Response to Reviewer 1

Review for "Improved basal drag of the West Antarctic Ice Sheet from L-curve analysis of inverse models utilizing subglacial hydrology simulations" by Höyns et al.

The original reviewer comments are in black and our responses are highlighted with blue color.

Overall answer to RC1:

Dear Reviewer,

thank you for all of your remarks. Your suggestions will help us to further improve our manuscript! Even though we were encouraged not to create a revised version yet, it was easier to incorporate the changes already and facilitate the response to your comments, which is why we have already prepared a next manuscript version. We have tried to incorporate most of your comments into this revised version.
Both reviewers suggest major revisions before publishing the manuscript. In the future version we will try to reorganize the structure of the manuscript and make it far more efficient to read, especially for the method and results sections. As both reviews suggest to remove or condense the section on subglacial lakes we decided to remove it from the upcoming manuscript version. In addition, we will work on the presentation of our figures and in particular on the figures showing the L-curves and their cost function curvature. We will add a new part to the discussion section in which we want to discuss our results with regard to the regularized Coulomb friction law. Further, we will add a description of our selection method for the λbest, λmin and λmax of our L-curves.

In this manuscript, the authors analyse ensembles of basal drag inversions with a range of different values for their regularisation parameter, using six different variants of sliding laws; Weertman and Budd sliding with m=1 and m=3, with two different formulations of effective pressure within the Budd sliding law, Nop and NCUAS. For each sliding law, an L-curve is produced and the authors identify a "best" value for their regularisation, along with maximum and minimum acceptable values. They explore various methods by which to improve the appearance of their L-curves, and carry out a sub-domain L-curve analysis which reveals differences in optimal regularisation values for areas of the domain with different glaciological settings.

Comparisons are made between the outputs produced by the "best" regularisation values in each case, and also between the maximum and minimum values in the authors' preferred case of non-linear Budd sliding using NCUAS, for which differences in the structure of the output are discussed. Finally the authors attempt to validate their preferred basal friction field by drawing comparisons with previous studies of bed structures and potential subglacial lake positions.

General comments:

There are a lot of ideas presented in this manuscript, some of which are more convincing to me than others. The model setup and majority of the methodology is thoroughly explained, and entirely appropriate to the topics being addressed.

Thank you for your remarks!

The main highlight of this manuscript for me was the idea of sub-domain L-curve analysis. The idea that different regularisation may be needed in areas which contain fundamentally different physical settings, and therefore major differences in the ice flow, is an important one. I would have been interested in seeing far more focus on this aspect of the work. Comparisons in those sub-domains of the outputs using the locally-defined best regularisation against using the globally-defined value would have been illuminating, and I think this very promising avenue should be explored more in the future.

Yes, locally defined best values for regularization is a really interesting point and worth to be explored more in the future. We briefly comment on that in the detailed responses below.

Another highlight was the results displayed in Fig.15, showing very clearly why it is important to consider regularisation carefully, and the difference it can make in a model.

Thank you!

In general, I felt there was too much focus on the fine details of L-curve shapes without much explanation of why readers should care about this. The details of the L-curves and convergence are likely to be specific to ISSM, while other models have different inversion processes, and different regularisation parameters. Some models require choices for multiple parameters, which require multi-dimensional “L-surfaces”, and would require entirely different analysis. To my mind, the more interesting and useful comparisons for a wider audience outside of ISSM modellers were made when it was shown what the effect of different regularisation values was on the actual outputs of the inversion (ie. basal traction fields). I wasn’t convinced of the importance of how smooth the curves were from which the values came, or the statements being made regarding the comparative shapes of the curves.

We agree that our manuscript is relatively detailed and in some cases particularly important for ISSM users, such as the values specified for the convergence criteria. To show that those details are also suitable for other models one would need to prepare comparisons between those ice-sheet models and algorithms like Barnes et al., (2021) did. Nevertheless, the details of the L-curve analysis, whether on the shape or the regularization of our results, also apply to other models or regularization parameters. And it is true that our analysis is not based on multi-dimensional L-surfaces, but that was not the goal and is definitely beyond the scope of our work. In addition, Reviewer 2 argues that these details for ISSM are very helpful and also useful for the reproducibility of our work. Nevertheless, we try to shorten the manuscript in this direction and to highlight the main results here.
Regarding the comment on the smoothness of L-curves, we can argue that whenever our resulting L-curve had many outliers there was something wrong with the numerics or with the underlying model choices we did for those runs. So the smoothness of our L-curves gives us a kind of measure of whether everything in the background is working properly with the model or the numerics, just as the subdomain L-curve analysis does it in more detail. In our opinion, the comparative shapes of the curves for linear and non-linear sliding are conspicuous, as they show an explicit different behavior and the steeper L-curves for non-linear sliding helps

us to bracket the range from λmin to λmax smaller. Both aspects are also something observed by Wolovick et al. (2023).

I was also not particularly convinced by the comparison with proposed subglacial lake positions. Partly because it is not actually known whether lakes really do exist in these locations, but mostly that I do not see much meaningful correlation between the basal drag fields and the lake outlines in the figures.

We agree with the reviewers opinion on the lake candidates. Therefore, we will remove this section in the revised version, which was also recommended by Reviewer 2.

This manuscript appears to rely a lot on reference to Wolovick et al. (2023), but I think the most important parts of methodology should be restated here. In particular, the procedure for picking λbest, λmin and λmax, which are crucial to the results of this manuscript, is buried in an appendix in Wolovick et al. and should be stated clearly to help readers understand the L-curve figures.

We will add a description of the picking procedure of λbest, λmin and λmax to the new manuscript version.

Overall, I believe that this manuscript contains some very interesting and important work, but that it is somewhat hidden amongst a wide variety of ideas which could be organised in a better way, and some of which seem specific to the model being used. I think this manuscript could benefit from being edited down to a more concise form, and that some restructuring could benefit its readability.

Thank you for all the comments and suggestions. We will revise the manuscript in terms of structure and language, as well as shorten it to emphasize the important points, especially in the results section, as also recommended by Reviewer 2.

I would recommend major revisions before this manuscript is accepted for publication.

Specific comments:

Abstract

Lines 10-12: It’s a bit unclear what “best” and “worst” mean in terms of L-curve behaviour. Perhaps saying Pine Island produces the smoothest curves would be a better way of wording this?

We agree with your comment and have edited the wording in the revised version.

“Pine Island Glacier exhibits the smoothest curves, and the slow-flowing areas such as Roosevelt Island reveal rather poorly shaped L-curve behavior for the basal drag inversion.”

Lines 16-17: I don’t agree with using “more accurate” here. The choice of parameter, as the authors point out, is a balance between matching and regularising the observations. Accuracy could be interpreted as getting the closest fit to observations, which isn’t the point being made. The regularisation parameter itself cannot be defined as being accurate or not, as there is no physical real-world comparison to make.

Thank you for the remark, we have clarified the wording in the revised version.

“The analysis suggests that non-linear friction laws are preferable to linear sliding because of reduced variance of the overall inferred friction coefficient, faster convergence, as well as steeper L-curves leading to a more well-defined corner region.”

Line 18: “improved performance” in terms of what? Convergence, smoothness of L-curves? This should be specified.

We get an improved performance in terms of the total model variance ratio, faster convergence and smoother L-curves when NCUAS is included. We have specified this in the revised version.

“We show that a Budd-type friction law incorporating effective pressure from a subglacial hydrology model rather than a simple geometry-based approximation achieves improved performance in our inverse model in terms of total model variance ratio, as well as faster convergence and smoother L-curves. ”

Line 20: Should this be a comma rather than two separate sentences?

Yes, we have combined both sentences in the revised manuscript.

Section 1. Introduction

Lines 35-36: Could you expand on what exactly is meant by “the majority of high velocities are cause by sliding”, or provide a reference? Is this just saying that sliding causes higher velocities than flow driven by surface gradients, or is it claiming that changes to basal sliding have a larger influence than other processes such as mass balance?

We have intended to express that the fast velocities in the ice streams are mostly caused by basal sliding and have added a reference to the work of Engelhadt and Kamp (1998) as an example. The text now reads as:

“In the context of a better understanding of ice sheet processes, it is particularly important to examine the distribution of friction at the ice-bed interface, as this process has a major influence on the ice velocity, particularly in the fast-flowing areas (e.g., Engelhardt and Kamb, 1998). Since the distribution of friction underneath the ice sheets is difficult to observe directly, ice flow models are used to determine the basal drag.”

Line 43: I’d say “represent” rather than “compute”. The parameters from inversions broadly represent a combination of several factors (as is pointed out in the next paragraph), rather than being a value which physically describes one thing.

We have changed that in the revised version.

Lines 87-89: This reads as if the Budd law doesn’t account effective pressure. Should be reworded for clarity.

We agree, thanks for the remark. We have reworded the sentence.

“In the literature, instead of a Budd-type friction law (Budd et al., 1979), a Weertman friction law (Weertman, 1957) is often used (e.g. Morlighem et al., 2010, 2013; Joughin et al., 2004; Ranganathan et al., 2021) in which no effective pressure is taken into account.”

Lines 98-100: Isn’t the inverse problem in glaciological models always ill-posed, regardless of sliding laws or domains? Could this sentence be explain further?

Yes, that is true, but adding regularization to the minimization problem makes the problem of course less ill-posed. And the idea behind it is that we might have regions that are problematic to treat in the modeling, e.g., rock outcrops. If these regions have more outliers for small λ values in the L-curve, the problem may not be regularized well enough and still be ill-posed. But if it shows a good L-curve, we can assume that it is rather well-posed. Overall, we wanted to express that some areas are less ill-posed due to regularization and others may be more ill-posed because more regularization has to be used. However, we recognize that the spelling is misleading and remove this point from the text. Also in lines 292 and 356.

Section 2. Method

Some restructuring in this section could be beneficial, as there appears to be some repetition, and jumping between topics.

We agree, thank you for the remark. Restructuring of this part has been done and incorporates now also changes suggested by the other Reviewer.

Lines 108-115: I don’t think this section is necessary. ISSM can be introduced at the start of 2.1 instead, and other parts stated in the relevant sections.

That is true, we have deleted this paragraph and added the ISSM content into the section 2.1.

Lines 134-139: This could be moved into 2.3, with the rest of the inversion description.

We have moved this sentences into section 2.3.

Lines 166-176: The discussion of B fields seems to interrupt the description of the forward model. It could be better placed in the subglacial hydrology section as it is being discussed in relation to CUAS.

We have added this paragraph to the subglacial hydrology section.

Lines 189-194: Why not introduce both effective pressure parameterisations in the same section (ie. move this to the next section).

We have also moved this paragraph into the subglacial hydrology section.

Lines 233-236: These two explanatory sentences belong in the introduction (in fact, I think they restate something already in the introduction)

We have removed those sentences at this point and rearranged the regularization part in the introduction.

Line 243: Given how important the values of λbest, λmin and λmax are for this work, I think the method should be shown here rather than just a reference.

As mentioned before, we will add a description of the picking method for the L-curve analysis in the revised version.

Lines 276-280: Should the detail about the forward model solver be under the Forward Model section?

Yes, that is a good hint. We have added this sentences to the forward model section.

Section 3. Results

Lines 304-305: Why not just use a full range of [10-3 104] for all L-curves?

Yes, this would be a possible option, but we deliberately shifted the range of the L-curve further upwards using a linear sliding law in order to avoid small λ values below 10^-2 as we only obtained outlier models for such small λ values, which are not useful. This can be explained by the fact that these outlier models for small λ values reflect that a too small weighting of the regularization term increases the non-convexity of the inverse problem again. Since we are dealing with an ill-posed inverse problem adding regularization the problem gets more convex and with that easier to solve for the optimization algorithm (convex functions have only global minima, e.g. Rockafellar (1970)). But when using too little weight for the regularization term the problem is still non-convex (global and local minima), which make it again more difficult to solve.
On the other hand, we do not need more points above 10^3 for the L-curve using non-linear sliding laws, since the vertical λ-limb of all non-linear experiments are already very pronounced. However, we are particularly interested in the corner of the L-curve, i.e., extending the λ range upwards would only lead to more computing time and costs, which we would like to avoid. For the sake of simplicity, we have shifted the range for the linear sliding law upwards and kept the number of λ values the same, but still get pronounced limbs in the flat and vertical curve. Because of this, we argue that we do not need to run any further models as they would not give us any new information.

This was stated in the methods section in the manuscript and was probably disconnected from the results section. We have moved this part into the results section (see also RC2 suggestion).

Line 313: How are outliers identified? I assume these are points which lie over a certain distance from the smoothed tradeoff curve, in which case this should be specified. Or are they simply the inversions which struggled to reach convergence? In particular, for Fig. 8(d,g,h) it isn’t clear why some of the outliers shown shouldn’t be part of the curve.

We define inversion runs as outlier models if they are not fully converged. We further declare models as outlier if they are above a certain distance to the nearest data cost or regularization cost model by choosing a threshold value. Reviewer 2 also noticed this, so we have added a description of the definition of outliers into the revised version. Indeed, for Fig. 8(d,g,h) it is a bit unclear, we will check the algorithm again in these cases and make sure that we have used the same threshold for all experiments.

Lines 318-329: Was the same smoothing of kinit and the same values of εgttol and Δxmin applied for all inversions for all sliding laws? Given that these can have an important impact of the results I assume a like-for-like comparison has been conducted, but it could read as if different smoothing was used in different cases.

No, this is not the case, we had to smooth kinit further for the use of non-linear sliding in our inversions. For linear sliding we therefore use only one averaging iteration, but for the non-linear sliding law, we use of three iterations. We also chose different values for εgttol and Δxmin for linear and non-linear sliding (εgttol=10^-3 and Δxmin=10^-5 for the three linear sliding experiments and εgttol=10^-6 and Δxmin= 10^-4 for the three non-linear sliding experiments) to ensure convergence of the model.

We state this in the L-curve analysis part of the results section. Although a “like-for-like” comparison would be beneficial, this is not possible here, because a non-linear problem is inherently more difficult to solve than a linear one. Thus, more averaging iterations are used.

We have rephrased this as:
“We used nearest-neighbor averaging to smooth the initial drag coefficient kinit resulting from the driving stress (see Eq.(6)) for the runs with linear friction law. Unfortunately, further smoothing (three times nearest-neighbor averaging) was needed for the runs with non-linear friction to ensure convergence of all those runs. ... Admittedly, the mentioned adjustments, limit the comparability between simulations with linear and non-linear friction laws.”

Lines 335-342: As a more general point, is “best” the right word to use? The choice, as is always the case with L-curves, is quite subjective. I agree that the identified corner regions are a good guide for picking your regularisation parameter, and appreciate the more rigorous methodology behind picking a value rather than doing so by eye, but I would assume any pick within this region would be reasonable. Specifically, as noted here, in some cases the identified λbest is very close to λmin, and perhaps another argument could be made to say that a value halfway along the curve between λmax and λmax could be the best choice. Do you have any insight into what variability the choice within the corner region causes in the output? It would be interesting to see the differences between your λbest and a value one might pick by eye, to give more of an idea of how much this matters.

Perhaps it is not the best word, as it is of course not the best drag, because it is not the real one.

Of course any value between λmin and λmax is considered as reasonable and we state this in the text. For this reason, we have bracketed the range of acceptable λ values with λmin and λmax instead of specifying only λbest. You could also call it λoptimum (2^{nd} derivative of the curve), but of course that wouldn't be entirely obvious either. Based on the L-curve curvature (maximum curvature of the L-curve) an alternative name could be λmax_curvature, but we decided to stick with λbest according to the work by Wolovick et al. (2023).

Without any doubt, the selection of the picking method for the λbest value is still a “modelers choice” and we make it very clear that this is just our choice.

The range of acceptable λ values includes the choice of an arbitrary “heuristic” threshold for the curvature of the L-curve. We used a value of ½ of the maximum curvature to define λmin and λmax. In terms of our selection method for the λbest value in the range of λmin and λmax, this is the “best” value. But, as we have also mentioned in the manuscript, our picking method does not match in the case of Fig. 6 a and e (m=1, Weertman and m=3, Nop experiments) with the value one would choose from a visual point of view.
As we are aware of some limitations of our picking method, we would probably consider a new method, based on an integrated curvature, in future studies. However, since we particularly address Fig.6 f (or rather the corresponding experiment and analyze λbest, where the value of our method corresponds to the one that would be selected by eye), we would not like to lengthen the manuscript any further.
All in all, Fig. 15 shows what difference it makes to analyze λmin versus λbest versus λmax.

Lines 343-345: Is it necessarily to be expected that λ values would be close to each other? What were the differences in the other inversions mentioned here which display larger variability, and why should that make them less trustworthy?

Also Reviewer 2 has remarked this point and we have to admit that we have expressed ourselves incorrectly here. We cannot associate good results with the small variation of the λ best value ranges of all experiments. In addition, the subdomain L-curve analysis shows that the λ values do not necessarily have to be in the same range for different areas, as they require different amounts of regularization. We removed this point from the revised version.

Lines 346-354 & Fig.7: The convergence could be specific to the particular inversion process of ISSM with the chosen optimisation algorithm of M1QN3. Appendix B1 of Barnes et al. (2021) shows improved performance (greater minimisation of the cost function) achieved using an interior point algorithm (Byrd et al., 1999). Do you have any thoughts on whether you might find similar differences in convergence between your cases using a different algorithm such as this?

Yes, this could depend on ISSM and the M1QN3 algorithm and other algorithms might achieve better convergence here. However, we are relying on the implemented algorithms in ISSM. And as Morlighem et al. (2013) showed, a BFGS algorithm (quasi-Newton method), as used here, performs significantly better than for example a steepest-descent algorithm. And we also recognize that the M1QN3 algorithm achieves good convergence, but there are certainly improved algorithms which one could use. But of course we cannot say whether other algorithms, like the one shown in Barnes et al., (2021), perform better without further comparisons, which is beyond the scope of our work. But for sure, we can add a sentence to that in the revised version.
But we consider that it is beneficial to keep the convergence part in the manuscript. Reviewer 2 also found the details very useful and for the community using ISSM it can be helpful and increase reproducibility.

Line 355: I think this section should be 3.2. It contains enough to stand alone. I like the idea of running inversions on subdomains to find whether different regularisation could be required under different physical conditions. I think the values of λbest should be highlighted on Fig.8

as they are in Fig.6, to show the difference more clearly. I would personally also be very interested to see what difference would be seen in the inversion outputs going into a forward run, comparing a spatially-varying λbest to the global value.

We have renamed this section into section 3.2 even if we have moved some of the results from this section to the non-linear versus linear sliding section after being recommended by Reviewer 2. In addition, we will add the λbest values to the Fig.8.
It is not clear what is meant by simulations based on a “spatially-varying λbest”. The regional L-curves are based on the same inversion result (experiment NCUAS, m=3). It is just another post-processing step to analyze different geomorphological settings. One could do forward runs using the different λbest values and compare those runs with the “global” λbest but this is out of scope for this manuscript. As shown already in Fig. 15, higher λ values would result in more smoothing (more regularization).

If the Reviewer thinks of a setup with spatially varying amounts of regularization (λbest) based on, e.g., information from the regional L-curves for different geographical settings, this would be a very interesting idea. Unfortunately, we don’t see a practical way on creating such a map, yet. (Classification by ice speed as well as surface, bed and thickness slopes?)

Fig 8: This figure, and some that follow, could be made clearer by highlighting the corner region in a different colour, rather than the thicker line currently used. I assume the circles with white in the middle are outliers, but this is not labelled on the figure. Some of the L-curves presented in this figure could benefit from an extended range of values, as the corners appear very close to the end of the intervals used.

Thank you for pointing this out. We will edit the corner regions of the L-curves accordingly in the revised version.
Correct, the white circles show outlier models and we added a sentence about these in the figure caption.
We understand the reviewers point (extending the ranges for the L-curves), but here we argue again that an extended range is not associated with new information, rather only with increased computational costs. In particular, this is only the case for Fig.8 d,e,i. Those subdomains do not benefit much from the regularization.

Figs 9,10,12: The diamonds showing location of λbest is not clear. The values of λbest could also be shown.

In the revised version, we will adjust the diamonds and add λbest values in these figures.

Lines 464-471: I’m not sure I follow the reasoning for using the same λbest in both cases. The whole point in the inversion process is surely to pick an optimal regularisation for each case, in order that the output is most suited to the individual problem. I would think it was more instructive to compare the outputs that come from each experiment’s λbest value. Also, the numbers given do not correspond to the λbest values stated on Fig.6. If these are coming from different L-curves it should be made clearer.

Yes, the point of the L-curve analysis in the inversion process is to pick an optimal regularization for each case.
However, the point we intended to raise here is how Nop versus NCUAS in a Budd sliding law

account for the structure of k^2, since we are discussing k^2_W=k^2_BN. If we would use different λbest values in this case, one result might be smoother than the other and no longer comparable. The goal is to show that NCUAS captures significantly more of the structure of k^2B than Nop, which for a Weertman sliding law is simply taken into account by k^2W. In the ideal example or ideal effective pressure N, k^2B would be a constant.

The selected λ values used in the figure are the nearest (discrete) samples on the L-curve. We don't have a simulation performed with the estimated λbest values at this point. See section "3.4 Best drag estimate" for the simulation using the exact λbest value. We have clarified this in the revised version.

Fig.11: Why not also show the Weertman output?

Yes, we could of course also show the Weertman output, but the aim here was to show what effect the use of an effective pressure N (NCUAS or Nop) has on the basal drag coefficient result k^2. In the case of the Weertman law, a basal drag coefficient k^2 would account for every structure that we otherwise put into the effective pressure field for the Budd friction law. Therefore, we argue that a map of k^2 for Weertman sliding provides no further insights.

Lines 519-520: Could you explain how the chosen case is the "best estimate" of those on Fig.6?

We have added a description in the revised version of why this chosen λbest is the best estimate here, as also Reviewer 2 has noted this.

We now state:
"For this particular λbest, the cost curvature (Fig.9 d) corresponding to the L-curve in Fig.6 f (N_CUAS, m=3) has a clear peak at which our picking method chooses the λbest value. This would also be the position one would pick the λbest value based on visual inspection of the L-curve. "

Fig.15: This is a great figure for showing why regularisation choices are important, and the difference they can make!

Thank you!

Section 4. Discussion

Line 558: It may be desirable from a practical perspective, but for better results maybe picking a few different regularisation values based on ice speed or other obvious differences in the physical setting would be the way to go. I find this idea very interesting.

Yes, this is a very interesting point, as mentioned in a comment above, and definitely something to think about in the future, but this is outside the scope of this manuscript at the moment.

Figs 17,18: I don't personally find these to be particularly convincing. While there are a couple of places where the basal drag field lines up with the white lines, there are also places where the outlines cross areas of higher friction. What are the locations of these “possible” lakes based on? Is there a correlation with their positions in the effective pressure fields Nop and NCUAS? Is similar correlation seen for the Weertman inversion output as the Budd ones?

There is no particular correlation seen in the effective pressure field Nop or NCUAS with the lakes and it also does not make much sense to compare the lakes with the effective pressures, as these are not explicitly included in the determination of these fields. Also for the basal drag inversion output using the Weertman sliding law we can find similar correlations with the possible lake candidates. But, we realize that the analysis does not contribute much more to the output of the manuscript and Reviewer 2 also noted this point. Therefore, we have decided to shorten the manuscript at this point and to remove the comparison of the basal drag and the possible lake candidates from the revised version.

References

Barnes, J.M., Dias dos Santos, T., Goldberg, D., Gudmundsson, G.H., Morlighem, M. and De Rydt, J., 2021. The transferability of adjoint inversion products between different ice flow models. *The Cryosphere*, *15*(4), pp.1975-2000.

Byrd, R.H., Hribar, M.E. and Nocedal, J., 1999. An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, *9*(4), pp.877-900.

Wolovick, M., Humbert, A., Kleiner, T. and Rückamp, M., 2023. Regularization and L-curves in ice sheet inverse models: a case study in the Filchner–Ronne catchment. *The Cryosphere*, *17*(12), pp.5027-5060.

Our References

Barnes, J.M., Dias dos Santos, T., Goldberg, D., Gudmundsson, G.H., Morlighem, M. and De Rydt, J., The transferability of adjoint inversion products between different ice flow models, *The Cryosphere*, *15*(4), pp.1975-2000, 2021.

Engelhardt, H. and Kamb, B.: Basal sliding of Ice Stream B, West Antarctica, Journal of Glaciology, 44, 223–230, 875, https://doi.org/10.3189/S0022143000002562, 1998.

Morlighem, M., H. Seroussi, E. Larour, and E. Rignot, Inversion of basal friction in Antarctica using exact and incomplete adjoints of a higher-order model, *J. Geophys. Res. Earth Surf.*, 118, 1746-1753doi:10.1002/jgrf.20125, 2013.

Rockafellar, R.T. Convex Analysis. Princeton University Press. http://www.jstor.org/stable/j.ctt14bs1ff, 1970.

Wolovick, M., Humbert, A., Kleiner, T., and Rückamp, M.: Regularization and L-curves in ice sheet inverse models: a case study in the Filchner–Ronne catchment, The Cryosphere, 17, 5027–5060, https://doi.org/10.5194/tc-17-5027-2023, 2023.