) transects are not homogenously perpendicular to the centerline (general course of the stream) and b) ledges are often "attractors" to start or end transects. Thus, the user's choice is consistently biased introducing a
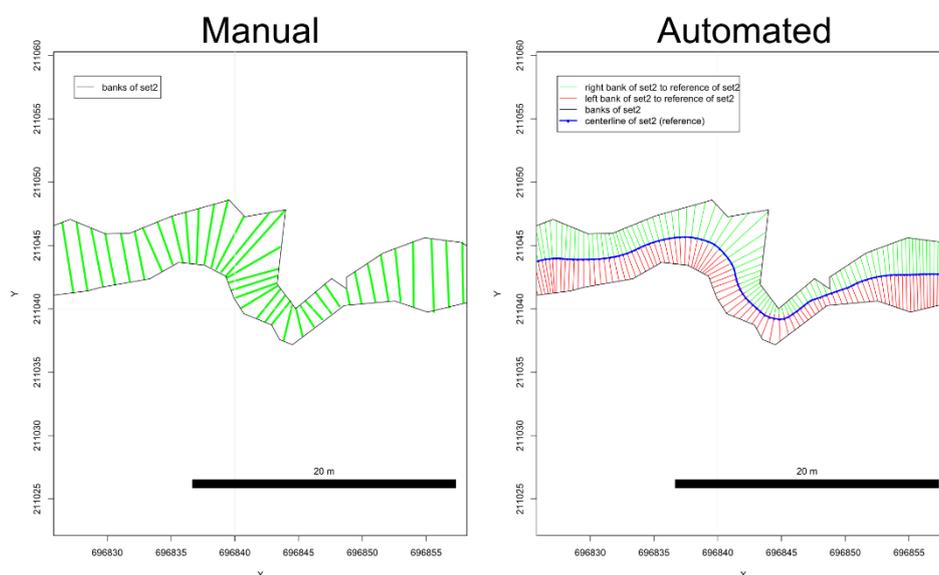
Figure 1: Comparison of methods

systematic error. However, we related to the comment and added chapter [8. Evaluation of the data quality].

- In general, as I said, I miss some comparison between the tools. Maybe one example can be shown in which the authors test all tools and give some values of computational times, input data needs, and some comparison between the outcomes. A table can summarize the results of this comparison exercise.

We took this very good advice and added a comparison of the tools in [8. Evaluation of the data quality].

- More importantly, authors have missed some important tools in their short review: the fluvial corridor toolbox, the Stream Restoration Toolbox (NCED), RivMap and the River Bathymetry Toolkit (RBT). I strongly recommend the authors to check these tools and the documentation and add some text about them. Again, it would be very interesting to add also these tools to the comparison exercise.

Thanks for pointing us to further tools. Generally, again, those tools are all GIS solutions with the mentioned drawbacks [59 and 71ff]. However, we were happy to test each of these tools individually. We address the products one by one next.

Fluvial corridor toolbox: with the demo data set (provided with the toolbox) the generation of a centerline failed on our system (see Figure 6). On our own channel shape the centerline generation did not work at all. Since we could not debug this with the documentation, no further efforts have been done.
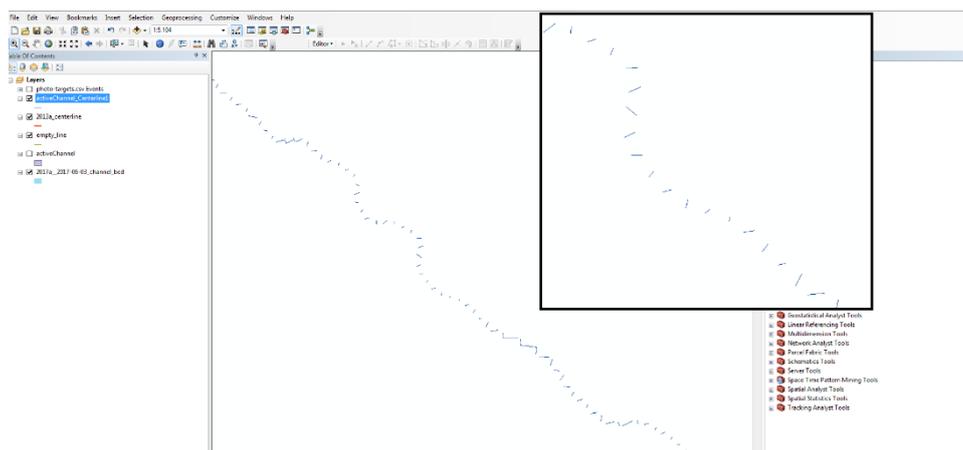
Figure 2: Resulting centerline on our system using the Fluvial Corridor Toolbox

Furthermore, the provided results of the demo data sets reveal that the channel width is not provided along the centerline but along a more generalized version of the stream course (see Fig. 7). For many applications a better control on the reference of the width generation may be required.
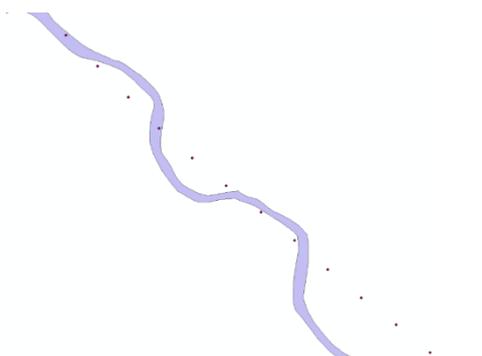
Figure 3: Points at which width was identified by

the Fluvial Corridor Toolbox

Stream Restoration Toolbox (NCED): For us the tool did not calculate a centerline correctly (see Fig. 8). We used 2m of spacing and a maximum of 500 points. The provided input variables are the river banks as requested. Since these where the only parameters and variables and no explanation on the resulting product were given, we did not further investigate this tool. Just as a side note, ArcMap 10.5 crashed three times unexpectedly during performing this task.
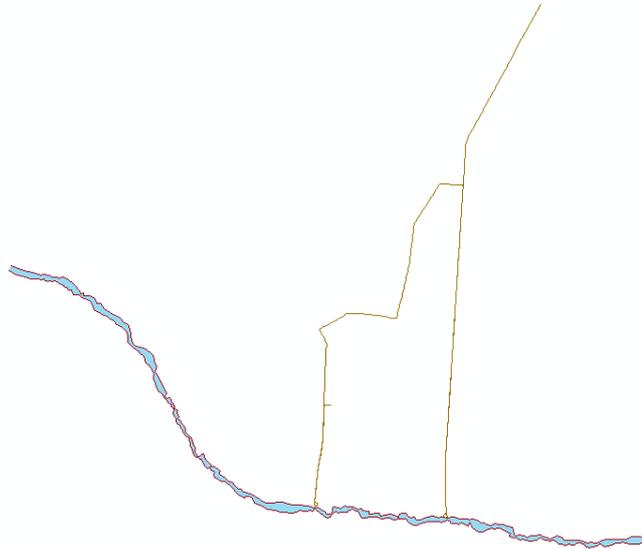
Figure 4: our given channel shape (red lines denote banks) and the resulting centerline (brown line) from the Planform Statistics Tool from the Toolbox Stream Restoration Toolbox

RivMap: RivMap is a very nicely documented MATLAB extension that we have not been aware of. We were able to extract local width information with this tool. We addressed the general drawbacks with proprietary products like MATLAB in [59ff]. We added this product to our review [84ff].

- A Table with the summary of all the available tools would be very useful for the readers.

Agreed, good suggestion. Done in [Table1].

- Here some useful links and references: Schwenk, J., A. Khandelwal, M. Fratkin, V. Kumar, and E. Foufoula-Georgiou (2017), High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali River, Earth and Space Science, 4,46–75, doi:10.1002/2016EA000196. https://ch.mathworks.com/matlabcentral/fileexchange/58264-rivmap-rivermorphodynamics-from-analysis-of-planforms Roux, C., Alber, A., Bertrand, M., Vaudor, L., Piégay, H., 2015. "FluvialCorridor": A new ArcGIS toolbox package for multiscale riverscape exploration. Geomorphology. doi:10.1016/j.geomorph.2014.04.018 http://umrevs-isig.fr. www.sedalp.eu/download/tools.shtml https://repository.nced.umn.edu/browser.php?current=keyword&keyword=5&dataset_id=15 http://essa.com/tools/river-bathymetry-toolkit-rbt/ https://www.fs.fed.us/rm/boise/AWAE/projects/river_bathymetry_toolkit.shtml

Thanks for the references. We incorporated them where suitable in [References].

- Why not using a tool extension for an open source GIS? Besides the limitation of using a commercial GIS, it is not very clear why the authors chose a statistical tool.

Basically, because the generated width information is later analyzed within a statistical tool anyways. Although the data is spatial data, most often the focus of interest is on the metric content, for example magnitude and relative location of features. Thus, we see no need to actually display the data graphically. Actually, for the export of channel metrics from GIS products represents an extra step which

– compared to one-platform solutions – represents decreased efficiency. Other than the drawbacks discussed in [59] and [71ff] no additional information has been added.

Our research on existing open source solutions did not lead to a product that fulfilled our requirements. If the reviewer would reference further tools, we are happy to review them. Moreover, we know that our solution is the only full-stack solution for the statistical language R. Since R has a growing community, requires very little investment (timewise) and has a flat learning curve.

We chose R as a tool that fits all of our requirements. We were careful in the manuscript to separate the description of the algorithm from the description of the implementation. It should be easy to implement the algorithm on other platforms. Clearly, it cannot be the requirement that we provide tools and interfaces for other platforms.

Inspired by this comment we added a discussion point in [451ff] of [9. Concluding remarks].

- Line 83: I would suggest that authors explain a bit better what are the limitations of these tools. Especially the description of the limitations in lines 87-92 would require some additional explanations and data or examples to support the authors judgements.

We had a figure in the earlier manuscript but removed it since we did not emphasize this particular product but make more of a general statement.

The limitations are in detail that transects are not perpendicular to the centerline but to a bank (Fig 9a). This introduces a systematic error that can lead to a large overrepresentation of the local channel width (Fig. 9a). Transects can be corrected but do not objectively represent what would be identified as the width of the channel. Currently, we do not aim to incorporate the figure back into the manuscript but are open to recommendations.
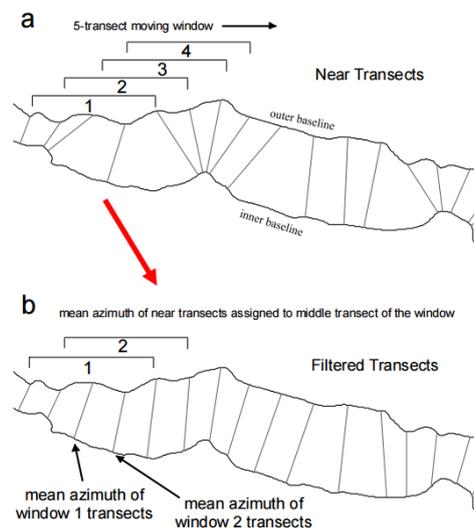
Figure 5: Transect generation of the AMBUR project

- Some information about the type of input data could very useful, such as the number of points or resolution (distance between them for example) that is required or desirable to interpolate correctly the polygon from the bank points. This is only partially mentioned in lines 339-341, but no details or suggestions are provided.

Thanks for the suggestion. Based on your supportive comments we were encouraged to fully revise the chapter *Initialization: Input data and parameters* where we describe the input data in detail [160ff].

- Why using bank points and not bank lines?

The data representation of a number of points and a number of vertices of a polyline is the same. Meaning, when exporting a polyline to an ASCII table it is the same as exporting points. We revised our data input description in [161ff] to clarify.

- If the input data is obtained by aerial picture interpretation for example, instead of by GPS survey in the field, the resulted data would be a line to represent each bank, or even a polygon representing the channel. Are these input data format allowed as inputs to the tool?

The explanation on the input required for cmgo is now revised. It is now clearer that we work on text files containing tabular data. One crucial part in the required data structure is that the points are ordered and have an attribute to which side of the bank each bank point belongs. Please see the updates in [161ff].

- Figure 1: instead of using a, b, c. . .g, the steps could be indicated (as in Table 1).

Very neat suggestion. Thanks! Updated in [Figure 1].

- The spatial scale is not needed in all maps.

Thanks! Updated in [Figure 1].

- Line 97: again the claim that other tools are not transparent or objective should be used with caution. I guess all these tools were also developed to be useful, reliable and they are still used by many researchers.

Your cited line [97 (original manuscripts)] read "Our aim in this paper is to develop a program that does not have the shortcomings of previous approaches and offers a transparent and objective algorithm."

In this sentence we do not judge generally that other algorithms are not objective. With "shortcomings" of the available products we refer to the shortcomings discussed in the review section, which are individual for each product. In the review section we assess the available products to the best of our knowledge and documented the drawbacks with care. However, we rephrased our assessment in [99] and appreciate other solutions as being useful and widely used in [53].

- Line 123: how is the centreline smoothed?

Good point. It is a simple mean filter with definable width. The information was already in chapter *4.2 Controlling the data processing* but we added the information also in line [137].

- What is the difference between smoothed and non-smoothed centreline? Some more explanations here would be beneficial. I think smoothed line is needed here (...)

This is absolutely correct. For all further calculations the smoothed version is used. However, this is also a setting that can be changed by the user. Both versions are stored in the global data object and the user can switch to the "original" version – entirely disabling the smoothing – at any time. More information on the centerline smoothing was already in chapter *4.2 Controlling the data processing* but we added further information in line [248] and edited [Caption of Figure 1].

- (...) because the initial input data are points, and therefore the conversion from these points to lines creates bank lines with sharp edges.

No. Sharp edges in the banks are natural conditions and occur regardless of the type of geometry (points or polyline) the banks are represented by. The necessity of the interpolation and smoothing can be demonstrated also with straight banks (Figure 10, [Figure 2]). Please see the figure and the caption for further explanation. We added further information in line [124] and Figure 10 of this reply as [Figure 2].
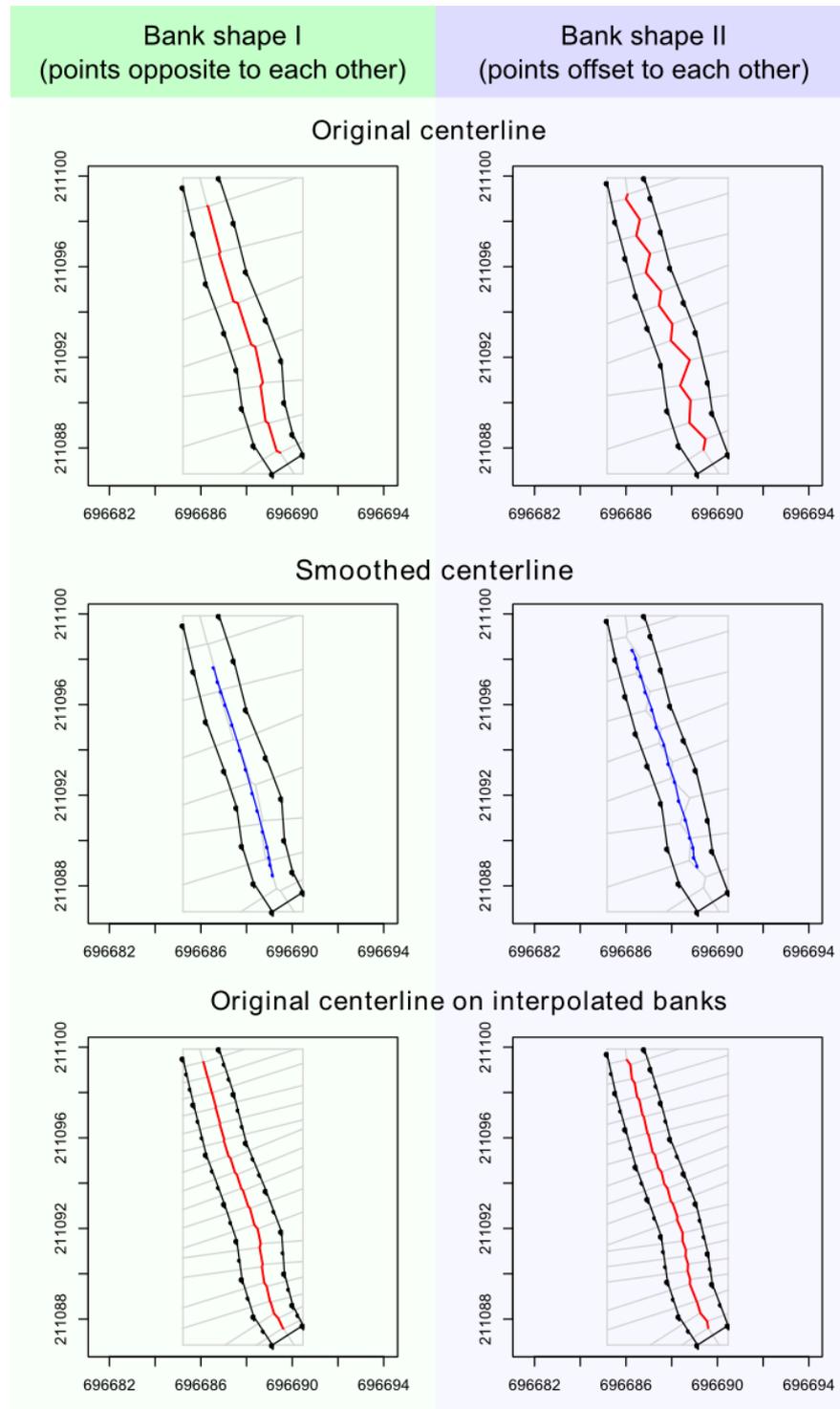
Figure 6: the plot shows to digitizations (Bank shape I and II) of the same channel stretch. They differ only in the arrangement of bank points which are mainly opposite (Bank shape I, left column) or offset (Bank shape II, right column) to each other. One can see how the offset negatively influences the shape of the centerline (top row). The problem can be overcome by smoothing the centerline a-posteriori (middle row) or interpolating between the bank points a-priori (bottom row). A combination of both methods is recommended and set as the default in cmgo.

- This would not happen if the initial input is the digitalized bank for example, (...)

No, it occurs regardless of the source of the bank geometry.

- (...) or would be reduced if the number of points is large enough. This aspect should be discussed more in detail.

Agreed, we did the changes in line [124] and added [Figure 2].

- Line 129 (and also before), how are the Knick points defined by the authors and identified by the tool? And what about the abundance of species? There is not explanation about these aspects, authors may provide examples to illustrate them. This paragraph is again repeated in lines 226-227, but again no explanations are provided.

Those are examples of geospatial data (e.g. quantities with associated x-y-coordinates) the user could have in hand. In cmgo, we offer a function that allows the user to project these metrics to the reference centerline for further analyses. Recently, we implemented also the identification of knickpoints in *cmgo* that is based on the scale-free approach developed by Zimmermann et al. 2008 (Zimmermann, A.E., Church, M., and Hassan, M. a., 2008, Identification of steps and pools from stream longitudinal profile data: Geomorphology, v. 102, no. 3–4, p. 395–406, doi: 10.1016/j.geomorph.2008.04.009.)

We added more information on that fact in line [17], [142] and line [250ff].

- Lines 206 and 224: how are these values selected? Some guidance would be useful.

They have to be selected by the user. Agreed, the guidance was a bit hidden in paragraph *6. Technical fails and how to prevent them* in line [335ff (original manuscript)]. We added a hint in line [228].

- Regarding the smooth see also my previous comment for Line 123.

Already handled above (updated [124] and added [Figure 2]).

- Lines 235-241: Is it not possible to select the distance between transects, instead of using the centreline points? I think this would be very useful. This could be done simplifying the centreline to one single feature line and then drawing the transects to a selected constant distance, or maybe defining the total number of transects.

We think it is not a good idea do decouple the centerline from transect definition at this point of the program routine. The centerline is represented by centerline points and the resolution (spacing) of these points should be a conscious choice of the user (see above, updated [123] and added [Figure 2]). It is normally a trade-off between computational power (too many points) and a poor representation of the channel or a failure of the centerline generation (too few points). Thus, the spacing of the centerline points represents normally the optimum of points necessary and required. Later, the metrics can be down sampled (with or without averaging) with the function CM.reduceCenterline(). The function has not been documented yet but will be available soon. It does exactly what you describe but on the raw data without losing information.

- Lines 241-248: I am not so sure about the assumption that the right bank is the right one with respect to the centreline, this would depend on the river orientation. This should be identify for each case, for example using the elevation values stored in the bank points, or selecting the point upstream and downstream. Otherwise, would be confusing to use the terminology right and left, as for rivers this is given always in the downstream direction.

We also name the banks looking in downstream direction (orographic naming convention). However, for cmgo the user defines the sides by giving the bank points initially the attribute of the side (i.e. "right" or "left"). Whatever he chooses, our statements always refer to this decision and thus remain valid. We included a general statement in the beginning to avoid confusion in line [168].

- Section 4.3: some of the described outputs could be illustrated in a figure here.

Very good suggestion. Updated in [273] and [Figure 5].

- Line 305: I think the sentence in blue and in German should be removed.

We do not see blue or German sentences at this location. No changes have been done in this regard.

- Figure 4: a legend is missing in the figure, although is explained in the caption could be useful to add it also in the graph.

Agreed. Done in (now) [Figure 6].

- Moreover, it is not very clear to me how the tool computes the time series. I understand that it needs a reference centreline, but then I would say that it uses the centreline for each survey and computes the differences between them, right?

Thanks for the comment. First, we changed the paragraph 5. from "Time series analyses" to "Temporal analysis of multiple surveys". You are right with the assumption that *cmgo* needs a reference centerline for the temporal comparisons. From these transects are built from which all metrics are derived. Currently, we can compare those metrics calculated for each survey (e.g. for each year) if they are derived from one centerline. The differences of centerlines is currently not implemented but is on our top 2 of the to-do items list. We have revised the paragraph 5.1 in [341ff].

- I would suggest to better explain for which applications the implemented method is suitable (river widening? Meandering migration?) and also compare it with other tools (e.g., channel migration toolbox). Discussion, implications and limitations are missing in the current manuscript. Section 8 is mainly a technical note describing the tool environment (i.e., R), access and download link, but not conclusions are addressed here. It would be very beneficial for the readers to end the manuscript with some general conclusions about the utility and benefits of this tool, based on the comparison with already existing tools.

Agreed. Thanks for this concluding remark which we used to further improve our manuscript. The limitations and applications are now addressed in the concluding remarks [437ff].