The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2019-156-AC1, 2019

The Cryosphere Discussions

Interactive comment on "Spatial probabilistic calibration of a high-resolution Amundsen Sea Embayment ice-sheet model with satellite altimeter data" *by* Andreas Wernecke et al.

Andreas Wernecke et al.

andreas.wernecke@open.ac.uk

Received and published: 5 November 2019

Dear Olaf Eisen, Thank you for giving us the opportunity to outline how we plan to address the referees' concerns.

Firstly, concerning the lack of a 'perfect-model' or similar benchmark of our calibration: We have now worked on this problem and will present our results, in which we employed a model test for which we select different members of the model ensemble (one at a time) and treat that respective member as observations. We can then see whether our methods are capable of finding, or at least incorporating, the correct model parameter values. To make this test more rigorous we add an error to the selected model run before calibration to see how the calibration responds to imperfections. We used different methods to generate said error, including homogeneous independent noise, independent noise with spatially varying variance equal to the variance we find in the real observations (which is what we use below), spatially correlated noise and using the real observations with the mean field removed (so that the mean can be replaced by the model run). We have not yet decided on the exact final form of this test (and are open to suggestions) but the findings in all cases are fairly similar.

We find that our current approach is not capable of robustly constraining the sliding law exponent or ocean melt forcing while the results regarding basal traction, viscosity and bedrock maps are supported by the benchmark test (see end of document). We will back this up with the emulator information criterion (indicating how much of a signal can be found for each parameter). We plan to address this issue in the manuscript by removing the ocean melt forcing and sliding law exponent from the calibration and instead using the uniform prior for the ocean melt and selecting nonlinear sliding by expert judgement.

This finding, and hence the decision not to calibrate sliding law and melt forcing, is supported by physical considerations: Due to the inversion of basal traction fields for each sliding law, the initial (t=0) model velocities are (nearly) the same. Only after those velocities change, the impact of different sliding laws becomes apparent. A change in ocean melt affects the ice dynamics in a commutative way: only once the ice shelf thickness has significantly changed, the ice dynamics upstream are affected. A change in bedrock, basal traction or viscosity have however a much more immediate effect on the ice dynamics and might therefore dominate the calibration on short time scales.

We believe this will also resolve the concerns of Reviewer #1 about the apparent preference of linear sliding.

Another concern is related to the model runs being interpreted as predictions even though no climate forcing is used explicitly. We will address this in the following way:

(1): expand the current discussion (in which we argue that climate scenarios are expected to have limited impact on ice sheets within 50 years) by highlighting the deep uncertainty in deriving local ocean melt forcings from global climate models as well as the fact that the ensemble encompasses this uncertainty by perturbing the melt forcing from halving to doubling; even though there is no explicit climate representation we do cover at least much of the uncertainty related to variability in ocean melt. This interpretation is further aided by the new decision not to calibrate the ocean melt with current observations. (2): we will avoid using the word 'predictions' but only 'projections' as the model is projecting the current state of the ice sheet into the future on the basis of a clearly stated set of assumptions. (3): We will move the focus of the study away from those projections and towards the development and validation of the methods in addition to further comparisons with more common methods (as requested by Reviewer #2).

To increase transparency (and concerns about the code being correct) we will upload the relevant Python code to a public repository.

Reviewer #1 further raised concerns about the absolute match/distance between observations and model runs as compared to the possibility of selecting the best among bad setups. We will address this by performing an initial history matching (ruling out implausible parts of the input space) and restrict the probabilistic calibration to those areas which are not ruled out. Preliminary results find that more than 10% of the input space and most of the likelihood distribution cannot be ruled out by a 99.5% probability interval (using satellite observations, not the model test). This reassures us that the chosen basis representation is adequate for calibration and that we did not underestimate the uncertainties (observational, systematic and emulator combined).

We agree that scaling an optimized input field, as has been done for the dataset we use here, is inferior to fully exploring the ice sheet response to more flexible, higher dimensional variations to the input fields. However, computational and methodological challenges make simple scaling approaches more feasible. The focus of this manuscript is on the spatial calibration and how to further constrain an existing model ensemble, not how to improve the initial design of ensemble experiments. We hope to address this concern by re-framing this study towards calibration validation and comparison with simpler approaches so that the ice sheet model data is seen more as a test case instead of predictions.

We are happy to investigate the sensitivity to the calibration period and clarify that Nias et al. (2016) optimized the basal traction coefficient for both bedrock topographies and sliding laws separately.

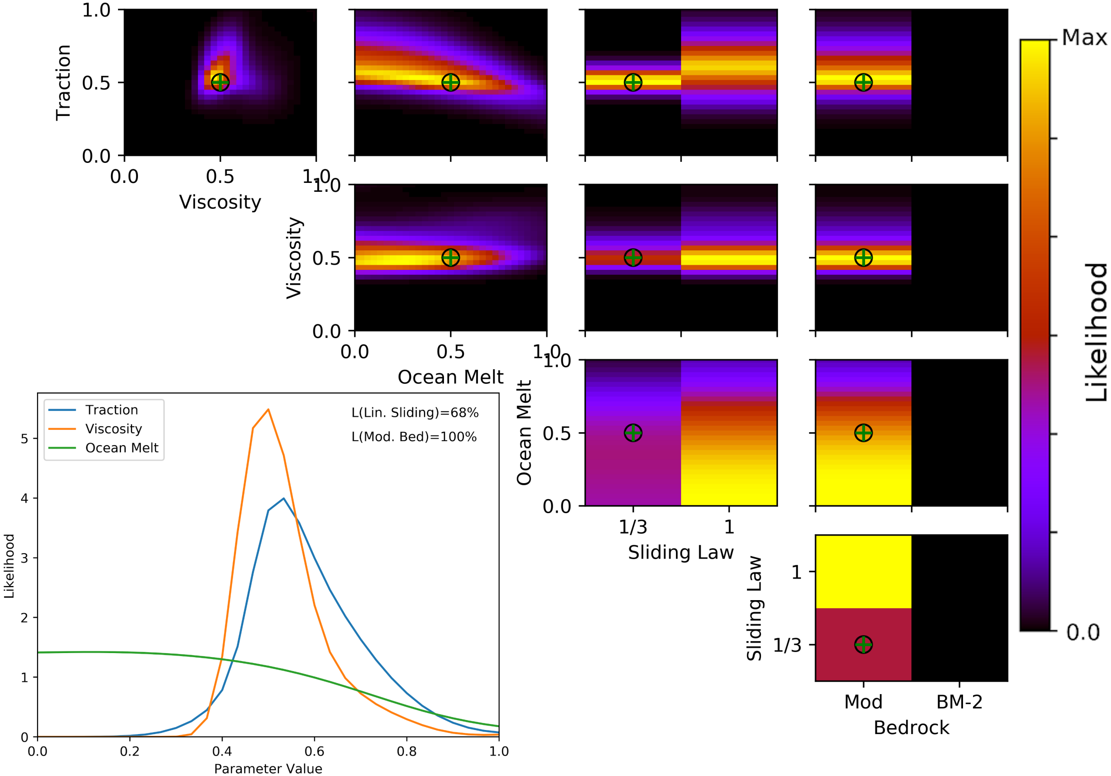
A: Synthetic experiment testing Here we present the results of our preliminary methods test, as introduced in the beginning of this document. At each location we calculate the variance from the real observations through the same period as used in the manuscript. The noise is defined here as spatially independent, zero-mean, normally distributed, random noise with before mentioned variance. For each selected model run (four of which are presented below) we generate 14 noise fields and add them to the single model dh/dt field and these 14 realizations are used in exactly the same way as described in the manuscript for the 14 periods of satellite observations. For comparison we added two markers to the likelihood figures: A black circle which represents the parameter values of the model run (i.e. the real values/target for calibration) and a green cross, representing the parameter values of the setup with maximal likelihood. Note that the green cross would always coincide with the top of the color scale (yellow) in five dimensional plots. Due to marginalization (summation) of three of the five dimensions (as done here for illustration) this is not true in two dimensions. In other words, the sum over many sub-optimal values can be larger than the maximum likelihood.

Caption for all figures: Likelihood of parameter combinations (evaluations of Equation 10 of the manuscript). Upper right panels show likelihood values marginalized to pairs of parameters, normalized to the respective maximum for clarity. Lower left panel shows likelihood values marginalized to individual parameters for the three scalar parameters (line plots), and sliding law and bedrock topography map (text and quotation within), normalized to an integral of one, consistent with Probability Density Functions. Differences between the four plots stem from the use of different model runs used for the test (in each case the central values for traction, viscosity and ocean melt are used, but bedrock and sliding law values are varied). The parameter values of each example are shown by the black circles, while the values of the set of parameters with highest likelihood are shown by green crosses.

As can be seen from the figures, marginal likelihoods of our calibration approach favour linear sliding in all shown test cases, including those from nonlinear runs. In addition, the ocean melt parameter is virtually unconstrained by the calibration (flat green lines in the line plots) or it is, as in the case of nonlinear sliding and modified bedrock, biased towards small melt factors.

In contrast, the basal traction coefficient and viscosity scaling factors have a strong mode and are always centered at, or close to, the real value of 0.5. The bedrock map is always clearly identified. Different values of basal traction and viscosity have been tested and show similar performance (not shown). The fact that the parameter setup used for the respective test is in all four cases attributed the maximal likelihood (green cross=black circle) supports our confidence in the implementation as the real parameter set is identified correctly as best fit. Relative ambiguity with respect to sliding law and ocean melt simply overrules this finding for marginalized likelihoods.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-156, 2019.



1.0
0.5
0.0
Traction
0.0
0.5
1.0
Viscosity
1.0
0.5
0.0
Viscosity
0.0
0.5
1.0
Ocean Melt
1.0
0.5
0.0
Ocean Melt
1/3
1
Sliding Law
1
1/3
Sliding Law
Mod
BM-2
Bedrock
Max
Likelihood
0.0
Traction
Viscosity
Ocean Melt
L(Lin. Sliding)=68%
L(Mod. Bed)=100%
Likelihood
Parameter Value

Fig. 1.

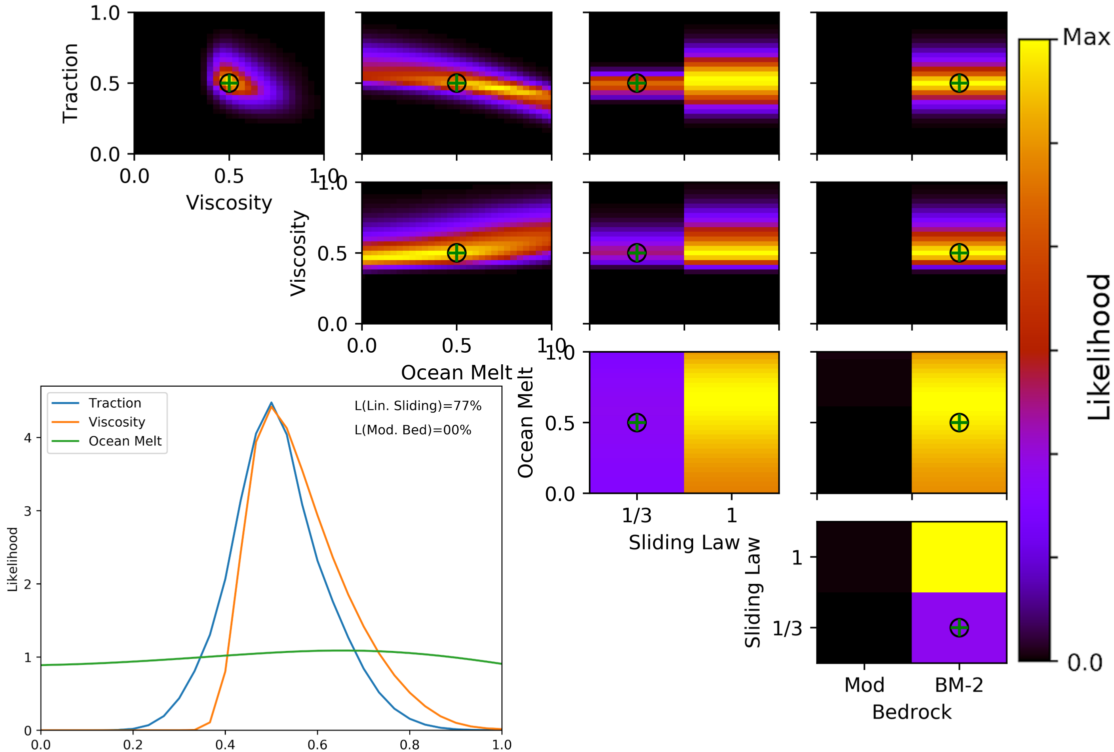
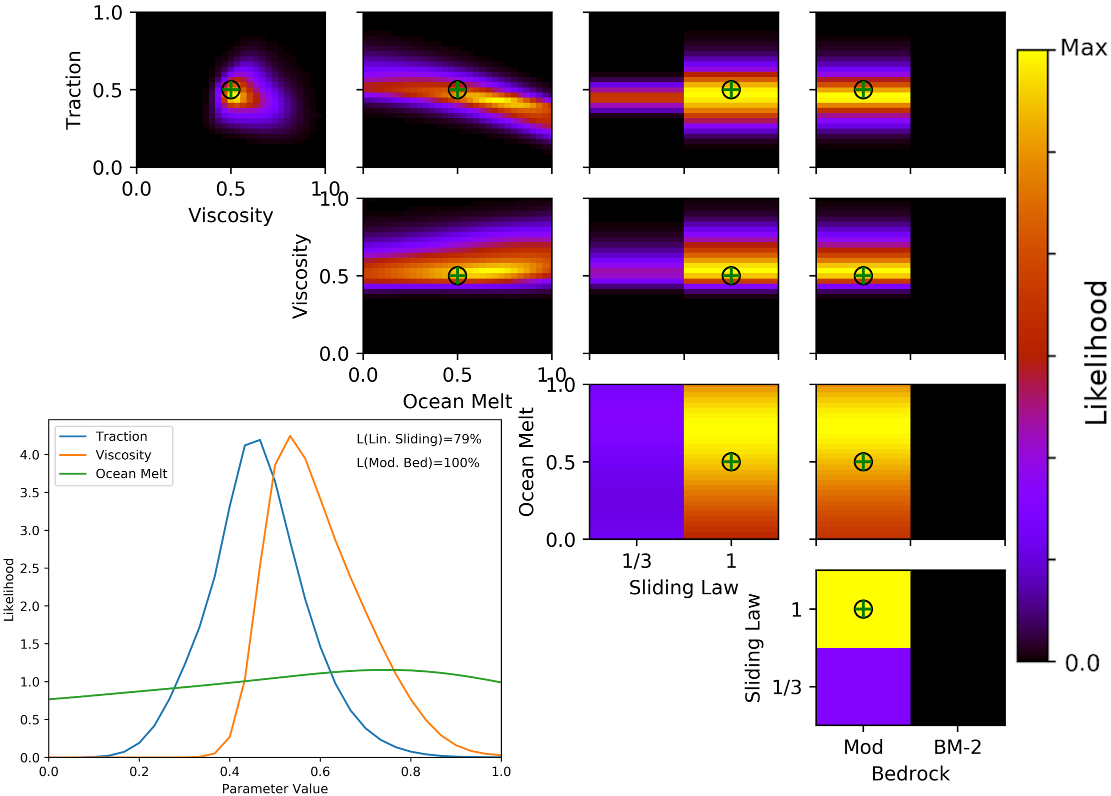


Fig. 2.

Interactive comment

Printer-friendly version

Discussion paper



1.0
0.5
0.0
Traction
0.0
0.5
1.0
Viscosity
Ocean Melt
Sliding Law
1/3
1
Bedrock
Mod
BM-2
Max
Likelihood
Traction
Viscosity
Ocean Melt
L(Lin. Sliding)=79%
L(Mod. Bed)=100%
Parameter Value

Fig. 3.

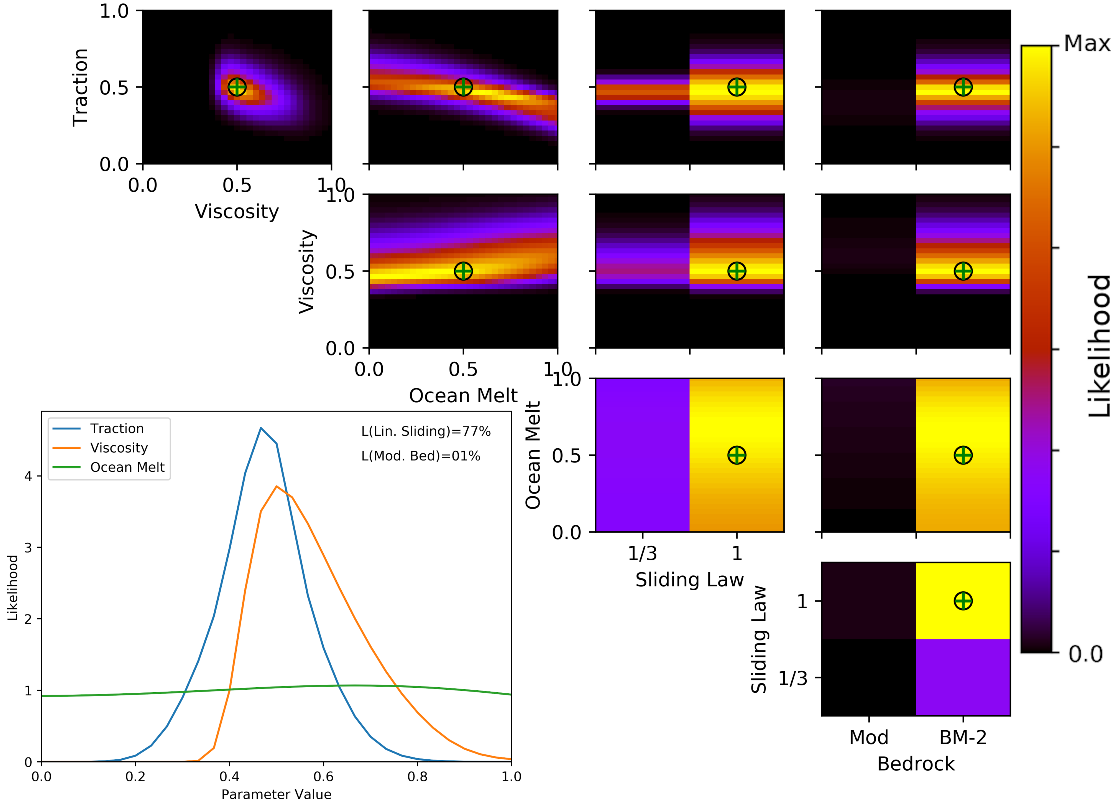

Traction
Viscosity
Ocean Melt
Sliding Law
Bedrock
Likelihood
Max
0.0
Mod
BM-2
1/3
1
Traction
Viscosity
Ocean Melt
L(Lin. Sliding)=77%
L(Mod. Bed)=01%
Parameter Value

Fig. 4.