Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-84-RC1, 2018
© Author(s) 2018. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development Discussions

Interactive comment on "The NASA Eulerian Snow on Sea Ice Model (NESOSIM): Initial model development and analysis" *by* Alek A. Petty et al.

Anonymous Referee #1

Received and published: 26 June 2018

The paper describes the development, calibration and validation, and sensitivity of a snow on sea ice model. Given that the model is likely to be used to retrieve ice thickness from CryoSat and ICESat altimetry, the model should be documented in the literature. However, the paper needs to be improved and some points clarified before it can be published. I would suggest that the authors consider reorganizing the paper to make it more readable and potentially shorter.

General Comments

While I recognize that a model cannot include everything, I am surprised that the authors do not include snow melt. I see this as a major shortcoming in the model. The top-left panel of Figure 3 shows, what might be interpreted, as a melt signal between March and April. I suspect that melt is also a factor in the North Atlantic sector. A

warming Arctic is almost certainly likely to have melt earlier. This years warming 'spike' likely caused melting in some sections of the Arctic. While this might not have resulted in a loss of snow mass, it would have increased snow density and caused a reduction in snow depth - refreezing/metamorphosis is another process. The authors should address in more detail leaving these key processes out.

Another concern is the use of the Polar Stereographic grid. The model tracks snow volume but only sea ice concentration, and not area, appears in Equation 1. Are the authors assuming that Polar Stereographic grids are equal area? This is not the case, they are conformal but not equal area. Cells at 70 N are about 10% smaller than cells at the pole. Maybe I am missing something here, and maybe this erroneous assumption might not have a big impact, but the authors should satisfy both themselves and readers that the choice of grid does not have an impact.

An obvious question, given the prevalence of Warren 1999 snow depths in sea ice thickness retreivals, is how different are the results presented here from W99? I would argue that W99 is the current benchmark for evaluation of snow depth products. The authors should include some discussion on this topic. I think it would also be useful to put the uncertainty reported in this paper ($\sim$ 10 cm) in the context of thickness retrievals. Ten centimeters is at least 30% of snow depths in the Arctic and similar to inter-annual variability in snow depth. While there is clearly room for model improvements, quality of precipitation fields and other forcing fields also come into play. It would be good to discuss these issues.

I find the model formulation as described in section and equations 1 through 8 confusing. This may be because I think of the process in a different way to that described here. Hopefully, my interpretation in the following paragraph, whether right or wrong, will help improve the model description.

I see the model as analogous to the evolution equations for ice thickness (or any other tracer), where the change in snow depth is the sum of a dynamic component (-del *dot* (hu) [I think] not del(hu)), and a static “snow depth evolution” component that represents snow accumulation, ablation by wind, and compaction (h^{acc}-h^{wp}+h^{bs} for the “new” snow layer and h^{wp} for the “old” snow layer). What I find confusing is that h^{acc} is the product of snow accumulation (snowfall/density) and sea ice concentration. Shouldn’t what I call the sum representing the “snow depth evolution” component be multiplied by concentration (A) not just (Sf/density); e.g. A*(Sf/density -h^{wp}+h^{bs}) [Note my comment in the paragraph above – this only applies to a uniform grid]. Similarly, should h in –del *dot* (hu) be Ah.

Maybe this is what the authors mean by “we track snow volume” – e.g. (Ah) – and “An effective snow depth...”. However, I would suggest that it is Ah that is the effective snow depth because this term represents a mean grid cell snow depth (including open water areas). Whereas, h can be thought of as a physical snow depth because it represents the process of accumulation, wind ablation and compaction at point on an ice floe. I think what I describe is a conceptual difference rather than an error in the model because h^{wp} and h^{bs} are tuned, so the wind packing and blowing snow coefficients can be thought of as including sea ice concentration, i.e. you can change the model description in the paper without changing the model.

Two further issues are with model calibration and the use of the ensemble mean snowfall. If I understand correctly, the parameters are tuned using ERA-Interim snowfall and these “best” parameters are applied to model runs with MERRA, JRA55 and the Ensemble Snowfall. I would suggest this is the wrong approach. The calibration process will compensate for biases in ERA-Interim snowfall. However, biases in the other reanalyses are different. A different “best” parameter set should be expected for each reanalysis snowfall product. Conceptually, model equations, parameters and forcing data are all part of the Model. Using parameters obtained for ERA-Interim, might have detrimental effects on snow depths when other reanalyses are used. I would recommend that the MERRA, JRA55 and (maybe) ensemble runs should be calibrated separately.

With regards to the Ensemble snowfall, the assumption behind an ensemble average being a better estimate is that individual ensemble members bracket “reality”. Is this the case with reanalysis snowfall? If all ensemble members are biased in one direction, the ensemble average will also be biased. My understanding is that both ERA-Interim and MERRA precipitation are both biased high compared to land stations. Based on Fig 11, JRA55 is also high. So is the ensemble snowfall an improvement over the individual ensemble members?

There needs to be more detail about the ensemble snowfall was generated. For example, I can envisage ERA-Interim, MERRA and JRA55 all having snow but the location of this event being shifted by one or two grid cells. On one side of the event, while ERA-Interim might have no snow, MERRA and JRA55 do have snow. By taking the median, snowfall from MERRA or JRA55, would go into the ensemble product. On the other side of the event, MERRA might not have snow but ERA-Interim and JRA55 do, so snowfall would go into the ensemble product. This would result in a larger region receiving snow. How do you deal with that situation.

With respect to the flow and structure of the paper, sections 4 and 5 seem repetitive, especially where model sensitivity to reanalyses is discussed. Essentially, sensitivity analyses for both periods give the same results. It makes sense (to me) that if you are going to compare the two time periods then the discussion and plots are merged. This would reduce repetition, make it easier for readers to compare the two time periods, and maybe shorten the paper.

Many of the figures could be improved. In many figures, the colors are not sufficiently distinct, dark purple and dark blue. This is the case with figures where dots are used. Maybe get rid of the black borders to the symbols. Also use different symbols. For many of the line graphs, increasing the weight of lines in the legend and in plots would help. Also consider whether or not you need to show the spread. The overlapping shading obscures the lines showing the means. Using shading works for two, or possibly three, series, especially if they are separated, but it starts to detract from a plot and not convey the information you want it to with more series. For example, the North Atlantic plot in Figure 11: I can't see the lower limit of the JRA55 spread or the upper limit of ASR. In many cases the spread is not discussed in the text. If you still need or want to show spread, you could just show May 1 snow depth spread as vertical bars off to the right hand side of each panel. The issue of including plots in figures but not describing the plots in the text occurs in several figures (e.g. Fig 9). I would suggest that if it not discussed, then don't include it.

Specific comments

L3, P5. "...to avoid complexity of snow melt processes". As I note in General Comments, Fig 3 shows what could be interpreted as melt. I would like to see more justification. Furthermore, a simple temperature index approach could have been used to account for melt.

Equation 1. See General Comments.

Equation 3. Should this be –\nabla \dot (hu)

Equation 4. Should the divergence be -h\nabla \dot u

Equation 5. You have a wind speed threshold for wind packing but not for blowing snow. Why is this? Studies for prairie environments indicate blowing snow initiates above $\sim$ 4 m/s, which is similar to your wind compaction threshold.

Section 2.5. Suggest this section is moved to 2.1 as an introduction to the modelling framework. This sets up the discussion of the parameterizations of the accumulation and sink terms.

L12, P11. An advantage of reanalyses is that they produce consistent outputs. Mixing and matching fields from different reanalyses breaks this consistency. How similar are the ERA-Interim winds to MERRA and JRA55.

L17, P11. "We linearly interpolate..." Do you mean bilinear interpolation? See General

Comments. This needs more detail.

L13, P14. Are significant amounts of snow in summer likely to be present in recent years? The data in Warren 1999 is 30 years old at a minimum. Are there observations from N-ICE or other field campaigns to justify non-zero initial snow depths. Furthermore, how do you initialize new sea ice? This needs to be explained.

L24, P14. "The snow depth is distributed evenly over the old and new snow layers...". Is there a reason why initial snow depth was not just assumed to be dense old snow.

L19, P16. "We carried out initial model calibration..." For this study, you only calibrated the model once, right? Suggest drop "initial" throughout this discussion unless multiple calibrations were made.

L22, P16. "...calibration involved manually tuning NESOSIM to improve the general fit..." Was this fit judged "by eye" or was some metric used? Also were all years 1980 to 1991 used, or did you leave a year out for validation during this period. While I recognize that you validated for 2000 to 2015 using OIB data, measurement accuracy and conditions might be different between the two periods.

L14, P17. It looks as if there is a larger difference in snow depth between January and March. Modelled snow depths gradually increase, while observed depths appear to increase in accumulation rate. Is this a shortcoming of the snowfall products. Also, you should mention the decrease in depths in April that could relate to melting and or compaction.

L2, P18. It is difficult to believe a correlation of 0.6 for density given the spread of points in the plot.

L7, P19. "Including the blowing snow loss... but no significant change in snow density." My first thought here, is why expect any change in density? The only mechanism by which density can be influenced by the blowing snow parameterization is a reduction in the "new sow" depth. So how deep is this "new snow" layer and how quickly does it get redistributed to the “old snow” layer?

L7 to 15, P19. Maybe add that blowing snow loss in the central Arctic are small because sea ice concentration is close to 100%.

Section 4.2 and Figure 8. I am struggling to make sense of this section. I think part of the problem is that evolution terms are shown as cumulative, which makes comparison difficult: a big snow storm could deposit several 10’s of centimeters of snow, dominating the snow depth for the rest of the season. I think you can compare the magnitudes of the terms at the end of the season (May 1) (as you do in the text) but not during the season. To compare terms during the season, I think you need to compare the timestep change in each component. The comparison is not helped by the fact that it is very difficult to distinguish lines in Figure 8. The lines in the legend need to be thicker. I would suggest leaving snow volume out of the figure.

L12, P24. Prefer “advected” to “drifting”. For snow, drifting implies blowing snow.

Figure 9. I would suggest showing only the evolution terms that you discuss in section 4.2. Other plots can be put in supplementary figures. While I suggest you don’t show the snow volume, note that the units are a depth. Moreover, I think Ah_s (sea ice concentration * snow depth) is better thought of as a gridcell mean thickness.

L13, P26. Prefer “Soviet Station” to “old” period.

L13, P26. Given the spread in snow depths in the “New Arctic” and “Soviet Station” periods, are they really that different?

L8, P29. Maybe use “difference” instead of “bias” as you have no “truth”.

L2, P30. See General Comments. If Median-SF is biased it might not be that useful.

L21, P30. “regional variability” – suggest “regional scale”. Regional scale is contrasted with Pan-Arctic Scale.

Figure 17. Why does NESOSIM have zero snow depths but OIB has non-zero snow

depths. It is difficult to interpret the panels with the OIB datasets overlayed. Maybe just show All-years but with separate panels for SRLD, JPL and GSFC OIB products. The individual years can be included as supplementary figures and/or discussed in the text. Also maybe use dots rather than x’s to avoid symbols overlapping.

Technical Comments

Abstract, L9, P1. “Several simple parameterizations to represent key sources and sinks”. The number of processes is not large, so you might as well list them explicitly, rather than keeping the reader guessing :).

L22, P3. Suggest “availability” rather than “presence”.

L12, P4. “(Show later)” Give a figure number.

L14, P4. “Ice drift”. Suggest “Ice Motion” to avoid confusion with drifting snow.

L8, P5. “...our reanalysis data...” Suggest “...reanalysis fields...”.

Table 1. Add symbols for snow densities.

Figure 3. As snow depths and densities are binned, could the data in the upper panels be shown as “box and whiskers” or just boxes. That way readers can see the amount of overlap between depth and density estimates. I suggest you spell out Soviet Stations in the figures. Add 1:1 lines on the lower panels. It would be nice to have a single symbol in the legend.

Figure 4. Does No Initial Conditions (NO IC) mean the model was initialized with 0 cm snow depth?

L1, P26. Shouldn’t this be Section 5?

L13-14, P29. Reference needed.

Figure 16. Why two symbols in the legend. Also the colors are difficult to distinguish. Maybe no black border on symbols. Also use different symbols.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-84, 2018.