General comments:

Priergaard Zinck and Grinsted present a comparison of two models to estimate ice thickness distribution of the Müller Ice Cap in the Canadian Arctic. The thickness information is further used to select a best site for drilling an ice core. The brief communication paper is generally well structured and written and discusses a relevant topic. There are however some fundamental issues that make that I can at present not recommend this article for publication. My major concerns are related to the methods and concern both the simple SIA inversion and the PISM-based inversion. In my opinion, certain modelling choices are not justified and the comparison of the two methods in its current form does not give too much useful insight. Thorough revisions would be required to make the study worthwhile and would require a change of strategy. My major and minor comments are further explained below.

Thank you for the review.

First, we would like to emphasize that a primary goal of the study is to document our reasoning for choosing this particular candidate for a new drill site. There is no need for this to also yield new 'insights'. This is why we eventually decided to submit this as a Brief Communication. -Our editor informed us that the site selection would be sufficient to warrant this type of manuscript.

Nevertheless we hope to convince you that our new regression based semi-empirical approach is useful. The key insight is that fitting a SIA-motivated semi-empirical model to estimate ice thickness is equivalent to fitting a simple linear regression model. This may not give insights to the physical processes, but we demonstrate that it performs better. A 'semi-empirical' approach is of course more statistical, which is both a pro and a con of the method compared to more physical approaches. The statistical point of view can also be seen as a bridge to the inevitable machine learning methods which have already started to appear in the literature, and which we will also see more of in the future.

Lastly, we realize that we have not put enough focus on the fact that the two models actually agree very well in the area of interest, which we will make sure to do in the revised manuscript.

Major comments:

SIA inversion:

The only difference of this method and other SIA-based inversion methods (see e.g. Farinotti et al. 2017, 2021) is that this method takes more freedom to calibrate the inversion method. But by individually calibrating a, b and k, which are normally dependent on another as they all are a function of exponent n, the model is no longer *really the SIA model after calibration. So the separate calibration of a, b and k in practice leads to a model that no longer follows the same physics as the SIA and it is unclear what physics it does follow instead. Even when a relatively good fit with observed bed data can be found with this tuning of a, b and k, it does not give much confidence in good performance of the same model with these parameter values elsewhere. It would make much more sense to only calibrate parameters like Glen's exponent n or rate factor A, which would not change anything to the physics of the model.*

We understand that calling the final model "SIA inversion" might be taking it too far. Therefore, we have changed the name of it to "semi-empirical ice thickness model (SEITMo)". The motivation for not calibrating A and n directly is that the aim is not to achieve the true value of these, but rather have more freedom to achieve the best possible results on the ice cap near the ice divide.

Theory will only match reality if:

1. our physical description is perfect.
2. The input data (H_icebridge, Q and α) are perfect.

We know that neither of those two conditions are true. The SIA model is a simplification, and the assumptions going into it are clearly not valid near ice divides (the key region of interest in this study). Further, no input data is ever perfect. The different input data will be of varying quality and therefore should not necessarily have equal weight in the final thickness determination. It is therefore not surprising that the best fitting parameters will deviate from theory.

We note that A. Zinck has tested a physically constrained Bayesian calibration of a SIA model in the cited MSc thesis. We decided to not include this model in this manuscript because 1) It does not perform better. 2) it has higher complexity and cost. 3) Something similar has already been done in the literature [e.g. Brinker et al. 2016]. 4) manuscript length.

Please also note: In our revised manuscript we have decided to include a comparison to a more clean SIA based method that used velocities instead of Q (Zorzut et al., 2020). We include this method to satisfy reviewer2, and because it is a similarly fast method to our semi-empirical approach.

Mass balance uncertainty:

In the SIA inversion it is claimed that the mass balance has not much impact on the reconstructed ice thickness. This in reality is of course not the case since the mass balance is a major factor that affects ice extent and average thickness. The reason that it does not play a major role here is that any biases in the mass balance (as clearly seen in Fig. 1) are indirectly compensated for by the tuning of parameters a, b and k. But this *does not mean that ice thickness is not sensitive to mass balance, any offsets are simply calibrated away. From Figure 1 and the Discussion section it becomes clear that there is a major overestimation of mass balance in the HIRHAM product. This bias is ”calibrated away” / corrected for in the SIA inversion but the same is not done in the PISM-based inversion, where only parameters affecting ice flow speeds are used for tuning. It is hence not surprising that the PISM-based inversion provides worse fits to the bed data as the degree of tuning is much less. In other words, the degree of tuning of the two models used in this study is very different and makes it hard to draw any strong conclusions from the current results.*

> We agree that mass balance is important for some models like the PISM. We also agree that the degree of tuning is not the same for both approaches, and so we also agree that there are limitations to how strong conclusions we can draw from the degree of misfit. We therefore offer to simply completely remove PISM from the manuscript. However, it is important to realize - and we may not have communicated that clearly - that the computational cost of the PISM approach is so incredibly huge that it places severe limitations on how large a parameter space it is possible to explore. The PISM inversion (with 2 parameters) in the current manuscript took months. We are talking 3x4x10x2000 model years (N_ExN_{phi}xN_{iter}xN_{yrs}). Adding three variations for two additional parameters for a mass balance offset and scaling would take ~9 times as long. This is a very real practical concern that strongly favors the semi empirical approach (or approaches like Zorzut et al.). So we would like to keep the discussion of PISM in the manuscript as this could be important for other people. We will make sure to acknowledge in the revised manuscript that better fits could be obtained by also tuning the SMB.

PISM-based inversion & factor K:

In the Methods it is mentioned that the PISM-based inversion uses a factor K that varies between 0 and 0.5 from 500 m a.s.l. to 1000 m a.s.l. This implies that below 500 m a.s.l. elevation K=0 and the bed is not modified at all after every iteration. Effectively, that means that at these elevations the bed remains at the initial height which, correct me if I am wrong, is taken from the global estimates by Farinotti et al. (2019). In the Discussion it is argued that the overestimation of ice thickness of the outlet glaciers (which are mostly below 500 m a.s.l.) is a result of the lack of a calving criterion, but I do not think the calving criterion plays any role here since the bed is fixed anyway in these lower areas of the outlet glaciers. The large deviations of bed heights between the PISM inversion and the observations, as shown in Fig. 1, nearly all happen in areas below 500 m a.s.l. Effectively, we are looking at a comparison of the Farinotti et al. (2019) bed and the observations for a large part of the domain. My suspicion is that the choice of variable K values with altitude was based on problems to make the PISM based thickness inversion converge. The lack of conversion should however not have been solved by choosing a variable K value, but rather by correcting a bias of the mass *balance. A strategy with combined calibration of a mass balance correction and ice flow parameters, and a fixed value of K, would probably have yielded much more reasonable results. It is good to realize that a more detailed inversion technique, i.e. with a more accurate decription of ice motion and boundary conditions (mass balance, calving), should theoretically yield better thickness estimates than simplified approaches as long as the input data (DEM, mass balance and/or ice velocity) is of sufficient quality.*

> It is correct that the bedrock below 500 m is not modified after each iteration. The bedrock shown in Figure 2 and 3 is the end result after the last iteration, so it is not equal to the Farinotti bedrock but the bedrock after 2000 model years (calculated as ArcticDEM minus thickness). However, we agree that we should not include this part of the domain in the comparison. It should be noted that the semi-empirical approach is also not tuned to this area, but only the non-shaded part of Figure 2.
>
> Low elevations have too much melt to be useful ice core sites. Furthermore, the ice flow in these areas is a complicating factor when an ice core record has to be put into a climatological sense. This informed our design choices. We are simply not interested in good performance at low elevations, and thus chose to not adjust K for z<500m and thus did not have to worry about getting ablation right. This is critical for performance as it reduces the parameter space we have to search.

Specific and minor comments:

L5-6: See the second major comment above. The SIA inversion is not insensitive to mass balance (or at least it should not be), but mass balance biases are indirectly calibrated away through tuning of a, b and k.

> We agree that the SIA in it’s theoretical form is not insensitive to mass balance. However, our semi-empirical approach is insensitive to multiplicative errors in SMB. See also our comments above.

L12-13: Large ice thickness does not always mean very old ice at great depths, only if it is close to an ice divide.

> We will make sure to be more specific here and change it to:
> “Ice of great thickness and minimal horizontal flow is desirable to increase the probability of reaching ice dating back...”

This statement is based on experience from the ice core experts in Copenhagen.

L17-18: "However, field work constraints…" This does not connect well to the previous sentence, more to the one before that.

We agree with that. Thank you for pointing it out. In the revised manuscript we will swap the two sentences around and make sure they are better connected with a better flow.

L18: Please replace "to be clever" with "to be selective".

Changed.

L21: Please remove the obsolete bracket

Bracket removed.

L22: Remove the obsolete "in"

Done.

Figure 1: UTM x and y should be replaced with UTM Easting and Northing instead and the UTM zone should be mentioned in the caption. Furthermore, exponents in the SMB units are missing.

Changed to Easting and Northing and the UTM zone has been added to the caption.

The SMB exponents were lost during upload in the processing by Copernicus. We will carefully verify that they are not lost in the next version.

L27: “differs” --> “differ”

Changed.

L29-31: The new method is not necessarily less sensitive to mass balance, steady state assumptions and ice flow physics. It just collectively calibrates any biases due to these factors away. But the problem is that by doing so the physics of the model are also changed, which is hard to justify. See my first major comment above.

We disagree. It must be fair to state that the new method is insensitive to errors in the different factors, if the resulting biases are automatically calibrated away. This is at least what we mean when we say that it is insensitive. We also think we very openly and transparently explain that the insensitivity arises from how biases are calibrated away.

L36: “why” --> “and”

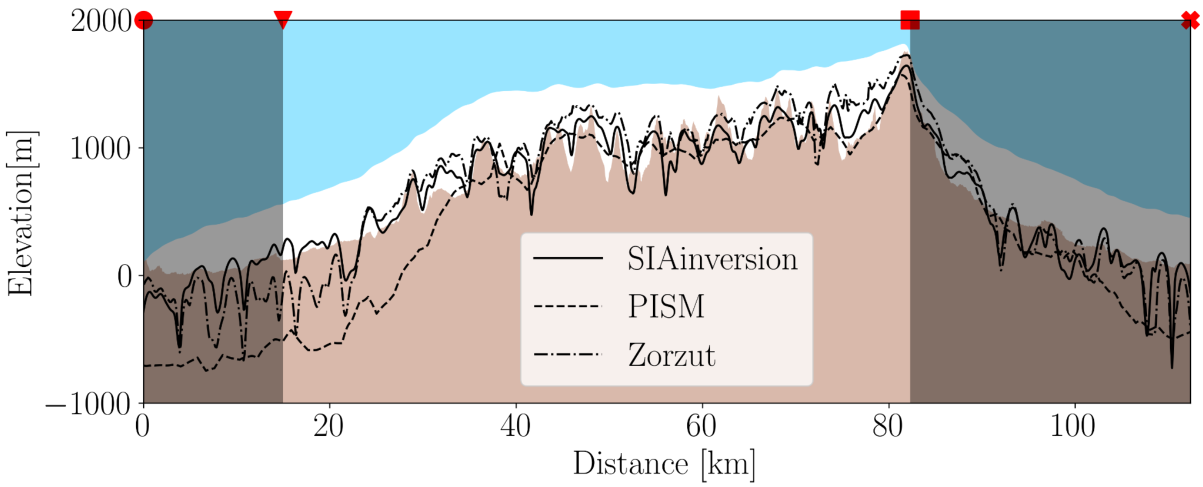
Changed.

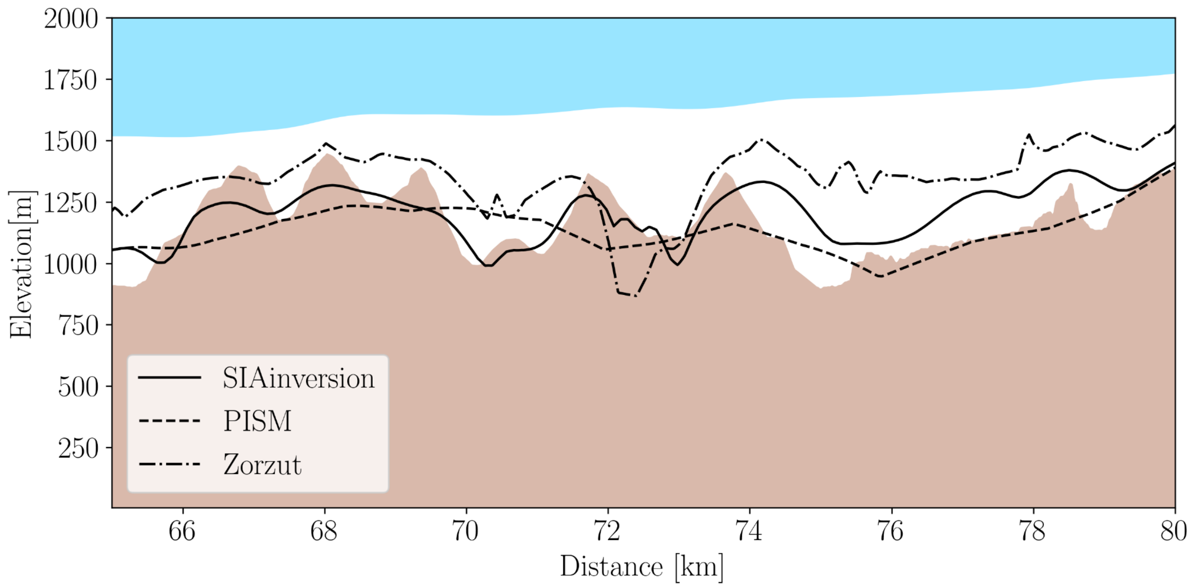
Section 2 (or Introduction): I am missing some information on ice velocities of the ice cap. I suppose these data could for example be extracted from online resources. A source like Its_Live (https://its-live.jpl.nasa.gov/) could potentially be useful. It would give an idea on potential sliding rates which is relevant to know because the non-sliding SIA is used in the SIA inversion, which may be a poor assumption of sliding is significant.

We have already tested a method using surface velocities in the cited MSc thesis. However, we found that this did not really improve the fit, and so we decided to exclude it from this manuscript to keep it short and simple (we used a custom velocity product that would require extensive documentation.)

As mentioned previously we will add the SIA method presented in Zorzut et al., 2020 which also includes sliding. But as you will see, this method also did not provide new insights or added value.

Again, we emphasize that our focus is the top of the ice cap where velocities are small. Small velocities result in a poor signal to noise ratio. This data quality issue severely hampers how well we can expect any method relying on surface velocities to perform. This is indeed what we find using the newly added Zorzut et al. 2020 method. At very high elevations we simply trust balance flux to be more informative than ice velocities.





L84: The slope threshold for ice thickness inversion is a critical parameter for SIA based inversions. The chosen value is however somewhat arbitrary which could be acknowledged as a source of uncertainty.

The slope threshold was chosen simply just to prevent any sinks in the surface elevation, as our method of calculating the balance flux cannot handle sinks in the surface. It is possible to find the optimal smoothing of the DEM (=resulting in the smallest misfit) using a brute force search given the very low computational cost of the semi-empirical model. We have not done that in this paper. We have, however, carefully chosen the gaussian smoothing with 2sigma=500m with exactly this in mind. This also implies that there are very few areas where the slope threshold is actually used. We will discuss our choice in more detail in the revised manuscript.

L88: ”why” --> ”which is why”

Changed.

L98-100: Under normal circumstances, the mass balance and its distribution in space should have a large impact on the thickness distribution. I suppose better fits could be achieved wíth a better spatial representation of mass balance.

The method is of course sensitive to the correct spatial pattern of Q. We have, however, compared the results of using a balance flux based on a constant mass balance against the HIRHAM SMB (MSc thesis). The ice thickness results were virtually identical (for the semi-empirical model). So, in practice this has little effect. Here it may be of interest to look at the mass balance vs elevation plot in figure 1.

L107: The PISM inversion is known to work best when a variable climate forcing is applied, i.e. when the modelled ice cap does not reach steady state. The fact that a fixed climate forcing is used here hence adds to uncertainty of the PISM inversion, which needs to be acknowledged.

We do not have sufficient info on past SMB to apply this here. But we will of course make sure to acknowledge this in the revised manuscript.

L118: See my third major comment above. Setting K to zero below 500 m a.s.l. effectively shuts down the inversion process in these areas. This is probably not desired.

It is correct that it shuts down the inversion process at low elevations. However, this is by design as explained above. We will be more careful about our PISM result interpretation and exclude the areas below 500 m from the comparison.

L122: ”combinations” --> ”parameter ranges”

Changed to: “the following possible combinations of parameter values”

L144: See also my first and second major comment above. In the SIA inversion the entire mass continuity equation, including ice dynamics and mass balance, is tuned through a, b and particularly k. In the PISM inversion, no tuning of for example the mass balance is done, only of a sliding coefficient and enhancement factor. That makes the comparison somewhat odd, since the SIA inversion is much more widely tuned to match the available thickness data.

We agree that this is important, and this makes the comparison unbalanced. However, as explained, computational cost limits how large a parameter space it is feasible to search. We will make sure to acknowledge in the updated manuscript that this comparison is not completely fair, and PISM would undoubtedly perform better if we also tuned SMB. We will also explain the computational cost issue.

L151-152: The larger RMSE for the PISM based method is not surprising. Based on Figure 2 this seems to be completely dominated by the errors below 500 m a.s.l. where the bed is not allowed to change in the PISM inversion. Above 500 m a.s.l. there does not seem to much difference between the two approaches (?). An additional mass balance tuning of the PISM method could help to make a better justified comparison of both methods.

We agree that the different models are not that different. This is actually very good news as this leads to agreement concerning which is the best drill site candidate. This is a major reason for estimating the thickness in multiple different ways. We realize that we did not make this point sufficiently clear, and will put an emphasis on this in the updated manuscript including a zoom-in of the cross section figure (Fig. 2) of the ice cap top.

L178-179: See also the third major comment above. I cannot imagine that the calving criterion is of much influence, since the bed is not allowed to change below 500 m a.s.l. (K=0), which means that a too large extent due to a lack of calving does not really play a role.

You are correct. The calving criterion mostly impacts the margin, whereas we are only concerned with the thickness at high elevation. By using K=0 for z<500m we save computational time as we do not have to include parameters for the calving approximation. However, the calving criterion still has a big impact on the time step chosen by PISM. We dont worry about misfit at the margin, and this allows us to pick a criterion that saves us computational cost.

L183: See also the second major comment above. There may be an overestimation of the SMB in HIRHAM5, but right now the overestimation of the ice thickness is a direct result of fixing the bed under these outlet glaciers to the Farinotti et al. initial bed. The apparent problems with the HIRHAM5 SMB data are exactly the reason why also in the PISM approach the mass balance should have been included in the tuning process, like is also done in the SIA inversion.

We agree that the low elevation misfit of PISM is not very informative. We intend to focus our misfit discussion to higher elevations. After all that is also our focus from a potential drill site perspective.

It is correct that tuning the HIRHAM SMB would have yielded better results at lower elevations. However, as mentioned before, it is not possible for us to test the full parameter space. Since our area of interest is around the top of the ice cap we do not worry too much about this.

L188-189: Again, the SIA inversion is not insensitive to mass balance, it just removes mass balance (and other) biases by tuning a, b and k. Similarly, mass balance biases could (and should) have been tuned also in the PISM inversion to enable a direct comparison.

We agree that it would have been nice to tune PISM SMB. It is computationally very expensive (as argued above). With this in mind we disagree that we “should” have tuned SMB too.

This limitation should of course be kept in mind in any comparisons. The comparison is not completely fair. We intend to further and more explicitly acknowledge this in the revised manuscript.

Regarding the SIA inversion (or semi-empirical ice thickness model) you are right that the SIA is not insensitive to the SMB, but the semi-empirical ice thickness model is, as mentioned previously.

Figure 3: ”UTM x” --> ”UTM Easting”; ”UTM y” --> ”UTM Northing”

Changed.