(Modifications to the response published in the discussion are in bold.) We are pleased by the positive and constructive feedback, and thank both Christian Kienholz and Bruce Raup for their reviews. In the following, referee comments are given in *italic*, and our response is stated in regular font. Both reviews included a number of minor technical comments. These are only listed and answered below if our implementation differs from the reviewers' suggestion.

Review by C. Kienholz

In certain places, additional explanations would be helpful (e.g., why did you introduce certain thresholds? why did you use a certain equation?), but I'm aware that this could lengthen the paper too much.

We assume that the reviewer mainly refers to the trade-off functions. We agree that the description of the latter is condensed. Nevertheless we would like to keep the description of the trade-off functions short. As the reviewer points out, adding explanations on why the various settings and equations were chosen would lengthen the paper substantially. Furthermore, we would also like to emphasize that the basic idea of our algorithm is to combine the two main conditions of maximizing distance-to-margin and surface slope by means of a set of trade-off functions. The here applied trade-off functions should be seen as just one out of various possible solutions to the problem of trading between surface slope and distance-to-margin.

p.2491 Title: I would remove the straightforward". Your method is rather sophisticated. Something like "The length of the glaciers in the world – global application of an automated method for the calculation of glacier center lines" may be an alternative. Although implied by "glaciers in the world", the global can be emphasized, as none of the previous methods have been applied on a global scale.

We have changed the title to "*The length of the world's glaciers - a new approach for the global calculation of center lines*". We wanted to avoid the word "automated" here since it can be understood in different ways (see the following reviewer comment).

p.2492, l.5 It would be beneficial to define your use of fully automated in the paper. Strictly speaking, the approach to derive the physical centerlines is not fully automated. As shown in the Greenland quality analysis, a small percentage of the centerlines is erroneous (with errors that are difficult to quantify), no matter the quality of the DEM and outlines. However, regarding the derived lengths, the term "fully automated" seems appropriate because here, you can easily work with uncertainties (as done on p. 2506).

We would like to thank the reviewer for pointing out that "automated", "semi-automated" and "manual" can be understood as being judgemental (e.g. "manual" would be worst and "automated" best because it requires the smallest investment of time). Our use of "automated", however, refers to the complete absence of manual interventions and does not imply any level of quality. To clarify this point we have included the following statement in the manuscript : *"In the context of this paper, "fully automated" refers solely to the complete absence of manual interventions, it does not imply that automation partly or completely replaces manual corrections as performed by by Le Bris and Paul [2013] and Kienholz et al. [2014]."*

Furthermore we have replaced "fully automated" by "automated throughout the entire manuscript. We have also removed several occurrences of "semi-automated".

p.2492, l.5 It would be beneficial to define your use of fully automated in the paper. Strictly speaking, the approach to derive the physical centerlines is not fully automated. As shown in the Greenland quality analysis, a small percentage of the centerlines is erroneous (with errors that are difficult to quantify), no matter the quality of the DEM and outlines. However, regarding the derived lengths, the term "fully automated" seems appropriate because here, you can easily work with uncertainties (as done on p. 2506).

See our comment above.

p.2493, l.3-13. In case of 2D applications (unlike 3D flow modeling), don't the terms "flow line" or "trajectory" per definition refer to the 2D projection of the actual 3D flow line or trajectory? To me, this is somewhat similar to the use of the term glacier area. In a 2D application, area always refers to the projected area, rather than to the actual 3D surface or slope-corrected area. While it is difficult to derive the 3D flow line from measured data, its 2D projection can be approximated to a certain degree, using surface velocity fields or even elevation contours. Following the explanations of the different uses of the term "glacier length", what is the definition of glacier length used in the paper? Quality criteria are given later on (p. 2502), but it would be beneficial if you quickly defined the type of line envisaged.

We agree with the reviewer that real 3D flowlines can be approximated in 2D and that terms like "flow line" refer to a 2D projection in applications as ours. Nevertheless, our statements should highlight that any 2D projection involves a certain level of ambiguity because the three dimensional glacier flow field is generally incompletely known.

It is a valuable suggestion to define what kind of line we envisage. We have added a brief statement to the manuscript: *"In the remainder of the paper we refer to model output as "center lines". Thereby we refer to a very similar concept as introduced by Kienholz et al. [2014], with the exception that calculated lines in accumulation areas adhere less strictly to the glacier center and can take some of the characteristics of flow lines."*

p.2496. l. 19. The numbers of inaccurate RGI glacier outlines must make up more than 1% of the total?

The reviewer is right in pointing out that we would need to also provide a definition of how to distinguish "accurate" and "inaccurate" polygons. Such a definition would go beyond the scope of our study and thus we have simply removed the statement "... *(estimated as 1 % of the total)* ...".

p.2499. l. 17. If confluence occurs in the ablation area, the lines might get deflected to the opposite margin of the glacier, due to the convex topography and the increased weight of the slope component. How robust is the approach in dealing with this?

Additional weight to the slope component was introduced because center lines had a strong tendency of getting attracted by the large distance-to-margin values in the center of confluence areas. Adding weight to the slope component in areas where the glacier widens resulted in a good performance of the algorithm in confluence areas with no tendency of getting deflected to the margin. The problem of lines being deflected to the glacier margin nevertheless exists but is generally related to rather narrow glacier tongues in the ablation area that are strongly convex, often also exaggerated due to DEM inaccuracies.

p.2500. l. 15. How effective is this approach in correctly separating marine-/lake-terminating and land terminating glaciers? While the correct separation is not necessarily required for your centerline application (it is ok to have suggested endpoints for wide, land-terminating glaciers), it is useful for studies that need to distinguish the different glacier types. If your approach works reliably, you have a useful side-product that should be mentioned, e.g., in the Conclusions.

The reviewer mentions an important point here and the second reviewer addresses the same point from a slightly different angle. If DEM and glacier outlines are of high quality *and* represent the same point in time, then this simple approach is quite reliable in distinguishing marine/lake terminating glaciers from other glaciers. Even the margins of wide land terminating glaciers tongues (e.g. piedmont type) have generally larger elevation differences than a calving front. However, we encountered a number of cases where the simple approach does not work correctly either due to flaws in the DEM, the glacier polygons, or, quite frequently, because the DEM and the polygons do not correspond to the same point in time. The latter problem is particularly related to marine/lake terminating glaciers because terminus retreat due to calving can be much faster than the melting of land-terminating glaciers.

We have provided additional statements in the manuscript discussing potential issues due to non-contemporaneous data sets (see our reply to Reviewer 2, first comment).

p.2501. l. 14. Can F2 cause the centerlines to lie outside the actual glacier perimeter?

The reviewer is right that the filter can result in center lines that run outside the glacier polygon for short distances. Further improvements of the filtering are required. We have added a brief statement to the manuscript: *"The filtering introduces a small risk that short sections of a center line fall outside of a glacier polygon"*.

p.2502. l.2. It would be of interest to have some kind of sensitivity analysis. For example, how different are the lengths if you set c_0 to 0.5 instead of 0.6? However, given the number of parameters, such an analysis may go beyond the scope of this paper.

In the process of developing the algorithm various parameter settings were tested. Changing c_0 to 0.5 instead of 0.6, for instance, does have only a marginal influence on the calculated lengths. However, the calibration of the algorithm is of a qualitative nature. In this sense a statement like "Changing c_0, c_1, ... by X % modifies glacier length by Y %." is of limited significance because the modification in glacier length could at the same time result in violation of, or better agreement with, the qualitative criteria listed in Section 4. We agree with the reviewer that a sensitivity analysis would be desirable, but such an analysis should not rely on glacier length alone and would thus require to first operationalize the qualitative criteria listed in Section 4. The latter would go beyond the scope of this manuscript.

p.2502. l.16. Flexibility is needed to deal with inaccuracies, but you might also need some flexibility to deal with rough surfaces that are real, e.g., due to debris cover.

The reviewer is right in stating that surfaces of debris-covered glaciers can be much more uneven than debris free glaciers. Since surface undulations due to debris cover have a repetitive character, they are well addressed by (i) the sink-fill and smoothing of the glacier surface, and (ii) by maximizing the search radius as implemented in the algorithm. Difficulties might arise where only certain sections of a glacier are debris covered, such as for instance elevated medial moraines. However, the size of such real surface undulations is mostly smaller than elevation differences related to erroneous DEM data. Nevertheless, further testing is required for areas where contemporaneous DEM and glacier outline data of highest quality exist. This would allow assessing the performance of the algorithm while ruling out the influence of inaccuracies in input data.

p.2503. l.3. "coarse resolution and possible uphill flow"

We believe that using the words "uphill flow" might be misleading. Instead we added the following statement: *"... possible leapfrogging over real* *surface undulations".*

p.2504. l.11. This supports the interpretation of "fully automatic" given on p. 2. The derivation of the actual centerlines is not fully automatic, but the derivation of the lengths is if a certain error percentage is considered.

We agree with the reviewer's comment. Our algorithm is not capable of calculating every possible center line accurately. A certain percentage of the calculated lines is erroneous, but no automated mechanism is able to solve any task with perfect reliability. Because of attempting to calculate the length of all glaciers in the world, we ruled out manual corrections from the start and decided to rely entirely on automation. Hence, our study explores the level of accuracy that can be achieved if avoiding all manual interventions.

p.2505. l.18. Is any influence of the debris coverage noticeable? The rough topography in debris covered areas may also stop centerlines too early.

An influence of the uneven debris covered surface is not very likely because the DEM has been sink-filled and smoothed. Furthermore, surface undulations due to debris cover are often small compared to the width of larger glaciers. On wide glaciers the algorithm tends to apply long search distances which maximizes the likelihood of finding cells at lower elevation and continuing the center line. For these reasons we assume that on the one hand a continuous debris cover is rarely the reason for stopping center lines halfway. On the other hand, chaotic middle moraines as they can be found on surging glaciers might deflect the calculated center line away from the center. To investigate whether such issues really exist would require testing with contemporaneous highest quality DEM and glacier data. To prepare such data sets for regions with challenging glacier geometries, involving glaciers with and without debris cover as well as calving glaciers (e.g. Alaska, Patagonia, Greenland) would go beyond the scope of this study.

p.2508. l.17. "Averaging over perimeters of a few grid cells". If a smoothed DEM is used (as done on the global scale), such an averaging effect is included?

The reviewer is right, the smoothing already includes such an averaging effect. The statement is superfluous and has been removed from the manuscript. Instead we have added a sentence to once again highlight that the trade-off functions could also be designed differently: *"We have presented one possibility of implementing the basic conditions and designing trade-off functions; alternatives to improve accuracy and efficiency certainly exist. This concerns in particular the trade-off between surface slope and centrality where alternative designs should be explored. One might also imagine involving additional conditions, such as the direct inclusion of observed glacier flow fields."*

p.2508. l.21. The relatively coarse spatial resolution of up to 250 m enables a high computational efficiency of the algorithm. Might the coarse spatial resolution (combined with the applied filter functions) yield centerlines that don't strictly lie within the glacier perimeter? (which can be problematic for certain follow-up applications)

The reviewer addresses this potential problem here and once again in connection to the applied filtering. Furthermore, the question is raised if the relatively coarse resolution is one reason for the short computation time. The calculations will take longer when higher resolution DEM data is used. The computation time is more or less a linear function of DEM resolution because search radii are basically independent of cell size and the number of cells contained in a search ring grows more or less linearly with DEM resolution.

Center lines always progress from the center of a grid cell to the center of another grid cell. Evaluation of the output shows that the unfiltered center lines always lie within the glacier perimeter. Exceptions can only occur if the center of a glacier grid-cell lies outside of the actual glacier perimeter, whereby the latter is defined by the vectorized glacier polygons. In such cases the rasterization might be inappropriate for the application.

p.2509. l.16 - 23. These explanations also imply that Tables 1 and 2 should be compared with certain care.

Yes, this is correct. To underline this issue, we have now mentioned the number of glaciers also in the caption of Table 2. Now it should be more obvious to the reader that the two similarly looking tables refer to different glacier samples.

p.2512. l.13. Do you also have a data set that distinguishes marine-terminating/lake-terminating from land-terminating glaciers? If so, mention it here.

We do not have such a data set. For certain regions, for instance Greenland, the glacier inventories distinguish between marine and land terminating glaciers [e.g. Rastner et al., 2012]. Such information has not been used in algorithm development and application.

p.2512. l.23. A few more explanations in this appendix A1 may be beneficial. Why did you use certain equations, why certain thresholds?

Please see also our reply to the first comment of the reviewer. We have added a few brief descriptions but would like to keep the appendix short.

p.2520. Increase the size of the figure, if possible.
p.2524. Increase the size of the figure.
Both figures should become bigger in the final manuscript due

to the different page format of The Cryosphere.

Figure 9: erroneous scaling relationships.

We would like to thank Reviewer 1 for notifying us after the submission of his review that the scaling relationships shown in Figure 9 were erroneous. The updated version of Figure 9 shows the corrected scaling relationships. **The corresponding numbers in the text have also been modified.**

Review by B. Raup

Have the authors quantified how important it is to have the glacier outline and DEM be contemporaneous? What errors can result if they are not?

The reviewer points to an important issue when working with DEM data and glacier polygons. We have qualitatively analysed the difficulties related to non-contemporaneous input data and found that problems are mostly related to marine- or lake-terminating glaciers. We have expanded the first paragraph in Section 4.2 to provide more background information. Furthermore, we have explained why Figure 4 shows glacier outlines for two different points in time: *"For the vast majority of glaciers, the center lines end where envisaged, but there are a few locations where the automatic approach suggests erroneous endpoints. Example 3 in Fig. 5 illustrates the difficulties in defining representative end points along complexly shaped calving fronts of retreating marine-terminating glaciers. The most frequent source of erroneously suggested end points is that DEM and glacier polygons are not contemporaneous. Example 4 shows an erroneous ending due to the complete absence of topography at the lowermost part of the glacier tongue. While the glacier polygon represents the year 1999 [Rastner et al., 2012], the DEM apparently shows the situation a few years later after a pronounced retreat of the glacier terminus. Non-contemporaneous DEM and glacier polygons are shown in more detail in Fig. 4 where lateral margins are at 0 m a.s.l. for a few kilometres inland, indicating an absence of topography in the DEM. Indeed, analysis of 2013 satellite imagery shows that the glacier has retreated for about 8 km since the year 1999. Problems related to non-contemporaneous input data mainly occur on marine-/lake-terminating glaciers because (i) terminus retreat due to calving can be much faster than the melting of land-terminating glaciers, and (ii) calving leaves a perfectly levelled surface behind."*

The center lines generated by this method, at least as shown in Figure 3, seem to be represented by a small number of vertices. Can the algorithm be tuned so that more vertices are produced for smoother center lines?

It is true that the center lines are represented by a relatively small number of vertices. This is a direct consequence of maximizing search radii when finding the next vertex. The settings of the algorithm can be changed to reduce search radii which will increase the number of vertices. The calculated center lines, at least when using a low- to intermediate-quality DEM, will not become smoother because shorter search radii increase the sensitivity to surface undulations.

Nevertheless, we believe that the distance from one vertex to the next is in most cases appropriate, as for instance illustrated in Figure 4. Narrow glaciers show closely spaced vertices while wide glacier tongues have a larger spacing. Creating more vertices on the large glacier would not result in a substantial gain in information and center-line accuracy. Visually more pleasing center lines could be achieved by implementing some sort of curve fitting algorithm rather than simply connecting vertices with straight lines.

The descriptions of the trade-off functions are short enough that they could be in the main text, rather than in an appendix.

We have placed the detailed description of the trade-off functions in the appendix for two reasons: (i) The description consists basically only of variables, simple equations and numbers. Placing them inside the main text would unnecessarily disrupt the text flow. (ii) The applied trade-off functions are just one solution to the problem of trading between slope and distance-to-margin. More straightforward designs likely exist and this should be underlined by showing the details of the trade-off functions outside of the main text. We believe that this was not made clear enough in the manuscript and thus we have also added statements in the main text highlighting the need to further experiment with the design of the trade-off functions.

It would be good to state in the introduction something like this: The GLIMS Glacier Database stores center lines for glaciers, and currently contains center lines for about 2300 glaciers. As the RGI is merged into GLIMS, it is expected that center lines from an automated method such as this will be put in the GLIMS database.

We believe that the suggested statement does not fit well into the introduction which focuses somewhat more on length rather than the center lines in a global context. Instead we have placed a similar statement in the "Discussion": *"Glacier length is an important, yet missing parameter in global glacier inventories [Paul et al., 2009]. The GLIMS Glacier Database, for instance, stores center lines for glaciers, and currently contains center lines for about 2300 glaciers."* We believe that mentioning the future merging of RGI and GLIMS is beyond the topic of our manuscript. Furthermore, we also would like to avoid a statement such as *"...it is expected that center lines from an automated method such as this will be put in the GLIMS database."*. This might be misunderstood as promoting our own method over other algorithms that are at least of equal quality.

- Is the F1 filter the same as, or different from, the Douglas-Peucker algorithm? Why not simply use the D-P algorithm?

We have not tried to apply the Ramer-Douglas-Peuker (RDP) algorithm and we think that the RDP cannot replace F1. The latter removes the vertices that see their two neighbouring vertices under the most acute angles. The RDP does somewhat the opposite because points that are furthest away from a direct line connecting start and end point (or start and end points of segments) are kept. However, exactly these points tend to form the most acute angles with their neighbours and should be removed by the filtering. Nevertheless, we would like to thank the reviewer for this suggestion as it could be very interesting to explore the effects of implementing the RDP elsewhere in the framework of the algorithm.

Is there a plan to make the code available to the community?

The code is distributed to anyone who is interested in using or modifying it.

It is good to have the comparison with the semi-automated method of Kienholz, et al. (2014) in Table 2. It would also be good to show a sample of Kienholz outlines in a figure with outlines from this method, as in Figure 4, for visual comparison.

We have added one more figure showing a direct comparison of longest center lines calculated by Kienholz et al. [2014] and by our method.

3-5: Change to "because of the link between length and glacier flow"

We prefer a different formulation here since "... because of the link between length and glacier flow ..." could imply that there is a physical relationship between the two. Instead length is just linked to flow because of being defined so.

8-12: What does "suggested" mean? Please clarify.

We are aware that the word "suggested" is mentioned out of context here. For this purpose we added a reference to Section 3.4 where the concept of "suggested end points" is explained.

10-2: Shouldn't the limits of the summation be 1 to 3 (not 0 to 3)?

We thank the reviewer for pointing out this ambiguity. Indeed, 0 to 3 is correct, but we missed to point out more clearly that the basic trade-off factor c_0 is also meant to be part of the summation. We tried to make this clearer by explicitly mentioning the symbol c_0 and adding a reference to Fig. 2b where the summation is shown in full.

12-11: Clarify what is meant by "slope glaciations". Please use standard terminology, for example from the "Illustrated GLIMS Glacier Classification Manual"

Here and throughout the entire manuscript we have replaced "slope glaciation" with "ice apron" as suggested by the source mentioned by the reviewer.

12-12: It is unclear how the last two criteria are considered only in certain cases, given that this is a fully automated algorithm. Please clarify.

These are the criteria used to qualitatively assess the model output during the process of calibration. They should not be mistaken for criteria directly implemented into automated calculation. Instead the criteria are implemented indirectly by means of design and calibration of the trade-off functions.

13-3: Change "optimization" to "increase"?

With "optimization" we mean a reduction of computation time. To make this clearer we have replaced "optimization" with "reduction".

13-6: Is it c_0 to c_2, or c_1 to c_3? See also comment 10-2.

Again, thank you for pointing out this ambiguity. We actually mean c_0 to c_3. We hope this becomes clearer now due to the changes we made on page 2499, line 2.

14-2: Change "masked" to "created" or "compiled"? Meaning is unclear.

We have replaced "masked" with "hidden" since the meaning was obviously unclear.

22-15: Restate the four conditions briefly, for the benefit of people reading only the conclusions.

We would like to avoid the listing of technical details in the conclusions and thus prefer to not explicitly re-state the conditions. Instead we have expanded the statement about the algorithm to the following: "*We have presented a grid based and fully automated algorithm to calculate glacier center lines. Our method is based on the two conditions of maximizing surface slope and distance to the glacier margin. A set of three trade-off functions analyses the local morphometry of the glacier and depending on the latter, flexibly controls the weight of the two basic conditions.*"

Fig. 4: It seems irrelevant to have glacier outlines for two different years in the figure.

Showing the glacier outlines for two different years was done with the purpose of illustrating issues related to temporal disagreement of DEM and glacier outlines. However, the reviewer's comments made us realize that we missed out on mentioning the impact of such issues in the manuscript. We have done that now as explained in our response to the first comment of Reviewer 2.

Fig. 5: The center line on the large glacier that terminates at about x=655, y=-2260 seems to end in the wrong place. It would be worth adding a short comment about that (whether it is correct as-is or not).

The reviewer is right that this particular center line ends not at the point that we consider to represent best the center of the calving front. The example illustrates the difficulties in defining and finding a representative end point along the often complexly shaped calving fronts of retreating glaciers. We have modified the figure and added a comment to the manuscript.

Fig. 6: Please use a color other than yellow, as it's very difficult to see.

We tried to use the same colors as Kienholz et al. [2014] to ease direct comparison to the similar figures in the aforementioned study. We would like to keep colors similar but have changed the yellow both in Figs. 6 and 7 towards a more orange tone and the previously orange tone towards a more red colour.

Fig. 8: For some regions, there must be previously known "longest" glaciers. Do these results agree?

To our knowledge this is the first systematic assessment of global glacier length, and also the first that uses the regions as defined in the RGI. We have not performed a detailed comparison to previous sources as most available information on glacier length lacks comprehensible documentation of source material and methodology. For a few regions, previously measured longest glaciers and our results can be compared and are in agreement: Aletsch glacier is considered the longest glacier of the Alps, for New Zealand Tasman glacier. Potanin glacier, for instance is considered the longest glacier of Mongolia and the Altai, but is not mentioned as the longest glacier of "North Asia" because the latter is a geographic unit created quite recently for the use in the framework of worldwide glacier modelling and glacier inventories. The situation is similar with glaciers like Fedchenko and Siachen. They are mentioned as longest glaciers, but the mentioning generally refers to different geographic units as used in the RGI. Fedchenko again is frequently mentioned as the longest glacier outside the polar regions, but whether this is correct depends on the question if southern Alaska belongs to the polar regions or not. Interestingly, Bering glacier is rarely mentioned in the context of "longest glaciers in the world", while sources from the Soviet Union seemed to have stated already decades ago that Bering glacier is the longest mountain glacier in the world.

Fig. 9: Include actual data on the graph for at least one region.

The figure should illustrate the scaling laws and we believe it might be puzzling if actual data from only one region are shown. Adding all the data would make the figure unreadable, the same might even be the case with data points from only one region since the number of glaciers per region is also large. The manuscript mentions the R^2 values (p. 2507, l.9) and for these reasons we would prefer to keep the figure as it is

References

C. Kienholz, J. L. Rich, A. A. Arendt, and R. Hock. A new method for deriving glacier centerlines applied to glaciers in Alaska and northwest Canada. *The Cryosphere*, 8:03519, 2014. doi: 10.5194/tcd-7-5189-2013.

R. Le Bris and F. Paul. An automatic method to create flow lines for determination of glacier length: A pilot study with Alaskan glaciers. *Computers and Geosciences*, 52:234–245, 2013.

F. Paul, R.G. Barry, J.G. Cogley, H. Frey, W. Haeberli, A. Ohmura, C.S.L. Ommanney, B. Raup, A. Rivera, and M. Zemp. Recommendations for the compilation of glacier inventory data from digital sources. *Annals of Glaciology*, 50(53):119–126, 2009. doi: 10.3189/172756410790595778.

P. Rastner, T. Bolch, N. Mölg, H. Machguth, R. Le Bris, and F. Paul. The first complete inventory of the local glaciers and ice caps on Greenland. *The Cryosphere*, 6:1483–1495, 2012. doi: 10.5194/tc-6-1483-2012. URL `www.the-cryosphere.net/6/1483/2012/`.