Review of ‘**Estimating snow depth on Arctic sea ice based on reanalysis reconstruction and particle filter assimilation’** by Li et al.,

Review by Alek Petty

Summary

This study presents estimates of snow depth on Arctic sea ice from an updated version of the NASA Eulerian Snow on Sea Ice Model (NESOSIM) and a particle filter data assimilation scheme to combine the model estimates with satellite-derived snow depth data (RA-5VLSTM). The results were compared primarily with snow depths collected by NASA’s Operation IceBridge and also some more limited Ice Mass Balance, and a MOSAiC snow depth buoys. The results were also compared with a Kilic et al., (2019) snow depth dataset produced from regression of IMBs to passive microwave data, and also the modified version of the Warren climatology.

General comments

In general, I think the approach of this study was good – use a new data assimilation approach to constrain NESOSIM output and potentially improve its ability to simulate snow depth on Arctic sea ice. However, I have a number of concerns about this study which I detail here:

1. NESOSIM is an open-source model (https://github.com/akpetty/NESOSIM) so community development is actively encouraged - e.g. adding new parameterizations, data assimilation modules etc., into the official code base. The framing of an Improved NESOSIM was thus slightly odd, although obviously this could also be a language/communication issue. The ‘Improved’ nature of this model framework was also somewhat underwhelming. The atmosphere loss term included as 1 of only 2 ‘improvements’ in this version of NESOSIM has already been integrated into NESOSIM (v.1.1, https://github.com/akpetty/NESOSIM/releases/tag/v1.1). The authors made a note of this term being introduced already but made no link to the official code repo and still included it in your own ‘improved’ version. This means the only new parameterization introduced here (to generate the Improved NESOSIM framework) was the simple degree day temperature/melt parameterization. I think this parameter inclusion makes broad sense (we’ve considered something along these lines ourselves) but i) it was not actually clear that this specific parameterization helped improve the simulation of snow depth as most of the validation occurred in winter/spring and ii) this could have been communicated as a simple added parameter to NESOSIM. I think the atmosphere loss term was much more significant and we’ve found this to be a useful additional tuning factor, although one not well constrained by observations. Indeed most of what this study is doing is bias correcting towards the OIB quick-look data. On that note, I didn’t see any information about making the code available (e.g. the degree day melt model or the particle assimilation approach) which was surprising considering the authors utilized extensively an open-source model for much of this work.

2. A big issue is that quick-look OIB snow depths are used as truth, with bias corrections/model calibration carried out to improve the fit to this dataset, essentially. However, deriving snow depths from Snow Radar data collected by OIB is challenging (Kwok et al., 2017,) and wide differences exist across the different products. We make a big point about this in the original NESOSIM paper (Petty et al., 2018, P2018). More recent research has shown that OIB QL is ~5 cm thinner than the consensus from the three 'final' products analyzed in P2018 (Petty et al., in prep), see preliminary figure below. These are (since 2013) quick-look data, supposed to provide a basic overview of sea ice conditions, not really a reliable dataset for validating models/retrievals.

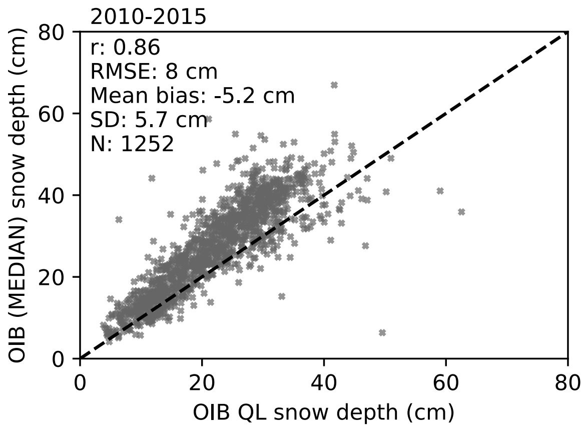

Figure 1: Comparison of the median snow depth from the three different OIB snow depth products used in Petty et al., (2018) and the quick-look (QL) OIB snow depth data. Data are gridded to a 100 km polar stereographic domain before the comparison.

There were also plenty of other parts of the study where data uncertainties are vaguely described and, in some cases, described with worrying levels of certainty ('The satellite-derived snow depth contains an uncertainty of 1 cm,').

3. I was hoping this paper would provide a much deeper explanation and insight into particle filter data assimilation, but the paper provided only really a minimal description of this. In no way is the approach reproducible. It also left me feeling unsure how much the authors understood about the approach and how best to implement this. The particle number sensitivity test did not feel satisfactory.

4. The RA-5VLSTM dataset was used as the only input to the data assimilation system but the citation linked to is just a data portal that I was unable to translate, so really there is no background to how this data was obtained and how well it agrees with other snow depth datasets that exist. My guess is that the INESOSIM-PF run tracks this observational dataset quite closely, but it's unclear if that's a good thing or not.

Specific comments

The statistics of RMSE and MAE include the bias – so really all the statistics presented are highly sensitive to the presence of a bias. Most of this study seemed to involve basic bias correction (which is somewhat understandable considering the large uncertainties in snow) but limits the impact of the results presented. Generally I think it is not a good idea to express RMSE/bias changes as percentages. Just stating the change in absolute terms is easier for the reader to assess.

L73-74: this particle filter methodology and motivation needs to be much better described.

'Section 3.2 Two snow depth retrieval methods' – why are these not in the data section? They are previous data not really created in this study - one 'retrieval' - the multi-linear regression to passive microwave data from Kilic et al., (2019) and then the Warren 1999 (W99) quadratic fit to in-situ snow depths.

It is also confusing that you use the W99 with snow depths halved over FYI as well, and refer to this as an NSIDC product (taken from the CryoSat-2 implementation of this). This was also used in P2018 and is typically referred to as the modified Warren climatology. This was referenced in P2018 (and I think was first introduced by Laxon et al., 2013). I don't think it should be referred to as an NSIDC product particularly.

L201: 'However, the updated algorithm has not been debugged' is a bit of a strange way of framing this. The code is on GitHub (version 1.1) so you should ideally cite that more clearly, as it is exactly the same as the 'improved' atmosphere snow loss term used in this study.

L240 – 'Therefore, the model was insensitive to β.' – really, it's just insensitive in the regions where we have observations (e.g. in the central Arctic). P2018 showed how in more marginal seas it has a bigger impact.

L243 – now a bit confused regarding the parameter you're looking at here. I think it's beta but why choose that if NESOSIM is *less* sensitive to this parameter?

Section 4.1.1. – the problem here is that you're fitting to quick-look OIB now depths that are likely biased. To accommodate the product uncertainty in P2018 we looked at the different algorithms and noted the wide-spread made it hard to calibrate.

Figure 3 – this is not really a great way of showing differences/biases between runs as the lines all look basically the same.

Figure 4 - Bimodal NESOSIM output is interesting, what's going on there? I think P2018 showed weak evidence of bimodality.

L280-287 – but you don't seem to use the F labelling in the figures/tables?

Section 4.2 – these descriptions were generally quite unclear ‘superiority of the assimilation results ‘

L353 – 355: ‘We obtain the error in the Kilic19 snow depth based on the OIB-measured snow depth from 2018 to 2019.’ This is a very bad idea!

Section 4.3 seemed superfluous. You compared against one snow depth dataset and the modified Warren climatology but without much context to guide this.

Figure 8a - I think something odd is happening in Figure 8a for that big start of October jump in snow depth. This needs to be looked into.

L444-445: ‘The satellite-derived snow depth contains an uncertainty of 1 cm, and the NESOSIM snow depth uncertainty 445 reaches 5 cm (Petty et al., 2020).’ Not sure where this is from. An uncertainty of 1 cm on what I assume are your snow depth measurements can’t be right.

References

Kwok, R., Kurtz, N. T., Brucker, L., Ivanoff, A., Newman, T., Farrell, S. L., King, J., Howell, S., Webster, M. A., Paden, J., Leuschen, C., MacGregor, J. A., Richter-Menge, J., Harbeck, J., and Tschudi, M.: Intercomparison of snow depth retrievals over Arctic sea ice from radar data acquired by Operation IceBridge, The Cryosphere, 11, 2571–2593, https://doi.org/10.5194/tc-11-2571-2017, 2017.

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Far- rell, S., and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophys. Res. Lett., 40, 732–737, https://doi.org/10.1002/grl.50193, 2013.

Petty, A. A., M. Webster, L. N. Boisvert, T. Markus (2018), The NASA Eulerian Snow on Sea Ice Model (NESOSIM) v1.0: Initial model development and analysis, Geosci. Model Dev., doi: 10.5194/gmd-11-4577-2018.