Response to Reviewer comments to manuscript "Exploring the utility of quantitative network design in evaluating Arctic sea-ice thickness sampling strategies"

July 23, 2015

We thank the editor and the reviewer for their careful inspection of the manuscript. We find the presentation of this innovative study to a community of observationalists and modellers challenging, and it was very helpful that you pointed at spots where we could further clarify.

We added the following sentence to the acknowledgements:

The authors thank two anoymous reviewers and Christian Haas for valuable comments on the manuscript.

In the following we address first the comments by the anonymous referee and then those by the editor.

1 comments by Anonymous Referee #2

1. *General: Inconsistent use of sea ice versus sea-ice. I favour sea ice (which then makes sea ice - ocean more readable), but up to the authors discretion.*

   O.k. We changed to "sea ice" throughout.

2. *P. 4, l. 20: Acronym (GFDL) is introduced, but never used again.*

   Acronym removed.

3. *P. 10, l. 2: In replicating these variables in condensed way → In replicating these variables in a condensed way*

   Fixed.

4. *P. 14, l 5: I believe some words were inadvertently left out here: Original manuscript: For both, region 7 and 8, Fig. 9 shows the direction in which a change of tau yields the largest increase in ice thickness. Latest revision: For both, region 7 and 8, Fig. 8 a change of tau yields the largest increase in ice thickness.*

Thanks. There was an accident during the revision. The original version is reconstructed in the final manuscript.

5. *P. 14, l 4: (cmp 9) – I assume the authors mean (compared with Fig. 9). The abbreviation does not help readability. Also note that the bracket is not closed – missing ")".*

   Thanks. The sentence is revised and reads now:

   B2F transect with respect to changes in various impact factors (relative to the default settings used for Figure 9) for the NOB target region.

1.1 Open questions (for which I don't necessary expect an answer)

1. *Why is B2F only superior for snow thickness, but C2F for ice thickness and area?*

   This question relates to Figure 6. We extended the text discussing this Figure in bold face:

   "Figure 6 shows the performance of each transect for improving forecasts for the target region covering the coastal ocean from Bering Strait to Prudhoe Bay (BS2PB). They show similar performance because this target quantity is temporally averaged from May to 20 August. B2F is superior for snow thickness and C2F for ice thickness and area. **This can be explained by the sensitivity of these three target quantities (Figure 7). Relative to ice thickness and area, snow thickness has a larger sensitivity to the initial (ice and snow) conditions (in particular over region 8) than to the surface forcing and the process parameters. And the initial snow thickness over region 8 is, of course, better observed by B2F (which crosses this region) than by C2F.** As an additional test case we evaluate the combination of the two transects, which clearly shows their complementarity."

2. *Figure 12: I find the increased sensitivity to perfect forcing and perfect model over perfect forcing intriguing given the almost zero sensitivity to perfect model otherwise seen. Does this indicate that the forcing introduces so much random error as to swamp any model error? Only when forcing error is eliminated, does the model error then show some sensitivity.*

   Yes, this is remarkable. We have added an extra sentence to clarify:

   Interestingly, combining the perfectly calibrated model and the perfect atmospheric forecast assumptions doubles the uncertainty reductions compared to the perfect atmospheric forecast assumptions alone. **In this case all the observational constraint can fully act to reduce uncertainty in the inititial conditions.**

   The point is that in the perfect model + imperfect forcing case much of the "power" of the observational constraint is "wasted" to also constrain the forcing.

3. *Figs. 3, 6, 9, and 12: There is a lot of white space. Do the authors really wish to have the sensitivity scale from 0-100% in all cases at the expense of being able to differentiate some of the smaller sensitivities?*

   Yes, it was a deliberate choice to keep the scale constant, because we want to highlight the larger scale effects (also between Figures) rather than small differences.

4. *Figures 4, 7, and 10: The authors have done their best to improve the readability of these figures. However, I fear all but the most eagle-eyed of readers will be dependent on the (scalable) electronic versions of these figures for all but the most trivial of interpretations.*

   Thanks. We'll point out to the copy editor that these sensivity plots should be displayed as large as possible.

2 comments by Christian Haas

Thank you for the revisions of your manuscript and for addressing most points raised by the reviewers. As one reviewer has requested major revisions I have sent the revised manuscript out for review again. However, to speed up the processes once that review has been received, I would already like to make a few additional suggestions as the Editor of your manuscript and would be glad if you could consider them in your revisions. This is an interesting paper and very promising approach, and the material is generally well presented.

1. *However, the text is quite compact and I feel that some aspects are still unclear. In particular, you should still better describe what the uncertainty reduction is and how it can be interpreted, despite the new paragraph on page 11, l7-13. Please could you expand this paragraph and better explain what uncertainty reduction really is and how it relates to the terms in equation 1. Maybe you could also include an example, e.g. that uncertainty reduction of xxx % compared to the initial uncertainty of xxx means that parameter xxx is known better by ??? How does this relate to the parameters in table 1?*

   We have added the following text after the definition of the posterior uncertainty in the target quantity Eq. 4:

   Evaluating Eq. 4 for the prior uncertainty $\mathbf{C}(x_0)$ instead of the posterior uncertainty $\mathbf{C}(x)$, i.e. for a case without observational constraint, yields a prior uncertainty for the target quantity:

   $$\sigma(y_0)^2 = \mathbf{N}'\mathbf{C}(x_0)\mathbf{N}'^{\mathrm{T}} + \sigma(y_{mod})^2 \,. \quad (1)$$

   We define the term uncertainty reduction relative to $\sigma(y_0)$, i.e. by

   $$\frac{\sigma(y_0) - \sigma(y)}{\sigma(y_0)} = 1 - \frac{\sigma(y)}{\sigma(y_0)} \quad (2)$$

For example, if $\sigma(y)$ is 90% of $\sigma(y_0)$, then the uncertainty reduction is 10%, i.e. we have increased our knowledge on y by 10%.

Table 1 shows the prior σ for the control variables, their squares fill the diagonal of $\mathbf{C}(x_0)$.

2. *Your use of the term observations is also confusing. Particularly with regard to page 7, l14-17, page 11, l20, and p12, l5. Please check that these sentences are consistent and maybe include more explanatory text.*

   Yes, the use is consistent. On p7 we discuss the role of observations in Eq (3), on p11 the potential observations for which response functions were calculated and of which we could, in principle, select subsets for evaluation, and on p12 the two sets of potential observatations that were actually selected for evaluation in this study. For clarification we have added explanatory text (in bold face) to the sentence on p11, which now reads:

   It provides response functions for **potential observations of** each of these three observables, for each surface grid cell, and for each day of the simulation period (i.e. about 5 million possible observations **of which subsets can be selected for evaluation**) with a user-defined data uncertainty.

3. *Similarly, it is unclear when and why you have used ice concentration and when not, e.g. p9, l14-16, and p11, l17. Why was ice concentration not used? Or was it not used as input variable, but as predicted variable?*

   Yes, exactly. We added text in bold face to clarify:

   For all target regions delineated in Fig. 1, we use spatial averages of **the three simulated quantities:** ice concentration (fraction of area covered by ice, regardless of the 15% floor used in the definition of ice extent), ice thickness, and snow thickness.

   And where we define the "observations":

   The "observations" consist of model output of ice and snow thickness at each grid cell that intersects with the transect as indicated in Fig. 1. The default case specifies a data uncertainty of 30 cm for both quantities. **Sea ice concentration is not observed.**

4. *The abstract is very general. Can you include some quantitative results? Also I wonder if NASA OIB needs to be mentioned in the abstract. Arent the results likewise applicable to other thickness data from the same transects, e.g. from submarine missions?*

   In the abstract we deliberately wanted to be general and advertise the merits of the method in a qualitative way. The reason for mentioning that the transects are derived from NASA OIB was to indicate that this theoretical study has some link to reality.

5. *The long flight lines raise another question. How would the results be different were the flight tracks much shorter and only focused on the study regions and regions slightly further upstream? Can you comment on that?*

   We can only speculate about the effect of shorter tracks. Our guess is that a reduction at the "Fram" end would be rather tolerable than at the other end, closer to the target regions. A thorough analysis may be one of the topics of a follow-up study.

6. *In contrast to one of the reviewers, I do believe that regional studies with less extensive observational data would be feasible too? At least when relative short prediction times are considered. Wouldnt only ice conditions be relevant in regions and ice fields which could drift into the study region during the prediction period?*

   Yes, the short prediction time is crucial. Then things should work as they do here.

   This is why (in the previous revision) we had left the regional application in the conclusions and just added the bold faced text:

   Furthermore, rather than operating Arctic-wide, the same concept can be applied on smaller regional scale**, when the forecasting period is short enough to ensure that the main influence factors can be appropriately simulated within the model domain**.

2.1 Some minor comments

:

1. *P3, l24: what is the name of the satellite mission?*

   added in brackets: **... (the Orbiting Carbon Observatory, Crisp et al. (2004))**

2. *P4, l 14: approach THAT operates*

   fixed

3. *P4, l24: is the southern boundary the Atlantic? Needs to be clarified.*

   We added the text in bold face: At the southern boundary **(near 50° N)** an open boundary condition has been implemented ...

4. *P5, l11: rephrase precipitation is added to the snow mass? What happens anyways when air temperatures are at or above freezing, i.e. during the summer?*

   Text extended (bold faced):

   When atmospheric temperatures are below the freezing point, precipitation is added to the snow mass**otherwise to the ocean**.

5. *Spell out PHC (p5), TAF (p8).*

   done

6. *P6, l18: why does this guarantee full consistency?*

   The text has been revised to clarify this:

   The control vector $\tilde{x}$ that minimises Eq. (1) achieves a balance between the observational constraints and the prior information. The minimum is determined through variation of the control vector (hence variational assimilation) comprising initial and boundary conditions and process parameters. In contrast to sequential assimilation approaches, which result in a sequence of corrections of the state predicted by the model, the variational approach guarantees full consistency with the dynamics imposed by the model, as it provides an entire trajectory through the state space of the model in response to the change in the control vector. In the case of our model this means that we infer a trajectory that assures conservation of mass, energy and momentum (except at the lateral domain boundaries). We note that, in this QND study, no minimisation of Eq. (1) is required.

7. *P7, l22: WAS perfect*

   corrected.

8. *P10, l12, and figure 2: Why do you describe 9 regions but only use 2 (or 3?) in the paper?* We do use all 9 regions, but regions 7 and 8 turn out to be those in which, for example, a change of boundary conditions has the largest impact on the target quantities (see, e.g. Figure 8).

9. *P12, l7: should the 30 cm in the text not be the same as the value in table 1, i.e. 0.5 m for ice and 0.2 m for snow thickness?* The values in Table 1 indicates the prior uncertainty we assign to the ice thickness and snow depth fields we start our model run with. In contrast the 30 cm refer to the uncertainy in the observation.

10. *P13, l22: Do you mean LOW impact? Why is it remarkable?*

    The point is here that over 10 days B2F does a good job for ice concentration over Chukchi. This looks surprising as it is remote of Chukchi and 10 days are to short. But it is no longer surprising as soon as we realise that the parameter $h0$ has a LARGE impact on ice concentration after 10 days. So what happens is that the B2F observations constrain $h0$ and this constraint is progagated to ice concentration after 10 days.

11. *P14, l10: do you mean INCREASES ice concentration?*

    We have added text in bold face to clarify:

    Parameter h_0, **which essentially determines the distribution of newly formed ice in the vertical vs the horizontal dimension,** has a negative impact: Increasing h_0 yields thicker newly formed ice and consequently reduces the ice concentration.

In other words: h_0 is the thickness that newly formed ice is assumed to have. If h_0 is increased it will reduce the ice area covered by this newly formed ice.

12. *P15, l4: compare to FIGURE 9, or cf Figure 9.*

    Text revised to read:

    B2F transect with respect to changes in various impact factors (relative to the default settings used for Figure 9) for the NOB target region.

13. *Table 1: precip, given in m/s?*

    That's o.k. it is volume per second per unit area (m3/m2/s=m/2).

14. *All table and figure captions. Please spell out the acronyms/abbreviations of variables etc. and just include them in parentheses such that the captions are self-explaining and do not require searching the text for explanation.*

    Only left Abbr. of target regions and flight transects (which are unique to this manuscript anyway and their replacement might rather confuse the reader) in the captions (if their replacement was also requested, we'll do that within the copy editing process). Remainder has been replaced already.

15. *Fig 1: BS2PB is light blue? How would results be different if flight tracks were shorter?*

    Light blue, yes: Corrected.

    For shorter transects see above.

16. *Figure 12: add annotation with case description to each panel, such that differences between their contents become immediately obvious.*

    Done.