Author response to Anonymous Reviewer #2 on: "The NASA Eulerian Snow on Sea Ice Model (NESOSIM): Initial model development and analysis" ***by*** **Alek A. Petty et al.**

Reviewer comments are in black, our responses are in blue.

We will also submit the revised manuscript and a word document highlighting the tracked changes we have made based on these comments.

This clearly written manuscripts provides a detailed description and exploration of the new NESOSIM snow model. NESOSIM produces gridded, daily snow thicknesses and densities for Arctic ocean sea ice during the accumulation season (defined as mid- August through April) given daily inputs of Arctic wide snowfall, sea ice concentration, ice drift, and near surface winds. Although the Arctic melt season may extend well into September, the model does not include thermodynamic or radiative processes, and this certainly limits its utility. Rather, the emphasis here is on the impacts of wind via wind packing and blowing snow loss to leads/open water. The parameterizations are fairly simple – winds exceeding a threshold can only decrease snow thickness and increase snow density. There are no snow drifts, for example, or sub-grid regions of bare ice which are present in other models. In addition, there are no snow-aging processes that may contribute to density changes. Still, the authors do a commendable job validating their model against observations and do a thorough evaluation of model sensitivity to the various snowfall reanalysis, ice drift and ice concentration products. This latter analysis highlights the true utility of the model – a simple framework for the inter-comparison of reanalysis-derived snow on sea ice data products.

We thank the reviewer for their time in producing this review and the useful comments provided.

Some specific scientific comments: The authors need to better place the work in scientific context and show how the work is unique. How is this an improvement over the simple models of snow depth forced from reanalyses? There are more complex snow on sea ice models (Lecomte-LIM, Liston-SnowModel, Hunke-CICE) which include some of the same processes (ice drift, dynamics, precipitation) yet rather than develop wind loss and compaction include some distinctly different processes (thermodynamics, radiative properties, snow ice formation, dune formation, ridge accumulation. . .). Are these models missing the "key sources and sinks"? There is also Dery and Tremblay (2004, JPO) that specifically looks at the effects of wind redistribution with an explicit mass flux into leads. Is your approach better? More useful? Consistent?

This is a fair comment, although we want to stress we don't seek in this paper to produce a model that is 'better' or 'more sophisticated' than the models currently available. Indeed this would be a real challenge considering the complexity of some models already available. But model sophistication isn't the only factor in producing reliable snow depths - the forcing is arguably more important, and there is still a high level of uncertainty surrounding the sensitivity of snow to various forcing data, especially snowfall. As such we sought o develop a model that will enable us to explore these sensitivities and increase model sophistication as needed based on these uncertainties.

We have moved up to the introduction and added more text regarding our motivation/philosophy (this was at the start of the modelling description) to make this clear from the outset and have also adapted the introduction to add the following before this statement:

" Due to these observational limitations, the sea ice community often relies on simple models of snow depth forced by reanalyses (primarily snowfall data) (e.g., Maksym and Markus 2008;

Kwok and Cunningham, 2008; Blanchard-Wrigglesworth et al., 2018). More sophisticated snow on sea ice models are available, such as SnowModel, a terrestrial snow model recently adapted for sea ice environments (Liston et al., 2018), as well as the snow parameterizations included in sea ice climate model components, such as CICE (Hunke & Lipscomb, 2010) and the Louvain-la-Neuve Sea Ice Model (LIM), which has recently undergone various improvements to its snow physics (Lecomte et al., 2015)."

What is the impact of excluding thermodynamic processes on your results? Does this change your conclusions about the impact of wind processes?

Unfortunately we really lack the data needed to better answer this question. It's likely that including thermodynamic processes (e.g. snow melt) will reduce the impact of the wind loss term, but that is already a very unconstrained process in this model.

We also refer the reviewer to our response on this subject of missing melt processes raised by Reviewer 1 and the newly added discussion of not including snow thermodynamics, the potential impact of this, and the future work section at the end of the paper.

Some misleading statements: First sentence of the abstract. . . . "produces daily esti- mates of depth and density of snow across the polar oceans". Not yet because of some important missing processes. Qualify with Arctic only and during the accumulation sea- son.

Agreed. Changed 'polar oceans' to 'Arctic Ocean through the accumulation season'.

Using old vs new snow in the text and figures. It's clear that there is intention to one day include snow aging, but for now there is only fresh vs compaction. The depth hoar densities of 150-250 is never used in the model even though paragraph 10 seems to suggest that it is. The old snow value is 350 kg/m3 which is not the average of the higher end of wind slab and depth hoar (325 kg/m3) but rather the average of the wind slab bounds.

The reviewer is correct that there is the intention to one day include snow aging in the model, but, for now, there is only "new" and "old" snow. We think there may be some confusion here as we take a weighted average based on the average ratio of depth hoar and wind slab from Radionov et al. (1997) and Sturm et al. (2002). The density values chosen for these layers are based on the literature on snow density observations (e.g., Radionov et al., 1997; Sturm et al., 1998; Sturm et al., 2002; Sturm and Massom, 2017) and implicitly include values for density layers that are not explicitly treated in the model. For example, the compacted layer uses a value of 350 kg/m3, which is a value calculated using the average ratios of depth hoar (40%, 250 kg/m3) and wind slab (60%, we found values of 410 kg/m3 in the referenced works above) within a snowpack. Admittedly, the ratio of depth hoar and wind slab seasonally changes and the model does not account for this seasonal change. However, we feel this is a better treatment than neglecting the influence of depth hoar altogether in the "old" snow layer. We have edited this description in the revised manuscript to make this clearer.

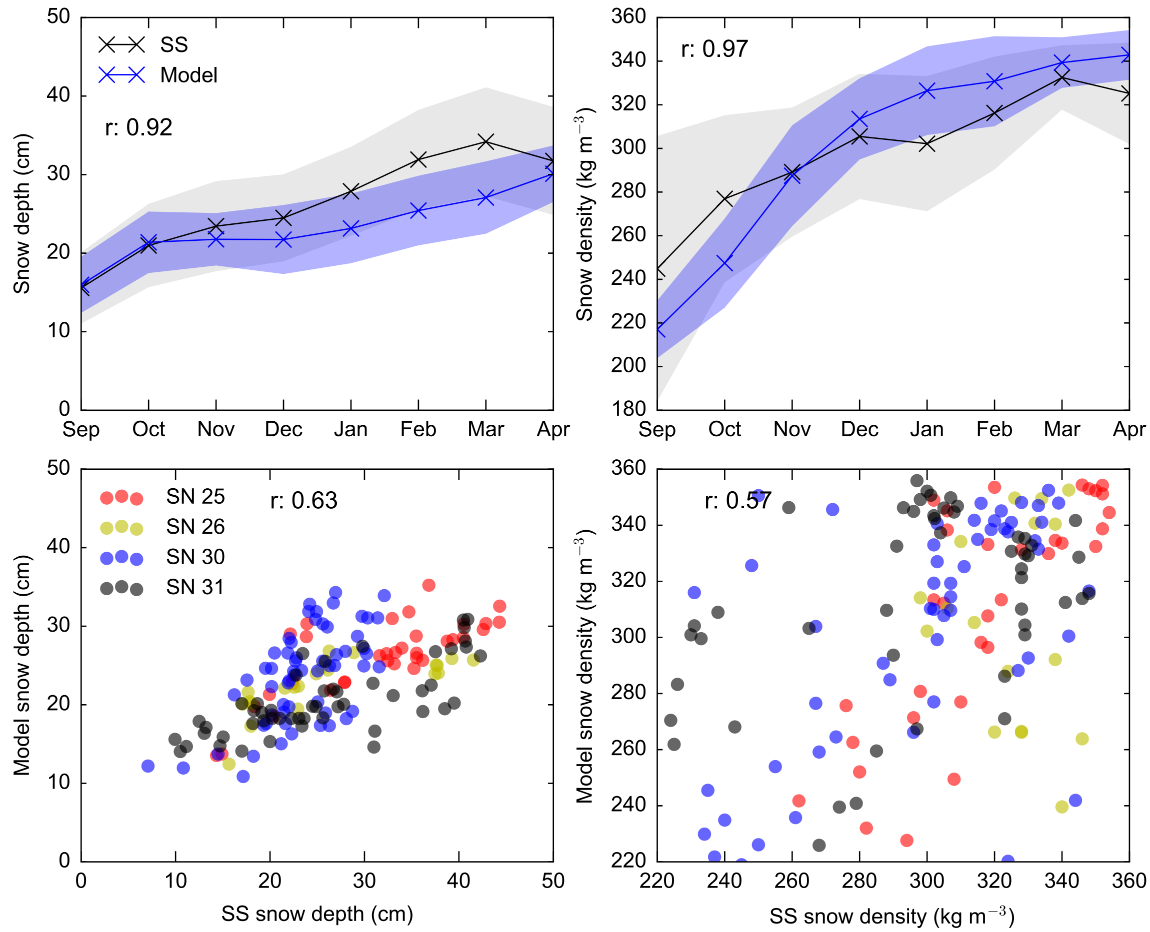
Perhaps future developments could be kept to a specific section to better clarify what the model does and doesn't do.

Yes, Reviewer 1 wanted this too, so we have added this to the final summary section. We refer the reviewer to this new subsection.

Snow density in NEOSIM is bounded by the two chosen snow density parameters (200 and 350 kg /m3) even though the observations referenced give values for dry snow of 150 and wind slab ~400 kg /m3 on average. Why exclude these possibilities at the outset? Instead of using an average value, doesn't it make more sense to use the upper and lower bounds given the nature of the parameterization? How sensitive is the model to these values?

We did experiment with these values, but found that this resulted in an over-estimation of the strength of the snow density cycle. See figure below. We have added this to the new manuscript:

"We did experiment with alternative snow densities (e.g. the wider spread of 150 and 400 kg m^{-3}) but found this provided worse correspondence with the seasonal snow density evolution compiled from in-situ Soviet Station data (introduced in Section 3.4)."



The late summer initial conditions integrate all the missing snow melt processes and for that reason, they are rather important. The paragraph on page 14 does a fairly good job motivating your approach, but it would be clearer if you showed the equations for hs(0) and hs(1) after summer melt. Also better explain how snowfall events factor into this parameterization and explain why keeping the same fraction for fresh/compacted snow is the right approach (or clarify if you do something different). It would also be informative to see the Aug 15 values in your figures 3 and 4. Are there Aug observations to help validate the IC parameterization and fig 2 in particular?

We think there may be some confusion here. We rerun the model every August, regardless of the prior spring snow conditions. The starting point for the model is a direct application of the initial condition snow depths.

We don't show these values in the figure because we start the model halfway through August and these figures are showing monthly means/spreads.

The equal split between fresh/old snow was due to (albeit fairly crude) reasoning that some of this initial snow may be from snow persisting through the melt season and some due to new snowfall through summer based on observations in Radionov et al. (1997). We explain and cite this in the manuscript.

We lack consistent observations of August snow depths to validate this component of the model, but we did look at the snow depth data from ice mass balance buoys – some show the presence of snow in August, while others do not, within the same summer seasons. While these observations suggest snow is present in August, these are point measurements and may not be wholly representative of basin-scale snow depth distributions.

Why absorb the timestep in the model equations? In (2) the parameter alpha has a timestep dependence that isn't explicitly called out and as a result, 0.05 is less meaningful. Better to define an alpha with units of per second.

Agreed, we have changed these equations to include a time step value, so these coefficients are now independent of the length of the time step. The units have been updated accordingly (Beta needs to be in units of per meter as we multiply by wind speed).

Are the differences between simulations with different snowfall estimates larger than the differences between time periods? Are the time period differences significant?

The spread between reanalyses makes it challenging to ascertain confidence in a physical system change between periods. Understanding the factors contributing to the differences between periods and reanalyses (e.g. the contribution from precip, freeze-up, other processes) is beyond the scope of this paper, but is a topic of ongoing work.

The potential changes in snow depth and their physical cause are part of a hopeful follow-up study which can elaborate on this in the detail required. We also refer you to our response to Reviewer 3 where we provide a brief summary of our thoughts about this.

Note also that to shorten the analysis description (and to prevent the focus being about changes in snow depth) we have merged the 1980s/2000s reanalysis sensitivity study sections and moved the 1980s figures of model evaluation to the Supplementary Information. The 1980s reanalysis sensitivity studies did not add much to our discussion and interpretation of the model performance so we have dropped this from the revised manuscript.

Fig 14 seems to suggest that ice drift is actually quite important but masked by basin or large regional averaging. Magnitudes of the differences are similar to the snowfall sensitivity. Impacts are near the ice edge (increase ice retreat?) and add to smaller (but still > 100 km) scale variability (potentially impacting melt-pond formation).

Agreed! We discuss this in the paper already so haven't added more here.

Technical corrections:

Table1. add the model variable in the table.

We have added all model variables to the table.

Define U in Eq. (5).

It was defined after Eq. 2 earlier.

Missing) in Eq. (6)

Added

Why is the bs term in Eq. (7) positive?

The wind loss term is now negative, so this should make more sense now!

Missing t dependence in some terms in Eq. (7) and the next equation Line 26 missing “of” in “one the better..”

Thanks, good spot. Added.

Add W99 to upper panels of fig3 Add W99 to fig 6 (a)

We decided not to add W99 to figure 3 as the legend makes clear we show data from the drifting Soviet Stations (hence SS in the legend) from 1981 to 1991. The W99 is a climatology of these data over the longer time period (back to the 1950s) so it is not exactly the same. Figure 6 shows no Warren or Soviet Station data so we didn't add this either.

In explanation for fig 8, NA region shows a small “decrease” due to convergence, not increase. CA shows a small increase but is not mentioned.

Changed this to: 'The NA region also shows a small (~2 cm) decrease (increase) in snow depth driven by snow/ice divergence (ice/snow advection), while the CA region shows a small (~2 cm) increase in snow depth driven by snow/ice convergence.'

Fig (9). What is (b) Ocean? Change (d) wind/leads to wind loss to leads. Is there an ice area cutoff for snow depth in (k)?

Ocean was the snowfall into the ocean. We have updated all these labels to make this a lot clearer, including the variable names and a link in the caption to the table which provides the model variables.

All snow depth results and budget fields shown in the analysis use a concentration cut-off of 15% and which we have now added to the revised manuscript at the start of page 9 after the model formulation description: 'Note that this bulk snow density is masked if the respective ice concentration is less than 15% or snow volume is less than 2 cm, while all snow budget terms presented here show data only when the concentration is above 15%, to prevent spurious results in regions of near open water conditions.'

Table 4 could be improved by adding the 1981-1991 time period for comparison, identifying the boxed regions as “reanalysis sensitivity”, “ice drift sensitivity”, “ice concentration sensitivity” and including in the description that the default configuration is MEDIAN-SF, NSIDCv3 ice drift and Bootstrap ice concentration

See comment above. The comparison between time periods is the focus of on-going work led by co-author of this paper Melinda Webster. As it requires careful evaluation of the uncertainty in forcing data and changes in these fields we feel we can't make a simple comment about that here. We have thus decided to drop the 1980s reanalysis figure (the 2000s figure serves this purpose just as well) and moved the 1980s budget figures to the SI.

End of p 30. Comment on the NA region in table 4 when ice drift is included. Seems to be important here too.

Yes, this was an omission and has been added to the text.

Fig 13, (d) NA does not appear to be consistent with table 4. The NODRIFT value on May 1 is around 34 cm in table 4 but seems to be much higher in this figure. Explain.

The table value was wrong and has been corrected. Although based on the tweak to the wind loss term (introduction of the wind action threshold for this term as well as wind packing based on reviewer recommendation), all values have been reproduced and are slightly different from their original values. Apologies!

Last comment before 5.3 does not seem correct. There does appear to be bias in the real-time products with respect to the peripheral seas .

Agreed, we have dropped this statement.

P 37. What are uncertainties in the OIB observations?

The recent STOSIWIG study (Kwok et al., 2017) stated that snow depth can be estimated from a Snow Radar echogram with an uncertainty of 'several centimeters' although this depends strongly on the ice conditions, the particular Snow Radar system being used, and various other factors (e.g. geolocation errors associated with the plane pitch and roll. We have added a sentence regarding this to the discussion at the end of Section 3, where the OIB data are introduced.

Mention the results from sensitivity of ice concentration in the summary. These were interesting and significant.

OK we added the following: 'We also briefly assessed the sensitivity of NESOSIM to the input concentration data, with our results suggesting that this choice of product (Bootstrap and NASA Team explored in this study) can have a significant impact and should not be overlooked.'