Comment on 'Spatial probabilistic calibration of a high-resolution Amundsen Sea Embayment ice-sheet model with satellite altimeter data'

presented by Wernecke et al.

In the manuscript, Wernecke et al., present a promising method to calibrate uncertainty distributions of mass loss derived from ice-sheet model simulations with spatial data. Their approach consists of a dimensional reducion by using the representation of the model simulation output in its corresponding principal component basis. The ensemble is then statistically emulated and calibrated in the principal component basis. This procedure is applied to an ensemble of simulations of the Amundsen Sea region published in (Nias et al., 2016). The approach presented is potentially of great value for ice-sheet modelling studies that aim to make sea-level projections. Before considering it for publication, I recommend additional analyses, a more detailed discussion of the capabilities and limitations of the method and reframing as explained in the comments below.

Major comments:

- p.1 l.9, l.11, & other: with some more analysis, this study can make a very good test case that demonstates the capabilities of the new method. However, it is problematic to say that in this study you are estimating future sea-level contribution or that you are making 'predictions' or 'projections', since your analysis is based on simulations with constant ocean forcing, excluding for example natural variability (e.g., Jenkins et al., 2016) or potential future changes in ambient oceanic and atmospheric conditions (e.g., Holland et al., 2019) depending on the different socio-economic pathways (RCP scenarios). Possible future evolution of surface mass balance is not considered and uncertainty in basal melting is based on a simple amplitude scaling, neglecting for instance the effect of changes in spatial melt rate distributions (discussed, e.g., in Goldberg et al., 2019) or uncertainties related to the basal melt rate parameterisation (see, e.g., Favier et al., 2019).

- p.5 l.11: the choice of calibration of dh/dt after running the model for 7 years appears random. Please explain this. Also, how would your results be influenced if your calibration was done after $1, 5$ or 10 years?

- p.12 l.3: my understanding of Nias et al. (2016) is, that inversion techniques were used to estimate the spatial fields of viscosity and basal traction coefficients. Were different inversions run for the different bed geometries and values of m? If the inversion was run only for $m = 1$, a better fit for $m = 1$ in comparison to $m = 1/3$ would not be a surprise as the parameter fields were optimized for this case. If this is true, your findings are maybe more due to the experimental design rather than being physically interpretable. Please

clarify this (similar for the bed topography and the other parameters) and, if applicable, consider it in the discussion of your findings.

- p.14 l.24-27 and Appendix B: you state that your method improves calibration with aggregated variables. It is interesting to see the effect on the different parameters (Figure B1), but to make this point clear, please add also the effect on the mass loss and grounding line probability estimates (similar to Figures 5,6).

Further comments:

- page 2 lines 22ff: there are a number of modelling studies with coarser resultion that do not require a parameterized grounding line for retreat (e.g., Schlegel et al., 2018). 'Regional' is maybe more appropriate than 'one glacier' (e.g. Arthern and Williams, 2017).
- p.2 l.28 and l.20: please check your use of 'predicted' versus 'projected'.
- p.3 l.23-29: emulation of model output was also used for example in Levermann et al. (2014).
- p.4 section 2.1: since basal melt is the driver of mass loss in the Amundsen Sea at present, more details should be given here, e.g., how do mass fluxes compare to observations?
- p.5 l.13: you could state here that your $y(\theta_i)$ is $\frac{dh}{dt}$.
- p.5 l.16: $\Theta = [0,1]^5 \subseteq \mathbb{R}^d$?
- p.5 l.21: shouldn't $S \in \mathbb{R}^{m\times n}, U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$, since U, V are unitary matrices and by definition quadratic? Please check also the other matrix dimensions.
- Section 3.1: a reference to Fig. 1 is missing.
- Figure 1: please give here more explanation, e.g., of 'unit length'.
- p.6, l.8: would it be an option to calibrate not only after 7 years but at all datasets from Konrad et al. (2017) individually as they find variations in the onset and propagation of surface lowering?
- Figure 2: in your reprojection of the mean observation, artifacts of thickening occur. How will this affect your calibration?
- p.7 l.1: a value of 0.6 seems to be rather large, please explain.
- p.7 l. 5: I cannot find where this is discussed in the results section?
- p.7 l.7: how is the training done? please give more details here.
- p.7 l.7: you could help the reader if you explain what the rows of $S'T'^T$ represent.
- p.7 l.12: I cannot find the definition of a Gaussian Process Emulator in the given reference.
- p.7 l.15ff: more details are needed here.
- p.8 l.16: eqn.

- Section 3.4: you are switching between observational errors and model errors in this section. It might be easier to read if you give and explain one by one.
- p.10 l.11: prediction, see above
- p.15 l. 28: 'the' too much
- p.16 l. 4: please specify 'uniform within the parameter space'.
- Figure A2: how are the quantities shown on the x and y axis obtained?
- Appendix B: It would be great to see also how your method compares to calibrations using a spatially aggregated, temporal evolution of mass loss as used for example for targeted parameter optimization in Golledge et al. (2019).

References

Arthern, R. J. and Williams, C. R. (2017). The sensitivity of west antarctica to the submarine melting feedback. *Geophysical Research Letters*, 44(5):2352–2359.

Favier, L., Jourdain, N. C., Jenkins, A., Merino, N., Durand, G., Gagliardini, O., Gillet-Chaulet, F., and Mathiot, P. (2019). Assessment of sub-shelf melting parameterisations using the ocean–ice-sheet coupled model nemo (v3. 6)–elmer/ice (v8. 3). *Geoscientific Model Development*, 12(6):2255–2283.

Goldberg, D., Gourmelen, N., Kimura, S., Millan, R., and Snow, K. (2019). How accurately should we model ice shelf melt rates? *Geophysical Research Letters*, 46(1):189–199.

Golledge, N. R., Keller, E. D., Gomez, N., Naughten, K. A., Bernales, J., Trusel, L. D., and Edwards, T. L. (2019). Global environmental consequences of twenty-first-century ice-sheet melt. *Nature*, 566(7742):65.

Holland, P. R., Bracegirdle, T. J., Dutrieux, P., Jenkins, A., and Steig, E. J. (2019). Climate forcing of the west antarctic ice sheet: Anthropogenic trends and internal climate variability. *Nature Geoscience*.

Jenkins, A., Dutrieux, P., Jacobs, S., Steig, E. J., Gudmundsson, G. H., Smith, J., and Heywood, K. J. (2016). Decadal ocean forcing and antarctic ice sheet response: Lessons from the amundsen sea. *Oceanography*, 29(4):106–117.

Konrad, H., Gilbert, L., Cornford, S. L., Payne, A., Hogg, A., Muir, A., and Shepherd, A. (2017). Uneven onset and pace of ice-dynamical imbalance in the amundsen sea embayment, west antarctica. *Geophysical Research Letters*, 44(2):910–918.

Levermann, A., Winkelmann, R., Nowicki, S., Fastook, J. L., Frieler, K., Greve, R., Hellmer, H. H., Martin, M. A., Meinshausen, M., Mengel, M., et al. (2014). Projecting antarctic ice discharge using response functions from searise ice-sheet models. *Earth System Dynamics*, 5(2):271–293.

Nias, I. J., Cornford, S. L., and Payne, A. J. (2016). Contrasting the modelled sensitivity of the amundsen sea embayment ice streams. *Journal of Glaciology*, 62(233):552–562.

Schlegel, N.-J., Seroussi, H., Schodlok, M. P., Larour, E. Y., Boening, C., Limonadi, D., Watkins, M. M., Morlighem, M., and Broeke, M. R. (2018). Exploration of antarctic ice sheet 100-year contribution to sea level rise and associated model uncertainties using the issm framework. *The Cryosphere*, 12(11):3511–3534.