Response to Reviewer comments to manuscript "Arctic Mission Benefit Analysis: Impact of Sea Ice Thickness, Freeboard, and Snow Depth Products on Sea Ice Forecast Performance"

May 9, 2018

We thank the reviewers for their careful inspection of the manuscript. In the following we address their comments point-by-point. We use *text in italics* to repeat the reviewer comments, normal text for our response, and **bold faced text** for quotations from the manuscript, with changes marked in colour. Where we use line or Figure numbers these refer to the manuscript version published in TCD.

We provide the revised manuscript (with and without changes highlighted) in the supplement.

1 comments by Anonymous Referee #1

This is a well written and detailed paper in which CryoSat-2-derived ice freeboard, sea ice thickness and snow depth products are used to assess a coupled ice-ocean models forecast performance for a region that includes the East Siberian Sea, Outer New Siberian islands and the West Laptev Sea. A comprehensive list of control variables ranging from atmospheric forcing, initialization fields and physical processes (e.g., density of sea ice) are used. The paper presents a very thorough description of the Quantitative Network Design (QND) and how it is used to assess the observational impact of remotely sensed ice freeboard on the uncertainty reduction on sea ice volume and snow volume. The substance of the study is highlighted in Figure 16 which shows the uncertainty reduction in the three areas for sea ice volume and snow volume when evaluating quantities such as sea ice thickness, radar freeboard, and lidar freeboard.

General Comments:

There is a wealth of information provided to the reader in terms of detailed tables and figures. The paper title is somewhat misleading as only the last day of the model forecast (May 28, 2015) is used in the evaluation. I expected a much longer period of analysis (e.g., weeks to months). Figures 14 and 15 (see comments below) are too difficult to read in their present format. The manuscript appears to include all relevant references.

We'll respond in detail below, where we address the specific comments, which take up all the above points again.

Specific Comments:

The paper title implies that the use of sea ice thickness, freeboard, and snow depth products will be used to assess sea ice forecast performance. However, I only see an evaluation of the model forecast on May 28, 2015. Could an extended period (entire month of May 2015) be evaluated? The Northern Sea Route is mentioned many times in the text. The maritime transport industry should have interest in how ArcMBA and the QND approach could be used to predict the ice conditions for June, July and possibly August as well. Could this work be extended in this manner for a future study?

Yes, the QND approach is flexible in that respect, but, of course, this requires that the corresponding target Jacobians for June, July and August be computed and the QND analysis be performed (which we consider beyond the scope of the present study).

Page 6 (line 9-10): The paper states "We perform these predictions for May 28, 2015, a point where there is still sufficient snow cover for our prediction to be relevant". How- ever, on page 12 (lines 23-24) the paper states "Note that on May 28 parts of the target regions are almost snow free already". How does this impact the first statement about "sufficient snow cover"?

Snow is already reduced strongly on May 28 especially in target region ESS (reduced by about 90%) but it still shows enough sensitivity to perturbations in the control vector. To clarify, we have changed the sentence on page 12 to read:

Note that on May 28 ~~parts of~~ over the target regions ~~are almost snow free already~~ a large fraction of snow cover of has already melted. The misfit to the modified Warren climatology in the target area East Siberian Sea is on the order of about $10cm$ (50% relative error) but much less for the other target areas.

Also in this paragraph, to make sure I understand; the model was spun up for a period beginning January 1, 1979. A restart file from March 31, 2015 was used to initialize MPIOM and the modeling system was run with data assimilation through April 30, 2015. The 4-week model forecasts begin May 1, 2015 and I assume are forced with the ERA-Interim reanalysis, but without any ocean/ice data assimilation? Is this correct?

We confirm that, in the spin-up, the model is driven by ERAinterim and no data assimilation is performed. We added a clarification:

~~As we base our QND experiments on simulations from April 1 to May 28, we~~ Next we address Arctic MPIOM performance over our assimilation and forecasting period (see Figure 4). We show the April mean and the May 28 mean of the modelled SIT and the misfit of the April mean thickness to that retrieved from CryoSat-2 (Figure 8). For a comparison of CryoSat-2 thickness to in situ observations we refer to Haas et al. (2017).

Please provide a more detailed caption for Figure 2 and provide some additional text about the trajectories (notional) depicted in this figure.

We added detail to the caption:

Schematic ~~Presentation~~ presentation of the QND procedure: Each coloured line illustrates a model trajectory that simulates from a given value of the control vector (x) counterparts of the observations (d_1 and d_2) and a target quantity (y). Through the model, the observations act as constraints on the control vector, which reduces its uncertainty from $C(x_0)$ to $C(x)$. This uncertainty reduction on the control vector translates into an uncertainty reduction in the target quantity from $\sigma(y_0)$ to $\sigma(y)$.

Page 12 (lines 17-18): Fig. 8c depicts the mean April 2015 misfit of the modeled SIT to AWI CryoSat-2 ice thickness. How does the April 2015 AWI CryoSat-2 data compare to NASA OIB for this period? Please provide an additional plot showing the NASA OIB data overlaid on the 2015 mean CryoSat-2 SIT. How does OIB compare with the AWI data?

The validation of the CryoSat-2 product is beyond the scope of this study, and to our knowledge no such comparison exists with CryoSat-2 from AWI. That is mainly due to the fact that OIB observations are remotely sensed as well (no direct observations); OIB uses Laser altimetry to measure the Laser freeboard and utilises, as the CryosSat-2 algorithm, assumptions on the sea ice and snow densities to calculate the sea ice thickness. In contrast to the CryosSat-2 algorithm, no snow depth climatology is taken into account but snow depth measurements from a snow radar. But these snow depth observations are uncertain as well. The estimation of the uncertainty of OIB snow depth observations taken by OIB is an active research area.

Better suited for a validation of the CryoSat-2 thickness are more direct measurements as, for instance, electro-magnetic (EM) in-situ measurements because these are (largely) independent of sea ice and snow densities. EM-thickness measurements deliver the thickness of sea ice and snow. We added to the manuscript a reference to Haas et al. (2017) who show in-situ EM-thickness observations compared to CryoSat-2 sea ice and snow thickness in the Lincoln Sea (see revised part of manuscript in response to above comment starting *Also in this paragraph* ...). The unpublished Figure 1 below (only shown in this response) depicts a scatter plot of areal EM-thickness observations and Cryosat-2 thickness observation taken in April 2017 during the PAMARCMIP2017 campaign on the Chukchi Shelf, Northwind Ridge and in the Lincoln Sea and Fram Strait.

Page 12 (line 22): There is mention of "modified Warren climatology", but no explanation on how the modified snowcover was used in the CryoSat-2 ice freeboard retrievals. Please explain and provide specific details.

The main challenge for sea-ice thickness retrieval with satellite altimeters is the parameterisation of snow depth on sea ice, which is not measured routinely. The current processors use a snow climatology instead of remotely-sensed data. Warren et al. (1999) established this climatology with results from drifting stations mainly on multi-year sea ice collected over the past decades. However, since the Arctic Ocean shows a significant higher fraction of first-year sea ice in recent years, the approach proposed by Kurtz and Farrell (2011) is followed and the

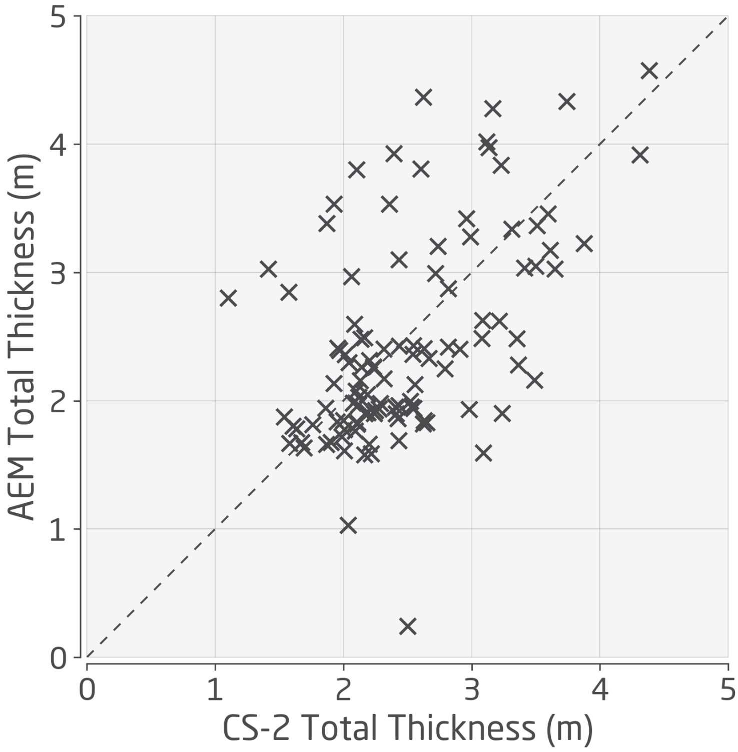


Figure 1: Scatter plot of CryoSat-2 total thickness (sea ice + snow thickness) and areal EM (AEM) total thickness in April 2017. Mean total thickness CryoSat-2: $2.71m$, mean total thickness AEM: $2.62m$, $R = 0.72$, $RMSE = 0.81$.

climatological snow depth values are multiplied over first year ice with a factor of 0.5. We revised the manuscript accordingly:

Figure 9 depicts the April mean and the May 28 mean of the modelled snow depth and the misfit to the modified Warren climatology (Warren et al., 1999) that is used in the CryoSat-2 retrieval (see Section 2.5). The main challenge for sea ice thickness retrieval with satellite altimeters is the parameterisation of snow depth on sea ice, which is still not measured routinely. The current CryoSat-2 retrieval uses a modified snow climatology that addresses shortcomings of the Warren et al. (1999) climatology that was based largely on data from drifting stations mainly on multi-year sea ice collected over the past decades, and hence is not reflective of a much younger, more seasonal Arctic ice cover. Given the increased fraction of first-year ice in the Arctic Ocean, the approach proposed by Kurtz and Farrell (2011) is used and the climatological snow depth values used in the retrieval are multiplied over first-year ice by a factor of 0.5.

Table 3 shows significant reduction in the uncertainties for SIV and SNV. I am surprised there is very little mention of these results in the text. Please expand on this in the text.

In fact we describe these uncertainty reductions in depth on the two pages following the presentation of Table 3, together with Figure 16.

Page 28: Graphs in Figure 14 are very difficult to read (too small). Perhaps graphs for Reg 1, 2, 3, 4, 5, 7, 8, 9 can be removed and the remaining graphs could be enlarged.

We think it is instructive to show sensitivities to the full control vector, so the reader understands “where the action is”. And then we followed the suggestion to show enlarged plots of the Jacobian rows for Region 6 and the model parameters in a new Figure.

Page 29: Figure 15 is a little easier to read than Fig 14, but still a challenge to read the individual plots.

Similarly, also here we show the full control vector first and then add a new Figure with Jacobian rows for Regions 5 and 6 and the model parameters.

Although mentioned briefly in the Summary and Conclusions, it would be of value to assess the impact from this study on the ice drift. Are there ice drift observations available in May 2015 to perform an analysis?

The system could indeed be extended by including either an ice drift product into the set of products to be evaluated or by using ice drift as an additional target quantity to be predicted by the model. The former would require an extension of the observational Jacobian and the latter an extension of the target Jacobian. Both could be topics for a follow-up study. An example of an ice drift product to be evaluated could be that of OSI SAF (Lavergne et al., 2010). We have extended the suggestions of possible ArcMBA extensions in the conclusions section.

Technical Corrections:

- *Page 2 (line 1): Spell out EO as this is first time referenced.*
  Done.
- *Page 2 (line 12): Dont spell out EO here.*
  Done.
- *Page 2 (line 14): Is there a better term for "rawer"? Also "rawer" is used in several instances through page 9.*
  Yes, "rawer" may not be ideal, now we also use "lower-level" to clarify (and in the following use either according to context):
  The ~~constraint from rawer EO products that~~ constraints from lower-level EO products (i.e. rawer products that more directly related to the actual measurement) that are used to derive SIT products may be even stronger, because ~~these rawer products~~ such products that conform more closely to the raw EO data are typically more accurate.
- *Page 2 (line 21): I suggest deleting phrase "products of further".*
  Done.
- *Page 2 (lines 22-23). Add comma after approach, and delete QND on line 23 to make 1 sentence.*
  Done.
- *Page 3 (line 3): LFB has already been defined.*
  Only in the abstract, and we think the journal policy is to repeat the definition in the main text, to be checked by the copy editor anyhow ...
- *Page 4: Figure 1 caption should read "Oval boxes", not "Ovals boxes".*
  Done.
- *Page 6: Figure 2 caption should read "presentation" (lower case p).*
  Done.
- *Page 7: Fig. 3 blue background is too dark. Please modify for better clarity?*
  We modified the plot.
- *Page 8 (line 10): Can a reference be given for "Gent and McWilliams style"?*
  Reference is given.
- *Page 9 (line 15): Replace "will be" to "are".*
  Done.
- *Page 9 (line 21): How do you come up with 34 years? Jan 1 1979 to March 31, 2015 should be 36 years.*
  Oh, an embarrassing error. Many thanks for helping us with the basic algebra!
- *Page 9 (line 24): Spell out OSI SAF.*
  Done.
- *Page 9 (line 28): delete "by" and put Lindsay and Schweiger (2015) in parenthesis.*
  Done.
- *Page 10 (line 5): typo xxx should be "regions".*
  Done.
- *Page 13 (line 4) remove "could".*
  Page 12. Done.
- *Page 21 (line 5): spell out EASE.*
  Done.
- *Page 22 (line 5): should be "For later use 'it' also lists".*
  Done.

2 comments by Anonymous Referee #2

General comment:
The authors present a formalism to assess possible benefits of different Earth Observation (EO) products for reanalysis Arctic sea ice data. The authors consider seven satellite products: sea ice thickness and free board, radar free board (derived from satellite data), and the hypothetical data laser freeboard and snow depth, the latter both in higher and lower accuracy. The question focused on in the assessment is how uncertainties of EO products are reflected in (user) defined variables, so called target quantities. An outcome of this study could be to identify those kind of EO products, which lead to the fewest uncertainties in the target quantities. The authors consider snow volume and and sea ice volume as target quantities. Sources of uncertainties are not only found in the EO products, but also in the model and experimental setup, such as initial and boundary conditions, parameterization and a formulation of the physics. To identify the impact of these onto the uncertainty propagation towards the target quantities, a so called control vector finds application in the formalism, containing representations of these sources.
Their findings are different for the target quantities:
Discussing the satellite EO products:
In an attempt to forecast sea ice volume with the MPI-OM, it appears most beneficial to use either SIT or RFB as EO product, compared to SIFB.
If one attempts to forecast snow volume, the results are different: it is most beneficial to use RFB, while SIT lead to highest uncertainties. SIFB appeared to be in the middle.
Second, using the hypothetical products:
The authors conclude, that using a hypothetical LFB product with low accuracy is better (for both SIV and SNV) than using SIT but could not reach the performance of RFB. Improving the accuracy of the LFB product improves the performance. Using an approach where any of the above EO products is used in combination with snow depth products leads to improved performance. Again, EO products with higher accuracy lead to improved performance.

As such, I consider the work the authors introduced to be a novel and valuable contribution in the process of optimizing the use of EO products in reanalysis and thus in prediction frameworks. However, I consider the presentation of the work poor, which strongly hinders an easy approach.

The manuscript lacks conciseness and does not follow basic rules of scientific writing. For instance, notions are either wrongly introduced (such as the Jacobian), or not explained , such as M and N or the Jacobians" or the perturbations, which appear to be crucial in the QND formalism. The explanation of the basic equations are erroneous and in the introduction of the sea ice-ocean model MPI-OM it is explained, that this model consists of the equation of the ocean while neglecting the sea ice. It is added later in the text. A reader not familiar with the set of equations will be confused. There are partly wrong explanations widely extended of topics irrelevant for the understanding of the proposed algorithms of the manuscript while relevant explanations are missing. Moreover, the captions of Figures do not (sufficiently) explain the graphs, graphs are lacking labeling of the axes, units are lacking, captions do not fit with the graphs/tables; Figures are neither properly explained in the text. The discussion of results (most likely shown in the graphs) lack references to the graphs at all, and if they refer to a graph (which might be quite complex), they do not explain, which bar and which of the many boxes in the graph they are referring to. This makes the argumentation very hard to follow.

There is a lot of jumping within the graphs, which are spread over the entire manuscript, such that the reader often finds himself in searching the graphs/tables, than in following the argumentation. I suggest to move them all to the end of the manuscript.

Moreover, the authors introduce the QND formalism, but in the development of the text it is not clear, what is precisely done. There are some indications on the procedure, for instance on how sensitivities are derived. It is not clear (for instance), how and when the EO products or the information on uncertainties are incorporated into the QND formalism.

Due to the poor/sloppy form and logic of the paper, I may have missed some principal issues that will appear better in a reviewed version of the paper.

We are glad that the reviewer recognises the novelty and value of the manuscript and appreciate the effort he/she put into further improving our manuscript. An iteration with the editorial office revealed that the reviewer inspected an earlier version of the manuscript, rather than the version published in TCD. As a consequence some of the reviewer comments were already addressed in the TCD manuscript. For a few comments we failed to identify the location in (either version of) the manuscript they refer to, and we agreed with the editorial office to ignore those. Furthermore the reviewer criticises presentation aspects that are out of the authors' control, because we need to follow the journal's guidelines. For example, the editorial office had confirmed that the Figures must not be moved to the end of the manuscript. We'll list more examples below. We hope that without these complications the reviewer would have come to a better rating of the presentation quality and, hence, the manuscript overall (given that the non-public part of the report in the journal's web interface explicitly states "Please note that this rating only refers to this version of the manuscript!"). These complications also render part of the reviewer's very long list of comments difficult to address. Having said this, we would like to stress that many of the reviewer's comments are very helpful and have led to significant improvements of the manuscript's readability (see detailed response below). We also note that sometimes different comments address similar questions. In such cases, in order to be concise and avoid redundancy, we tried to refer to responses already provided instead of repeating responses. Often this resulted in forward references, as, in writing the response, we moved backwards from the specific comments to the general comments.

Specific comments:
Comments on the arrangement of the manuscript
The current sectioning of the article is:
1. Introduction
2. Methods
2.1 QND
2.2 Target Quantities
2.3 Model
2.4 Control Vector
2.5 Data Sets and Observation Operators
3. Target and Observation Jacobians
4. Sea ice and snow volume uncertainty (Rename: Uncertainties in the target values")
5. Discussion
6. Summary and conclusions

This is unfortunate. For instance, in the methods subsections the authors use terms (such as the model", the control vector, ...) before introducing them. I suggest to first introduce the QND formalism, then to introduce the model, followed by the Data Sets and Observation Operators, Control Vector and Target Quantities. Beside, the model section (as also mentioned below) contains topics, that should be shifted into a separate section that contains a concise description of the experimental setup. This is missing so far. Yet, it is not clear to me, why hindcast experiments are discussed in this section. This is definitely not part of a model description and should be moved into a section, where results are presented and discussed.

The order of subsections of section 2 was deliberately selected. We first present the QND formalism in an abstract way (with all relevant terms: target quantity, model, control vector, Jacobians, mappings M and N). Then in 2.2. we specify the target quantities for our study, i.e. we start from our objective. When this is formulated, we can present in section 2.3 the numerical model we are going to use and can refer to the target quantities, to judge whether the model in appropriate. Based on the description of the model, we can describe the control vector (which depends on the model). Our goal is to minimise the uncertainty in the control vector through observations, so 2.5 follows naturally.

A set of clarifications (also in response to the detailed comments below) are inserted to support this logic. For example, to clarify that 2.1 takes an abstract point we included the following clarification (revised text shown below with response to content comment 4c).

And we have changed the section title of section 2.3 to "Sea ice-ocean model" to stress the distinction from the abstract model (introduced in section 2.1).

Section 3 belongs also into a section regarding the experimental setup. In such, it should also stated clearly (among a concise explanation of what and how the authors perform in the QND formalism), that and how hindcast experiments are performed and assessed. The authors should also consider to properly introduce M and N and what they call Jacobian, as these appear to be crucial part of the algorithm.
I suppose, that section 4 is meant to be a discussion on results of the QND scheme. If so, then it should be named along that line.
The authors mention that the mean state is of little importance although it obviously impacts the derivatives: the model bias is not accounted as model uncertainty and should lead to even more optimistic benefit analysis, even with larger control vectors.
This issue should be flagged upfront and in the discussions of the results.

For the experimental setup we have introduced a dedicated section (by splitting off the start of the results section) In the context of our manuscript, with focus on evaluation of EO data sets, the experimental setup consists in the description of the cases we investigate.

Section 3 takes an intermediate role: The Jacobians are a component of the QND system so they could have been presented under the method section. On the other hand they are interesting objects of study on their own. This is why we dedicated a separate section to their presentation. As mentioned above, all relevant terms (including M, N and their Jacobians) are introduced in 2.1. along with their symbols. Even if we later justify, why we have not merged our estimates of the model error contribution into the uncertainty of the target variables, but prefer to report it separately, we need to have the complete equation in section 2.1, so the reader knows where and how model uncertainty contributes.

Content

1. *Referring to abstract, l.7 and throughout the paper: It is not clear, what you exactly did in your experiments.*

   We hope the responses to the comments clarify this. Part of the problem may also be attributed to the fact that the reviewer did not read the TCD version of the manuscript (see above).

2. *Introduction p.3 l.15f: Do not refer to results in this paper in the introductory part! This section is dedicated to the documentation of already existing work and for motivating the content of the manuscript at hand. Instead of referring to your own (unpublished) work of this manuscript, cite (published) articles supporting your suggestions. If there arent any, I suggest you to reformulate your statements as hypothesis and provide reasons/indications for its validity.*

   The introduction of the TCD manuscript does exactly what you suggest: *...for doumentation of already existing work and for motivating the content of the manuscript at hand.*. No unpublished own work is referred to, and the problem is formulated.

3. *p.4 l.12 ff: I would slightly restructure the enumeration to something like (which you could refer to these by naming or referring to the numbering):*
   1. Structural uncertainty: caused by the representation of individual processes and their numerical implementation.
   2. Parametric uncertainty: of the constants in the parameterization of these processes
   3. Boundary value and forcing uncertainty: of relevant processes, e.g. uncertainties in the forcings such as surface winds or precipitation.
   4. Initial state uncertainty.
   In the following I would also rename "factor" as "uncertainty type". E.g. in l.19: it could be rephrased along the line: "The choice of the control vector is subjective. A good choice should take into account all input uncertainty categories (2. to 4. in the upper list)"

   To clarify we have revised the wording and use "category" for the above uncertainty types 1-4 but "input quantity" for the components of the control vector (of which more than one typically fall into any given category). We prefer to first describe the category and then (where applicable) define a name for it.

   As mentioned, the QND formalism performs a rigorous uncertainty propagation from the observations ~~via the control vector~~ to a target quantity of interest ~~through a dedicated modelling chain. Hence, it is worth recalling the four influence factors which produce~~ relying on the indirect link from the observations to the target variables established by a numerical model. We distinguish between four sources of uncertainty in a model simulation:

   (a) Uncertainty caused by the formulation of individual process representations and their numerical implementation (structural uncertainty).

   (b) Uncertainty in constants (process parameters) in the formulation of these processes (parametric uncertainty).

   (c) Uncertainty in external forcing/boundary values (such as surface winds or precipitation) driving the relevant processes.

   (d) Uncertainty in the state of the system at the beginning of the simulation (initial state uncertainty).

   The first ~~factor~~ category reflects the implementation of the relevant processes into the model (code) while the others can be ~~understood as~~ represented by a set of input

quantities controlling the behaviour of a simulation using the given model implementation. The QND procedure formalises the selection of these input quantities through the definition of a control vector, x. The choice of the control vector is a subjective element in the QND procedure. A good choice covers all input ~~factors~~ quantities with high uncertainty and high impact on simulated observations d_{mod} or target quantities y (Kaminski et al., 2012; Rayner et al., 2016).

4. *Be more concise and introduce the notions and used quantities and mechanisms thoroughly:*

   a) *p.4 l.26: Clarify what the "observational information" is. Is this the uncertainties in the observations?*

   We think at this point the general phrasing is fine, later in that section we'll be more formal. Also note response to phrasing comment 37.

   b) *p.4 l.28ff: A motivation for the use of the PDF covariance matrices, the assumption of their Gaussianity is lacking. Where is it used? Explicitly in the backpropagation step? As well, you have constants in the control vector, dont you (see Table 1, rows 1-31 out of 45)? How are they transformed into the required structure?*

   Indicate, how the PDF covariances are constructed . In this section it could be referred to Section 2.4 Control Vector. In that section (2.4), it should be mentioned, how the PDF covariance matrix is build for each type of entry. Currently, in this section it is explained, that a perturbation is added to the fields themselves and all the discussion is about the fields, but not about the control vector itself. This is confusing. Beside, it is lacking, which law the perturbations follow the N(0,sigma) would be a natural choice, but it is not mentioned, neither the size of sigma. Motivate the necessity of the perturbations.

   We hope that our above explanation of the logic behind the order of the sections (first general then specific) answers most of the difficulties. Also the (slightly revised) section 2.4 on the control vector (see response to phrasing comment 33) should (now) be sufficiently clear.

   c) *p.5 l3: "For the first QND step we use the model M as a mapping from control variables onto equivalents of the observations." - It is unfortunate to say "the model M" without introducing it before. If M is just the mapping from control variables to the observational space, then it might be better to write: "In the inverse step we use a mapping from the control vars onto the observational space. In the upcoming we refer to this operator as the model M."*

   Manuscript revised as follows:

   For the first QND step we use ~~the model~~ a mapping M ~~as a mapping~~ from control variables onto equivalents of the observations. In our notation the observation operators that map the model state onto the individual data streams (see Kaminski and Mathieu (2017) and Section 2.5) are ~~absorbed~~ incorporated in M. Here we refer to M as model.

   d) *In Section 2.1 QND explain concisely the role of the control vector, what the outcome for the target vector is dependent on the observation products and their (which?) additional information. Moreover, you list the sources of uncertainties for the model, but not for the EO products. Elaborate on these as well!*

   After the explanation of the terms in equation (3) on page 5: It is not clear, how the forecast/assimilation is involved. It is not clarified throughout the manuscript.

   It might be beneficial to introduce N more properly. It is not really clear to me, which role the control vector plays at this stage, not how it is involved in the QND structure.

   While M is a mapping from the EO product to the model equivalent, I guess, that the ocean ice model is already somehow involved here and some of the parameterizations etc (see uncertainty types) are involved (explanation, how this is done, is missing).

   In step one you thus estimate the sensitivities of this mapping (how?). In the second step, you basically aim to assess the propagation of the uncertainties within the sea ice ocean model (how?), if I understood you right. As an outcome of this step 2 you also get an estimate of the uncertainty quality of the model parameterization on the uncertainty of the target quantities. It is not clear to me, how/if the EO products are incorporated into the process.

   Particularly, it is not clear, how the scheme as sketched in Fig.1 is related to the procedure as sketched in Fig. 12, which comes into play without any motivation.

   These questions should be clarified.

   Most of these questions are clarified in the responses to other comments. Role of control vector: phrasing comment 33. Incorporation of EO products: phrasing comment 37.

   An elaboration of the sources of observational uncertainty is beyond the scope of this study, as long as it does not come in in the section of the retrieval chain between the rawer and higher level products we evaluate. Such sources of observational uncertainty are discussed in section 2.5.

The forecast of target quantities depends on the problem at hand, for the present problem it is explained in section 2.2.

MPIOM (including parametrisations) is presented in section 2.3, observation operators (including parametrisations) in section 2.5.

In step 1 we do not estimate the sensitivity of M, i.e. M' but we use it. The approximation of M' is presented in section 3.

e) *p.6 l.5: In the QND it is mentioned that there are two models involved: represented by the operators N and M. Moreover, in Section 2.3, a sea ice-ocean model is being introduced, that seems to be not incorporated into the QND (see the definition of M and N). This is confusing. A clarifying explanation on this is strongly desired.*
Moreover, the authors mention, that it is crucial to have a realistic propagation of the sensitivities of the uncertainties to the target quantities (via both, N and M, I guess), instead of a realistic representation of the simulation of the target quantities. I do not understand, how these two are disconnected. In particular, the authors compare model output with EO products, (see e.g. Fig.6-9) which contradicts their own argumentation. This needs to be clarified.
How do the authors access that the sensitivities are represented realistically?
4f) Figure 2: caption: Explain what it is seen, what are the shaded lines, what the darker? What do the x-axis and the y-axis represent? What are the units? Why are there two d_i involved and how and why at different time steps? This is explained neither in the caption nor at any point in the manuscript! What is contained in $C(d_i)$, what in $sigma(y_i)$?
These are basics. The graph is not self-explaining and does not help the reader to understand the graph nor the algorithm.
This confusion also occurs in p.7 l.7, where it has not been clarified beforehand, how the observations are incorporated into the "model" (whichever model). In the abstract you also talk about forecasting. How does this agree with a scenario which appears to be a reanalysis scenario? How is this Figure 2 connected to Figure 1 and how to Figure 12?

Link between MPIOM and equations of section 2.1 explained in revised first sentence of section 2.3:

~~The requirement on the dynamical~~ To simulate observation equivalents (M in Equation (1)) and target quantities (N in Equation (3)) we employ a coupled model of the ~~coupled~~ sea ice-ocean system~~is that it simulates in a realistic manner~~. The model is required to provide realistic simulations of the sensitivity of ~~the~~ observation equivalents and ~~the~~ target quantities to changes in the control variables.

Need for realistic model sensitivities: See response to comment 9. For clarification we also added an example:

To conduct a valuable QND assessment, the requirement on the model is not that it simulates the target quantities and observations under investigation realistically, but the requirement is that it provides a realistic *sensitivity* of the target quantities and observations under investigation with respect to a change in the control vector. ~~If these sensitivities ,~~ (As a hypothetical example we can think of a perfect regional tracer model that is run with an offset in the initial or boundary conditions for a passive tracer. The simulated tracer concentration will carry this offset, but all sensitivities will be perfect.) If the sensitivities of the target quantities and observations (i.e. the Jacobians~~,~~) are realistic, but the simulation of target quantities and observations incorrect, we can always make a valuable QND assessment with appropriate model uncertainty.

More detailed caption for Figure 2 was provided above (with changes to the manuscript pasted in) in response to a comment by reviewer 1. See also response to phrasing comments 20 and 25 (on change of symbols).

See response to phrasing comment 37 on inclusion of the observations and observation operators.

The forecasting scenario is described in section 2.2.

5. *Deducing from (5), where you define the uncertainty reduction as $(sigma(y0)-sigma(y))/sigma(y0)$, the posterior target uncertainty in equ. (4) is not $sigma^2$ but sigma! Moreover, it is confusing, that in the text above you mentioned, that you do not consider $sigma(ymod)$, and come up with it here.*

There is no role for $\sigma(y_{mod})$ in the formalism before equation 3. We added (before the equations that provide the squares of $\sigma(y)$ and $\sigma(y_0)$) a "via" to clarify that the square root has to be taken. For $\sigma(y)$ the resulting text change is shown with response to "Phrasing comment" 22.

6. *p.7 l.26: Here it is said that predictions are performed, but from the preceding it appears that (in some way) the incorporation of the EO products into the model appears in a reanalysis framework (see e.g. Fig. 2). It is not clear, how the QND procedure fits with the argumentation. What I make up from the preceding is that in some way you will use different types of observations and will get different SNV and SIV . If so, it is not clear how uncertainties/sensitivities are then derived. The entire procedure needs clarification!*

   We hope the clarifications we added in response to the other comments (on inclusion of observations, forecasting, etc...) have resolved these difficulties.

7. *Section 2.3 Model: The detailed explanation is not of relevance for the purpose of the manuscript. It is not relevant to explain, what an ocean-sea ice model is, and what the particularities for MPI-OM are. Just refer to Jungclaus et al. (2013); Niederdrenk (2013). Beside, the description has parts which are seriously wrong:*

   - *p.8 l.7 ff: A short explanation: Due to the complexity of the 3D Navier-Stokes equations, it is common practice to apply a couple of approximations, such as the the hydrostatic approximation or the Boussinesq approximation. You can skip that information, this is nothing special. What follows is incorrect and should be skipped due to the already mentioned non-explicitness of the MPI-OM with respect to the primitive equations and an equation for the balance of the thermodynamics.*
   - *Particularly, you introduce the MPI-OM by saying, that is consisting of the three balance equations which are solely related to the ocean (without mentioning) while skipping the second set of equations for the sea-ice component.*

   If you really want to make a distinction, then cite the articles related to the ocean models and those related to the ice models. You can discuss the relevant parts (like snow loading treatment in the discussion section, as you already do) when it is needed (and refer then in the discussion to the literature). Also, the discussion of the mesh is unnecessary. If it is really necessary (which I do not see) I recommend to mention the structure in short and provide a source. If there is anything particular you implemented due to the necessity of the algorithm, then mention it along the line "In addition to the standard MPI-OM we implemented... in order to ... based on [literature]".
   The part starting from p.7 l.30 to p.8 l.5 is OK. If I understand the authors correctly, then they use the last sentence in there to justify/indicate that the MPI-OM gives realistic dependencies. If this is the case, then I would formulate exactly this e.g. by "Thus, we consider the model results to be reasonably realistic." The remainder of the model description should be removed.

   Not all readers of the article are familiar with the MPIOM (and its development status), hence we consider a short presentation of the model relevant. Nevertheless, this model description has been shortened (even though it is unclear which parts the reviewer considered "incorrect"). We skipped the very general part about the MPIOM description but maintain the part about the recent development of MPIOM and the brief description of the ocean model because we think that it is essential for the readers to have some idea about the implemented processes.:

   ~~MPIOM is based on the primitive equations, a set of nonlinear differential equations that approximate the oceanic flow and are used in most oceanic models. They consist of three main sets of balance equations: A continuity equation representing the conservation of mass, the Navier-Stokes equations ensuring conservation of momentum, and a thermal energy equation relating the overall temperature of the system to heat sources and sinks. Diagnostic treatment of pressure and density is used to close the momentum balance. Density is taken to be a function of model pressure, temperature and salinity (UNESCO, 1983). Recent development of the model~~ Recent development of the ocean part of the model includes the treatment of horizontal discretisation which has undergone a transition from a staggered E-grid to an orthogonal curvilinear C-grid. The treatment of subgridscale mixing has been improved ~~by~~ through the inclusion of a new formulation of bottom boundary layer slope convection, an isopycnal diffusion scheme, and a Gent and McWilliams style eddy-induced mixing parameterisation ~~. Along-isopycnic~~ (Gent and McWilliams, 1990). Along-isopycnal diffusion is formulated following Redi (1982) and Griffies (1998). Isopycnal tracer mixing by unresolved eddies is parameterised following Gent et al. (1995). For the vertical eddy viscosity and diffusion the Richardson numberdependent scheme of Pacanowski and Philander (1981)

is used. An additional wind mixing proportional to the cube of the 10-m wind speed (decaying exponentially with depth) compensates for too low turbulent mixing close to the surface. Static instabilities are removed through enhanced vertical diffusion.

A viscousplastic rheology (Hibler, 1979) is used for the sea ice dynamics. ~~The thermodynamics is~~ Sea ice thermodynamics are formulated using a Semtner (1976) zero-layer model relating changes in sea ice thickness to a balance of radiant, turbulent, and oceanic heat fluxes. In the zero-layer model the conductive heat flux within the sea ice/snow layer is assumed to be directly proportional to the temperature gradient across the sea ice/snow layer and inversely proportional to the thickness of that layer, i.e. the sea ice does not store heat. The effect of snow accumulation on sea ice is included, along with snowice transformation when the snow/ice interface sinks below the sea level because of snow loading (flooding). The effect of ice formation and melting is accounted for within the model assuming a sea ice salinity of 5 psu.

Regarding the resolution, we have included in the text why this is important:

This setup achieves a spatial resolution as high as that of the EO products we analyse (in fact over the target regions the model resolution is higher) without major computational constraints, which allows an evaluation of the full spatial information content provided by the respective EO products. Here, we will refer to this particular model configuration as Arctic MPIOM.

8. *Remark on Section 2.3 Model: I understand that in this section the authors introduce the model and refer to related literature, introduce the forcing (though it should be indicated in the Section title as well). Starting from p.10 l.11, the authors describe the initialization of the MPI-OM. This belongs to the presentation of the experimental design. I suggest to separate the experimental setup from the description of the model. I suggest to dedicate a separate section with a clear description of the experimental setup, starting from initialisation, perturbation strategies of the control vector variables, etc.*

   As mentioned above we agree on an extra section for the experimental setup, but it addresses the observational cases we investigate. The ocean model, including its setup is regarded as a component of the system, the components of which we describe in section 2.

9. *p.10 l.20ff- until the end of the section: A motivation of the upcoming paragraphs is missing and I do not see the point why it is placed in the model description section. Place it into a different section with an appropriate title. Moreover, if you aim to present an assessment of the MPI-OM hindcasts due to observations and a discussion on their uncertainties, then indicate this in the abstract and motivate this in the beginning of a possible new section, where you perform this discussion.*

   We added a motivation for the validation part:

   For a successful QND assessment it is essential that MPIOM provides realistic sensitivities of the observation equivalent and the target quantities to the changes in the control vector (Equation (1) and Equation (3)). However, observations are not available to validate these sensitivities. The only validation of MPIOM possible is against observations of the state of the sea ice and ocean. In the following we present comparisons with selected observation based products first for the hindcasting period, and then for assimilation window and forecasting period.

10. *Alternatively, the authors could shortly indicate, that they consider the MPI-OM to represent the physics well, and present a summary. At this point this is not clear, how this discussion is related to the QND.*

    See response to item 9.

11. *p.10 l.26 and the discussion related: In earlier passages, the authors stated, that they are not interested in the realm of the model results, but rather in the sensitivities. This is not reflected/discussed in the comparison of concrete values against observations.*

    See response to item 9.

12. *p.10 l.28: "only small misfits": you should exclude the marginal ice zones out of this, as I consider a misfit of about 50% as noticeable. And it could be explained by stronger transport and errors in the advection schemes. As well, it is possible that in those regions there are different (weaker) tolerances in the accuracies of the observations.*

    We revised:

In March (panel d) and June (panel e) only ~~small~~ relatively small scale misfits to the OSI SAF ice concentration are found but they can reach up to 50% (here and in the following we use the term "misfit" for the model-observation difference)

We don't want to speculate about the reasons but just want to describe the performance of MPIOM.

13. *p.11 l.3: it is not clear to me that you look at hindcasts. Clarify this beforehand, for instance in a separate section explaining the experimental setup.*

    We introduced the hindcast in the paragraph about the model initialisation. We rephrased the beginning of the following paragraph:

    The hindcast with Arctic MPIOM has been validated against remotely sensed ice concentration from the reprocessed ~~OSI SAF~~ Ocean and Sea Ice Satellite Application Facility (OSI SAF) sea ice concentration product ...

14. *p.13 3f: How much sense does it make to compare multi-annual means in a period of sea ice decline? Is the interdecadal trend insignificant?*

    Indeed the value of a comparison of the mean state is limited in a strongly changing climate but we think that a more detailed validation is beyond the scope of this paper. For the QND approach only the state in April and May 2015 is of relevance which we discuss in Figure 8 with respect to SIT and in Figure 9 with respect to SND.

15. *p.15 l.2: Describe where the uncertainties are derived from and how.*

    This is exactly what the section does, it describes the PDF of the control vector, i.e. mean and uncertainty:

    For process parameters this standard deviation is estimated from the range of values typically used within the modelling community. The standard deviation for the components of the initial state is based on a model simulation over the past 37 years and computed for the 37 member ensemble corresponding to all states on the same day of the year. Likewise the standard deviation of the surface boundary conditions is computed for the 37 member ensemble corresponding to the April-October means of the respective year.

16. *p.15 l.5: If you want to be indepth: you could explain, why it is numerically cheaper to divide big vectors into several smaller ones. Or is it rather due to the fact, that it is beneficial to get to know where the uncertainties stem from? At least this was the impression in the extensive argumentation that comes later in the manuscript.*

    We added a bit of detail:

    The largest possible control vector in our modelling system is the superset of initial and surface boundary conditions as well as all parameters in the process formulations, including the observation operators. As described in section 3, the Jacobian computation requires an extra run for each additional component of the control vector. To keep our ArcMBA system numerically efficient, two and three-dimensional fields are partitioned into regions.

17. *p.15 l.12ff: How can uncertainties have diagonal form? It looks like what you mean by uncertainties also contains information about cross-correlations between the different control variables. Elaborate more on that, or repeat it here in a concise way. Otherwise it does not make sense. Uncertainties themselves will form no matrix but a vector.*

    Response to phrasing comment 20 should have clarified this.

18. *p.20 l.1 : What is the retrieval chain and how can this (as well as Fig. 12) be brought into agreement with the QND formalism introduced in Fig. 1. This section lacks explanation on how this incorporates into the QND formalism.*

    See response to phrasing comment 37.

19. *p.20 l.5: the Jacobian is a matrix which contains derivatives. This I do not see reflected in the right hand side of Fig. 12. What I see is that the observational equivalents of the left hand side products are being derived and so it seems compared. Maybe, sensitivities is a better word. Anyhow, I do not see this reflected in the graph. If it is a Jacobian, it could be useful to give a formula.*

    Revised formulation to:

The right-hand side of the graph illustrates how ~~this Jacobian is derived from the Jacobians of the~~ the equivalents of the respective products are simulated from the relevant model variables, which are denoted in violet colour.

20. *-p.20 l.7f: I do not see where you derive variables that describe the changes in the variable (due to changes in the control vector) this is why you have the control vector, right? Moreover, the comparison regarding the complexity is not clear and should be explained.*

    The observational Jacobian M' (sensitivity of observation equivalent with respect to control vector) is described in section 3. For incorporation of EO products into M see phrasing comment 37. The complexity refers to extra computations that require extra input, as described in the next sentence.

21. *p.20 l.12f: "SIT refers to the grid cell average, i.e. for the Jacobian...": grid cell average vs dividing by SIC is not coherent to me. Please correct.*

    Is is common practise that observers define sea ice variables on a grid as the mean over the ice-covered grid cell while modellers define the mean as the average over the grid cell. The model analogue of the former quantity can be calculated by dividing the latter quantity by the sea ice concentration in the grid cell, i.e. by SIC. No need for any correction of the text, we think.

22. *p.23 l.14: relating to the Beaufort Gyre: If this is the case, shouldnt there be then a negative correlation seen in those regions, 7 or 8?*

    We added:

    WIX is positive for eastward wind stress. A positive perturbation on WIX is most distinct in region 6 (but also evident in regions 7 and 8) and slows down the Beaufort Gyre which advects less sea ice into the target region (sea ice behaves, at least in April and May, to a large extent like a rigid body, i.e. the impact in regions 7 and 8 acts almost instantaneously on the target regions) resulting in a negative sensitivity.

23. *p.31 l9: it is not clear how your assessment is linked to this forecast. When did you apply your QND framework? In which period? How did you treat the nonstationarity?*

    We did it for one specific period as described in section 2.2.

The language could be improved throughout the manuscript. Here, I give suggestions to some of the parts, which I considered most worthy to be improved. The author should consider to use short and flat structured sentences.:
- abstract l.10: remove the institutes name, it appears awkward, just all derived from CryoSat-2".
See response to technical comment 20.

Phrasing&Structuring

1. *Abstract l.7 "observation impact (added value)": replace by "added value of observations" or "We assess the added value of different EO data products in terms of ..."*

   "observation impact" is a standard term in the data assimilation community (see, e.g., Todling (2013)). One occurrence had been removed already in the TCD manuscript (see below).

2. *Abstract l.9: "the assessments cover" replace by "We assess seven..."*

   Done as suggested.

3. *Abstract l.11f: concerning the phrases in brackets: I suggest to replace both by "(low and high accuracy)" each.*

   Done as suggested.

4. *Abstract l.20: "Providing" instead of "the provision of".*

   Done.

5. *p.2 l.7: Mention that this forecast is done with one particular model, namely MPIOM.*

   Done.

6. *p.3 l.7: Write instead "Forecasts of the ice and the ocean state are ...", as the sea ice-ocean models not only contain equations describing the dynamics of the system, as you also introduce later in the manuscript.*

   Done as suggested.

7. *p.3 l.8: A minor suggestion: formulate in a positive way: "In order to derive reliable forecasts, uncertainties in the model initial state, of the atmospheric b.c.s and in the parameterizations of physical processes should be minimized."*

   Done as suggested.

8. *p.3 l.9: remove "only". For instance, observations of bad quality are of no advantage. And improvements in modeling, parameterization etc. also contribute to improved model output.*

   Done.

9. *p.3 l.21: "observation impact" : change to "the impact of observations"*

   Was already changed in the TCD manuscript.

10. *p.3 l.23: optimized for what?*

    Sentence extended:

    The technique originates from seismology (Hardt and Scherbaum, 1994) and was first applied to the climate system by Rayner et al. (1996), who optimised the spatial distribution of in situ observations of atmospheric carbon dioxide to achieve minimum uncertainty in inferred surface fluxes.

11. *p.3 l.27f: "successfully demonstrated" sounds weird.*

    Replaced by "successfully applied for".

12. *p.4 l.9: Do not use control vector at this stage, it is confusing, when it is not introduced yet and does not lead to a further understanding. I would just skip it. Furthermore, maybe it is better to formulate, that with the QND formalism you are able to assess how the uncertainty propagates from the observations (raw data?) to a certain target quantity. To my mind it is not of interest at this point to add information on the modeling chain. It is just confusing.*

    Done, resulting text was shown above in response to comment 3 related to content.

13. *p.4 l.10f: I would remove "hence", as this is the 4 factors you identify. ("We distinguish 4 types of ..."). Remove "influence" at end of line 10, as it is redundant and confusing. Instead you could consider to use the phrase "sources of uncertainties".*

    Done, resulting text was shown above in response to comment 3 related to content.

14. *p.4 l.17: remove "(code)", this is redundant.*

    Yes, a deliberate redundancy to be really clear.

15. *p.4 l.22: Keep the message as short as possible to maintain comprehensibility. For instance remove "any potential model output" and replace ", for example a process parameter such as the albedo of the snow" by "(such as the albedo of snow)". The phrase "process parameter" only adds confusion.*

    First suggestion followed, second not, as we think it is useful for a better understanding of the generality of the concept to mention the "class" of quantities the albedo of snow belongs to.

16. *p.4 l.26f: A suggestion to rephrase: "In a first step, we reduce the uncertainty in the control vector by making use of a given inverse model and information (to be specified by the authors) on the observations." Then start a new sentence for the second step.*

    We split in two sentences, but prefer our wording.

17. *p.4 l.28ff: You could shorten it to "Within the QND formalism, we present all involved variables/quantities by probability density functions (PDF)." The explanation does not add new information.*

    Done, but "variables/" deleted.

18. *p.4 l.8: "based on algebra" sounds weird. I would just phrase it as "and is partly based on..."*

    Done.

19. *p.5 l.5: I would replace "absorbed" by "incorporated".*

    Done.

20. *p.5 l.7: "with covariance $C(x)$, i.e. the uncertainty is given by": There is an inconsistency. Why is $C(x)^{-}1$ the uncertainty, and why are the data uncertainties $C(..)$ and not $C(..)^{-}1$? I would rather replace that by: "with covariance $C(x)$, which is given by/defined as".*

    We have clarified our use of the term uncertainty where we introduce the PDF notation through the following additional text:

    In the context of these PDFs we will use the term uncertainty to refer to its full covariance matrix in the case of a vector quantity, and in the case of a scalar quantity or a given vector component it refers to the square root of the entry on the diagonal of the full covariance matrix corresponding to that particular vector component. In the latter case the uncertainty refers to one standard deviation of the marginal PDF corresponding to that component, and we use the notation $\sigma(d_2)$ to denote, for example, the standard deviation of the second component of d.

21. *p.5 l.12: Is "observational constraint" the correct word? Shouldnt it rather be the given uncertainty of the observations? Furthermore, to improve readability, use $C(dmod)$ instead of "the second term". Also mention here, that this is a subjective choice, instead of coming back to that 10 lines later when discussing different equations. For a better understanding, I suggest a reformulation from line 9: "where the data uncertainty $C(d)$ is a combination of two factors: [formula]. The term $C(dobs)$ expresses the uncertainty in the observations and $C(dmod)$ the uncertainty in the projection operator M. Its/Their (both?) formulation is a subjective choice.". For the formula (2) you could also shortly explanation/indication, why you used the quadratic form. I guess, the reason is smoothness and higher regularity due to the inversion step. Or do you aim to account more for larger uncertainties than for smaller? (Which is what the L2 norm does compared to the L1 norm).*

    Squares unintended (in fact left overs from an earlier version of eq 2 that was formulated in terms of the σ). We think it is o.k. to be general here and discuss the role of model uncertainty together for eqs 2 and 3. We included phrasing suggestions as follows:

    where the data uncertainty $\mathbf{C}(d)$ ~~combines~~ is the combination of two contributions:

    $$\mathbf{C}(d) = \mathbf{C}(d_{\mathbf{obs}}) + \mathbf{C}(d_{\mathbf{mod}}) \tag{1}$$

    The term $\mathbf{C}(d_{\mathbf{obs}})$ ~~with the uncertainty~~ expresses the uncertainty in the observations and $\mathbf{C}(d_{\mathbf{mod}})$ the uncertainty in the simulated equivalents of the observations $M(x)$~~:~~

    $$\cancel{\mathbf{C}(d)^2 = \mathbf{C}(d_{\mathbf{obs}})^2 + \mathbf{C}(d_{\mathbf{mod}})^2}$$

    . The first term in Equation 1 expresses the ~~observational constraint~~ impact of the observations and the second term the ~~prior information content~~impact of the prior information.

22. *p.5 l.14: replace "in the second step" by "in the propagation step" ... you already introduced that notion. As well it is now confusing, which model you consider. Better to first introduce the model and then what is done in this propagation step. A proposition: "The model N involved in the second, the propagation step, is the mapping from the control vector onto the target quantities. The Jacobian of N, (N) is used to estimate how the posterior uncertainties in $C(x)$ propagate to the" - I am confused here: before equation (1) you say, that $C(x)$ is the covariance of the Gaussian PDF of the posterior control vector. And here you say, that that $C(x)$ is the control vector. Use unique formulation.*

    Use of uncertainty clarified in response to comment 20.

    Revised phrasing according to suggestion:

    ~~In the secondstep, the Jacobian matrix $\mathbf{N}'$ of the model (now used as a~~ The mapping N involved in the second, the uncertainty propagation step, is the mapping from the control vector onto ~~target quantities and denoted by~~ a target quantity, y. The Jacobian matrix $\mathbf{N}'$ of the mapping N ~~)~~ is employed to ~~propagate~~ approximate the propagation of the posterior uncertainty in the control vector $\mathbf{C}(x)$ forward to the uncertainty in a target quantity, $\sigma(y)$ via...

23. *p.5 l.18: For improved readability, I would proceed chronologically in the order of occurrence of the terms (start from the beginning of the equations), introduce the meaning of the single terms and indicate subject choices then. My suggestion for p.5 l.18-p.6 l.4: "The first term, $NC(x)NT$, reflects the propagation of the posterior uncertainty $C(x)$ to the target uncertainties via the model N, while sigma(ymod) reflects the remaining uncertainties (see types 2-4 in the list above), that are not yet*

represented in the control vector. Like $C(dmod)$, this quantity is set due to subjective choice. In our work, we skip this term in order to sharpen the contrast between the EO products, and only mention two plausible estimates."

After the revision in response to comment 23, we think the presentation of Eq 3 reads well, no need for further modification. We do not skip $\sigma(y_{mod})$ but report it separately.

24. *p.7 l.13: "does not require real observations": This phrase is unnecessary. Instead you could just say, that the QND formalism can be used to assess/evaluate hypothetical...*

    Done as suggested.

25. *p.7 l.15: here you use d as the set of observations, and in l.7 you use d1 and d2. Introduce the definition of d (using vector notation) before you use it (or components of the vector without mentioning). For instance, Fig. 2 could be introduced after such a definition. A suggestion: First say, that it is possible to evaluate a network of observations, that do not need to have the same structure, nor be available on the same grid. In particular, this enables the study of the benefit of using hypothetical data networks. As is done in this work.*

    See response to comment 20. Figure revised from "$C(d_i)$" to $\sigma(d_i)$

26. *p.10 l.11: from a restart file a dd.mm.yyyy generated ...", remove (start time of ERA- Interim).*

    Done.

27. *p.10 l.15: The initial ocean state is assumed to be at rest, the initial sea ice...*

    We added in the manuscript that the sea ice is at rest as well.

28. *p.12 Fig 6: Explain what blue and what red colors mean! How is misfit defined? How do you assess with this comparison the sensitivities instead of real values?*

    Explanation of colours added to caption. For the rest see response to comments 9 and 12 related to "content".

29. *p.12 l.7: "is linear in time plus a quadratic time-dependent component, i.e. it does not contain year-to-year variability." this correlation is not clear to me. Explain or remove!*

    We think that the information is important to understand the assessment and refer to the Lindsay and Schweiger (2015) paper for a detailed explanation, see also response to next comment.

30. *p.12 l4: explain the ice thickness regression procedure.*

    A detailed description of the procedure would be outside the scope of the paper. We refer to the Lindsay and Schweiger (2015) paper.

31. *p.14 caption Fig. 8: Needs to be improved along the line already mentioned (Think of self- explaining!, colors, notions, etc.)*

    Done.

32. *p. 15 Add to the title of the control vector: "and Uncertainty specification".*

    The control vector is represented by a PDF, which implies that the section addresses both mean and uncertainty.

33. *Section 2.4: Give a little introduction into the purpose of the control vector. Do you gain information by using that one? What is the difference in the outcome when using a large or a small control vector? Somewhat trivial: Add, why you do not modify the control vector, while you do so with the observations.*

    We have revised the text as follows:

    ~~The definition~~ Criteria for the choice of the control vector ~~and the specification on prior uncertainty follows Kaminski et al. (2015)The components and their prior uncertainty~~ are presented in Section 2.1. The specification of prior, both mean (x_0) and uncertainty ($C(x_0)$), follow Kaminski et al. (2015), and is listed in Table 1.

    Exploring the sensitivity of the results with respect to the specification of the control vector could be the topic of a follow up study, as is mentioned in the conclusions (variations from year to year).

34. *p.15 l.2: Consider to add ", C(x0)," after "uncertainties". Moreover, I would shorten: "(2015), and are listed in Table 1."*

    Done.

35. *p.15 l.7ff: In order to avoid confusion, the part in the brackets where it is said that perturbation is added to the entire part of the simulation, should be put out of the brackets.*

    Done.

36. *p.15 l.11: Either use present tense ("results"), or reformulate: "Thus, the control vector contains in total 157 control variables."*

    We use present tense ("results").

37. *Section 2.5: This section is not understandable at all. As introduction of this section clarify where you apply the data sets and where the observation operators in the QND framework!*

    We added a sentence to the first paragraph:

    Recall that the (combination of) data set(s) enters the QND algorithm through its uncertainty $C(d)$ and that the observation operator is incorporated in the model M (see Section 2.1)

38. *p.16 l.8: when you use the word "link", you should say between what. Right now you only use from models state variables, but lack the to-part.*

    We added "to the respective data sets".

39. *p.17 check table caption against the table: column one lists the indices/place of occurrence of the quantities in the control vector, while column 2 the abbreviation.*
    To enable easier reading you could section the table in 3 parts, the first being process parameters, the second initial fields and the third forcing fields. You could remove the third column and section by horizontal lines and note the type by writing "process parameters" etc in vertical style left beside the index. Alternatively, insert additional rows that only contain "process parameters" etc as sectioning of the table.
    The last column can be removed and instead it should be explained in the caption, that the parameters are unique values, while initial and forcing are given in the control vector individually for each of the 9 regions (and refer to the figure 10 where they are introduced). Column 5 lacks units in most of the entries. Caption and head of table disagree.

    Units added where missing. Typos corrected. The last column is useful to identify the location of individual components in the Jacobian plots. But we have followed the suggestion to add horizontal lines to section the table into the three compartments.

40. *Fig. 12: What effect do the assessment boxes have? Which role do they have in the upcoming of the manuscript? Explain abbreviations in the graph, that have not been introduced yet, such as MSS.*

    The assessment boxes indicate where the model and the retrievals "shake hand". Definition of MSS added.

41. *The first time the notion "Archimedes principle" shows up, it could be shortly explained, if the author want to be self-explanatory.*

    We thought that Archimedes' principle needs no explanation in a scientific paper, but have now added a reference (to Guerrier and Horley (1970)).

42. *p.20 l.20: for consistency in notation, use formula for snow depth or write the following formulas in words, i.e. "densities of snow, ice and water".*

    Symbol for snow depth was not used to avoid confusion with modelled snow depth.

43. *p.20 l.21: add names of f_i, f_r and f_l. It has so far only once been mentioned in Fig. 11.*

    The names are introduced on p.16 l.10.

44. *p.20 l.28: motivate $-0.22hs/c$: what is this and where do you take the formulas from.*

    A motivation is given in l.28 but we rephrased the sentence.

45. *p.21, l5: remove "provided by AWI", and use: the CryoSat-2 product files used in this work.*

    Done.

46. *p.21 Caption of Fig. 13: time is missing (April 2015). l.6f: How do the uncertainties in the other times look like? I do not see how you incorporate the uncertainties into your algorithm. And: l.8: you introduced before the diagonal structure of the "uncertainties". So I would refer to that by "Recall, that we assume uncertainties to be uncorrelated in space".*

    We added the time in the caption. Observational uncertainty enters via equation (1) and (2)). Rephrased as suggested.

47. *p.21 l.10: give a justification/reason, why you use the threshold 0.7 for SIC.*

    All altimeter retrievals have problems for large open water fraction. We selected the threshold in analogy to the CryoSat retrieval. We added that in the manuscript.

48. *p.22 l.10: Does M refer to model MPI-OM? With respect to what is the derivative?*

    M' is the derivative of the simulated EO product with respect to the control vector and was defined in section 2.1.

49. *p.22 l.17: where do you derive sigma$_i$ from particularly for process parameters?*

    We had explained it at the end of section 2.4 and provided the extra text with our response to content comment 15.

50. *p.22 l.20: what is a 1-sigma change?*

    We rephrased the sentence (change by one standard deviation).

51. *p.23 l.2: It is easier if you explain, that this plot shows the sensitivities of the XXX due to changes in SIFB, LFB,...*
    What does that mean: "the Jacobian for April means of SIT over a point"? One entry in the Jacobian is: $\partial f_j/\partial x_i$. Explain, what f_j, what x_i is?.

    We had defined the Jacobian in 2.1 and had interpreted it as sensitivity in the preceding paragraph. We extended to read **a point in space**.

52. *In the caption of fig 14 clarify that each bars in the plot corresponds to the uncertainty/sensitivity (?) of one entry in the control vector due to the changes in the values XXX in the black dot! Then explain that for instance for SIT there are 4 bars for each region one for each EO product. It is very hard to read this figure without any further explanation.*

    Done.

53. *p.23 l.4: add information where you are referring your discussion to, for instance "SIT sensitivity (indicated as the XXX bars in the graph)" otherwise it is simply confusing. End of that sentence in l.6: add "in that region".*

    Done.

54. *p.23 l.9: this has not been indicated in your model description. Just give a reference here.*

    The dependence of the sea ice growth on the open water fraction is independent on the model formulation. We rephrased the sentence.

55. *p.23 l.17: "the various..." where do we see this in Fig.14? Do you still refer to this figure? Indicate which bars you are talking about! This applies for the entire section! Any statement you make refer to the corresponding bars!*

    We added:

    The various freeboard products exhibit high sensitivity to initial SIT and SND (orange, red, and green bars in Figure 14).

56. *p.23 l.28: what is the model N? Are you still in Fig.14?*

    N' is defined in section 2.1 and we write that we are on Fig 15.

57. *p.24 l.11: put "region 6" out of the brackets, as this is a particular feature of region 6!*

    Done.

58. *p.25 l.3: Is "derive" the right word? If so, say how you do this. Else, use "use"/"introduce". In any case, motivate your choice.*

    Yes, exactly: We explain immediately how we do it.

59. *p.25 l.4 remove: "and listed in the last but one row".*

    Done.

60. *p.25 l4: "model that perfectly simulates"...: where do you use this result and how?*

    Exactly here, to translate a thickness into a volume.

61. *p.25 l.6 "and listed in the last row": remove.*

    Done.

62. *caption of table 3: 4-6 are 3 columns, whereas prior and posterior are 2 values, confusing! Moreover, you could refer to the figure where they are depicted.*
    Is low or high accuracy used? Explain where you find "without additional product", "with product with low accuracy" and "product with high accuracy in the table".

    Caption explains clearly that uncertainty is given per region and target quantity, i.e. we have 2x3 values. Accuracy (or absence) of snow product in column 3 as described.

63. *p.26 l.8: better phrasing (and indicating what you are referring to): "the performance of SIFB (bars with magenta color in Fig. 16) is similar for ".*

    Was already revised in the TCD manuscript.

64. *p.26 l.11: Figure 14: ...green bars in (?). explain what you are exactly comparing! This applies for the entire manuscript and I will not further mention any further occurrences.*

    We added to the caption:

    The sensitivities of the respective EO product to the control vector ("observational **Jacobian**~~rows~~") **for** ~~a~~ **April means of** ~~**SIT, SIFB**~~LFB (orange bars)**, RFB** (red bars), SIFB (green bars), SIT (black bars) **and** ~~**LFB**~~ SND (cyan bars) **over a single point indicated by the black dot (and by** ~~yellow~~ black **cross on Figure 3).** The observational Jacobians with respect to the process parameters are shown in the left middle panel. The other panels show the observational Jacobians with respect to the initial and forcing fields (see **Table 1** for an explanation of the abbreviations).

65. *p.26 l.17: "has so good performance already": and l.20 "the first thing to note", l.22 "with uncertain assumption primarily": improve phrasing.*

    We rephrased: ~~**The**~~ This imbalance is lower for the **high accuracy LFB product**~~**has so good performance on SIV already,**~~, because this product already performs excellently on SIV such **that there is not much scope for** ~~**yet better**~~ further increases in **performance on SNV.**

    and

    ~~**The first thing to note is that the step**~~ First, we note that switching **from SIT to SIFB drastically reduces the performance for SIV.**

66. *p.26 l.20: which step? In which procedure? Refer to figure.*

    We replaced "the step" by "switching" to prevent confusion with the two-step procedure of QND formalism.

67. *p.26 l.23: (right hand side of Fig. 12) instead of on the modeling side of Fig....*

    Done.

68. *p.27 l.20 Remove "We need to" and "here". And put "(Equation (2))" at the end of the sentence!*

    Done

69. *caption fig 14 and 15: write instead the dependencies/sensitivities of xxx to xxx. For instance it looks like in Fig. 14 you depict the outcome of step 1 (inverse step, see your Fig. 1) meaning the sensitivities of the control vector to the EO products, while in Fig.15 you depict the sensitivities of the target variables to the control vector. (forward step 2 in your Fig.1) could that make sense?*

    Yes. Caption adapted.

70. *caption Fig.16: Uncertainty reduction due to what? Explain the different bars, the different color codes.*

    Done.

71. *p.31 l-5 which setup do you mean? Regarding the spatial resolution: It is clear, that it is finer than the target regions... why do you mention that here?*

    The setup of MPIOM (we added "of MPIOM" in the manuscript). By mentioning the size of the target regions we want to make clear that the sensitivity of the target regions is aggregated over many model grid boxes, and small-scale effects are averaged out.

72. *p.31 l.12: what does that mean that you are not resolving changes in the initial conditions? Does that mean that in the considered period of integration, the model state does not develop that much away from the initialization? Furthermore you emphasized several times in the manuscript that you are not interested in the real state but in the realistic representation of sensitivities. How does that fit here?*

    We are talking about the control vector. The initial and boundary condition have the full temporal and spatial scales included. The perturbations to the initial and surface boundary condition are per region, though. The model state can develop freely away from the initial state in response to the surface boundary conditions. We have not stated that we are not interested in the "real" state (we discussed the "real" state of the model in section 2.3) but we stated that the "real" state does not enter the QND formalism directly but only via the model sensitivities (which have some dependence on the "real" state, of course).

Technical corrections - compact listing of purely technical corrections, typing errors etc.

1. *Articles are lacking in many places, such as in (p.3 l.31), (p.4 l.28), (p.15 l.13), (caption in p.19), (p.23 l.4), (p.23 l.14), (caption of table 3), (p.25 l.21 and l.35), (p.26 l.2 and l.6), (p.33 l.14).*

   Difficult to follow as line references do not refer to TCD manuscript, some spots we could not identify (e.g. p.25 l.21 and l.35). Among those spots we could identify in the TCD manuscript, often articles were already present, in some cases we found that inserting an article not useful (p.3 l.31, p.4 l.28, p.15 l.13, (caption of table 3)), and in other cases we have inserted articles, we'll see with the copy-editor ...

2. *Check for doubling of words such as in p.32 l.14 (than than) and in p.13 l.6 (the the), p.24 l.8: "compared".*

   In all cases except for "compared compared' the TCD manuscript was already correct.

3. *Check commas, they are missing in several places, such as in: (p.4 l 9: as mentioned, ..."), (p.5 l.5: In this case, ..."), (p.14 l.9), (p.16 l.8: after In the following"), (p.20 l.12: after SIC in the brackets), (p.20, l.2 after "assessment").*

   Done.

4. *Fullstops are missing: end of eq (6) , (7), and eq. (10), and p.20 l.29.*

   Done.

5. *Put the Tables and Figures all at the end of the manuscript. The authors jump a lot back and forth between their Figures and Tables, some of them are placed in sections that are unrelated to the Figures/Tables. Having them all in one place would make it easier to follow the argumentation.*

   This would be incompliant with the journal style.

6. *Addresses of authors should be consistent in their structure. For instance, (1) has street name, while others only list the town and the country. Address (3): Danish writing of Copenhagen, which should be changed to English.*

   Done.

7. *Abstract l.21: clarify the abbreviation EO when used the first time.*

   Done.

8. *p.3 l.9: typo: parametrisation: correct to parameterization (or to parameterization if BE is used).*

   Parametrisation was no typo, see here for BE (https://dict.leo.org/englisch-deutsch/parametrisation). We have changed to parameterisation but will check with the copy-editor.

9. *p.9 l.10: "Recent EO products".*

   We think this refers to p 2. Switching to "recent" would change the meaning.

10. *p.3 l.12: "The constraints" (plural).*

    Done.

11. *p.4 l.19 and in other parts: use vector notation for vectors, such as the control vector.*

    We'll see with the copy editor.

12. *p.5 l.16: insert comma before sigma (y). otherwise sigma could be understood as target quantity.*

    ... is employed to ~~propagate~~ approximate the propagation of the posterior uncertainty in the control vector $\mathbf{C}(x)$ forward to the uncertainty in a target quantity, $\sigma(y)$ ~~:~~ via

13. *p.10 l.34: no new paragraph.*

    Done.

14. *p.10 l.34: "underestimates" instead of "is underestimating".*

    Done.

15. *p.10 l.35: "target regions".*

    Done.

16. *p.11 fig5: The figure does not add relevant information to the paper.*

    We think it is useful for the reader to get an impression of the spatial variation in the resolution of the model. On the relevance of the resolution we have commented above. But we have removed the "arctic zoom" panel.

17. *p.13 fig 7 (a-b) the color map is unfortunate. The reader does not see a lot of differences.*

    Maybe a problem with the printer? We had included isolines with a distance of 0.5 m to support readability, so we think they should be o.k. and will also cross-check with the copy editor.

18. *p.13 l 2 and p.23 l.11: no new paragraph.*

    Done.

19. *p.15 l 2: "prior uncertainties".*

    Added symbol to clarify.

20. *p.16 l.6: remove "by the AWI", this is not relevant here and does not follow common rules. Instead move "(Rickers et al. 2014)" after "Cryosat2 mission". Also remove "by AWI" in l.10.*

    To our knowledge there exist three different CryoSat-2 products (respectively derived by AWI, UCL, and NSIDC) and we would like to make clear which product we used.

21. *p.19: Grey coloring not explained in the caption.*

    We added the word "emphasise" for the other colours to make clear that grey has no special meaning.

22. *p.22 l.5: "For later use it..." and "and the three...".*

    Done.

23. *p.23 l.2: "a April means" - correct.*

    Done

24. *p.24 l.14: the prior row is the first row and not the third.*

    It is the third row if you take the two header rows into account.

25. *p.24 l.15: "uncertainties" - It is not only 1 uncertainty.*

    No longer present, dropped with introduction of extra section.

26. *p.24 l.17-25: I do not see why you list them here.*

    With the new section title "experimental design" this gets probably obvious. We have also changed to enumeration.

27. *p.25 l2: rows 3-18: say to which table you refrer to.*

    Table reference added.

28. *p.26 l7: regions 5 and 6.*

    Done.

29. *p.26 l.9: "In contrast to" instead of "By contrast to".*

    Done.

30. *p.27 l1: remove brackets.*

    Brackets make sense because for WLS this is only one.

31. *p.27 l.6: comma after technically. This sentence is way to long. Split it!*

    Done.

32. *p.32 l.4: "of a grid cell to a grid-cell average": use uniform writing.*

    Thanks, we'll see with the copy-editor ...

33. *p32, l.12: comma after assessment.*

    We put it after SIFB:

    In the assessment of SIFB~~Archimides~~, Archimedes' principle is applied in the observation operator, where the input quantities including snow depth are taken from the model.

References

Gent, P. and McWilliams, J.: Isopycnal mixing in ocean circulation models, J. Phys. Oceanogr., 20, 150–155, 1990.

Gent, P., Willebrand, J., McDougall, T., and McWilliams, J.: Parameterizing eddy-induced tracer transport in ocean circulation models, J. Phys. Oceanogr., 25, 463–474, 1995.

Griffies, S. M.: The Gent-McWilliams skew flux, J. Phys. Oceanogr., 28, 831–841, 1998.

Guerrier, D. and Horley, F.: Archimedes: Archimedes' Principle and the Law of Flotation, Discovering with the Scientists Series, Blond & Briggs, URL `https://books.google.de/books?id=FOa_AAAACAAJ`, 1970.

Haas, C., Beckers, J., King, J., Silis, A., Stroeve, J., Wilkinson, J., Notenboom, B., Schweiger, A., and Hendricks, S.: Ice and Snow Thickness Variability and Change in the High Arctic Ocean Observed by In Situ Measurements, Geophysical Research Letters, 44, 10,462–10,469, https://doi.org/10.1002/2017GL075434, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL075434`, 2017.

Hardt, M. and Scherbaum, F.: The Design of Optimum Networks for Aftershock Recordings, Geophys. J. Int., 117, 716–726, 1994.

Hibler, W.: A dynamic thermodynamic sea ice model, Journal Geophysical Research, 9, 815–846, 1979.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J.: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth system model, J. Adv. Model. Earth Syst., 5, 422–446, https://doi.org/10.1002/jame.20023, 2013.

Kaminski, T. and Mathieu, P.-P.: Reviews and syntheses: Flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle, Biogeosciences, 14, 2343–2357, https://doi.org/10.5194/bg-14-2343-2017, URL `http://www.biogeosciences.net/14/2343/2017/`, 2017.

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, Atmospheric Chemistry and Physics, 12, 7867–7879, https://doi.org/10.5194/acp-12-7867-2012, URL `https://www.atmos-chem-phys.net/12/7867/2012/`, 2012.

Kaminski, T., Kauker, F., Eicken, H., and Karcher, M.: Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies, The Cryosphere, 9, 1721–1733, https://doi.org/10.5194/tc-9-1721-2015, URL `http://www.the-cryosphere.net/9/1721/2015/`, 2015.

Kurtz, N. and Farrell, S.: Large-scale surveys of snow depth on Arctic sea ice from Operation IceBridge, Geophysical Research Letters, 38, 10,462–10,469, https://doi.org/10.1029/2011GL049216, 2011.

Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H., and Breivik, L.-A.: Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic, Journal of Geophysical Research: Oceans (1978–2012), 115, C10 032, https://doi.org/10.1029/2009JC005958, 2010.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, https://doi.org/10.5194/tc-9-269-2015, URL https://www.the-cryosphere.net/9/269/2015/, 2015.

Niederdrenk, A.: The Arctic hydrologic cycle and its variability in a regional coupled climate model, PhD Thesis, Unversity Hamburg, pp. 1–186, 2013.

Pacanowski, R. and Philander, S.: Parameterization of vertical mixing in numerical-models of tropical oceans, J. Phys. Oceanogr., 11, 1443–1451, 1981.

Rayner, P., Michalak, A. M., and Chevallier, F.: Fundamentals of Data Assimilation, Geoscientific Model Development Discussions, 2016, 1–21, https://doi.org/10.5194/gmd-2016-148, URL http://www.geosci-model-dev-discuss.net/gmd-2016-148/, 2016.

Rayner, P. J., Enting, I. G., and Trudinger, C. M.: Optimizing the CO_2 Observing Network for Constraining Sources and Sinks, Tellus, 48B, 433–444, 1996.

Redi, M. H.: Oceanic isopycnal mixing by coordinate rotation, J. Phys. Oceanogr., 12, 1154–1158, 1982.

Semtner, A.: A Model for the Thermodynamic Growth of Sea Ice in Numerical Investigastions of Climate, Journal pf Physical Oceanography, 6, 379–389, 1976.

Todling, R.: Comparing Two Approaches for Assessing Observation Impact, Monthly Weather Review, 141, 1484–1505, https://doi.org/10.1175/MWR-D-12-00100.1, URL https://doi.org/10.1175/MWR-D-12-00100.1, 2013.

UNESCO: Algorithms for computation of fundamental properties of seawater, UNESCO Technical Papers in Marine Science, 44, 1983.

Warren, S., Rigor, I., Untersteiner, N., Radionov, V., Bryaz-gin, N., Aleksandrov, Y., and Colony, R.: Snow depth on Arctic sea ice, J. Clim., 12, 18141829, 1999.

Arctic Mission Benefit Analysis: Impact of Sea Ice Thickness, Freeboard, and Snow Depth Products on Sea Ice Forecast Performance

Thomas Kaminski[1], Frank Kauker[2,5], Leif Toudal Pedersen[3], Michael Voßbeck[1], Helmuth Haak[4], Laura Niederdrenk[4], Stefan Hendricks[5], Robert Ricker[5], Michael Karcher[2,5], Hajo Eicken[6], and Ola Gråbak[7]

[1]The Inversion Lab, Hamburg, Germany
[2]Ocean Atmosphere Systems, Hamburg, Germany
[3]eolab.dk, Copenhagen, Denmark
[4]Max Planck Institute for Meteorology, Hamburg, Germany
[5]Alfred Wegener Institute, Bremerhaven, Germany
[6]University of Alaska Fairbanks, Fairbanks, USA
[7]European Space Agency ESRIN, Frascati (Rome), Italy

Correspondence to: Thomas Kaminski (Thomas.Kaminski@Inversion-Lab.com)

Abstract. Assimilation of remote sensing products of sea ice thickness (SIT) into sea ice-ocean models has been shown to improve the quality of sea ice forecasts. ~~Open~~ Key open questions are whether ~~the assimilation of rawer~~ assimilation of lower-level data products such as radar freeboard (RFB) can ~~achieve yet a better~~ further improve model performance and what performance ~~gain~~ gains can be achieved ~~by the joint assimilation~~ through joint assimilation of such data products in combination with a snow depth product. The Arctic Mission Benefit Analysis ~~(ArcMBA)~~ system was developed to address this type of question. Using the quantitative network design (QND) approach, the system can evaluate, in a mathematically rigorous fashion, the observational constraints imposed by individual and groups of data products.

We present assessments of the observation impact (added value) of different Earth Observation (EO) products in terms of the uncertainty reduction in a four-week forecast of sea ice volume (SIV) and snow volume (SNV) for three regions along the Northern Sea Route ~~by~~ using a coupled model of the sea ice-ocean system~~. The assessments cover~~, specifically the Max Planck Institute Ocean Model. We assess seven satellite products~~,~~; three real products and four hypothetical products. The real products are monthly SIT, sea ice freeboard (SIFB), and RFB, all derived from CryoSat-2 by the Alfred Wegener Institute. These are complemented by two hypothetical monthly laser freeboard (LFB) products ~~(one with low accuracy and one with high accuracy)~~with low and high accuracy, as well as two hypothetical monthly snow depth products ~~(again one with low accuracy and one with high accuracy)~~with low and high accuracy.

On the basis of the per-pixel uncertainty ranges ~~that are~~ provided with the CryoSat-2 SIT, SIFB, and RFB products, the SIT ~~achieves~~ and RFB achieve a much better performance for SIV than the SIFB product~~, while the performance of RFB is more similar to that of SIT~~. For SNV, the performance of SIT is only low, the performance of SIFB higher and the performance of RFB yet higher. A hypothetical LFB product with low accuracy (20 cm uncertainty) ~~lies in performance~~ falls between SIFB

and RFB in performance for both SIV and SNV. A reduction in the uncertainty of the LFB product to 2 cm yields a significant increase in performance.

Combining either of the SIT/freeboard products with a hypothetical snow depth product achieves a significant performance increase. The uncertainty in the snow product matters: A higher accuracy product achieves an extra performance gain. ~~The provision of~~ Providing spatial and temporal uncertainty correlations with the EO products would be beneficial not only for QND assessments, but also for assimilation of the products.

1 Introduction

Over the last few decades the state of the Arctic climate system has undergone ~~a~~ rapid change. Most pronounced are ~~decreases of the~~ major decreases in summer sea ice extent and ~~of the year-round~~ sea ice volume throughout the year. This transformation is affecting marine ecosystems and coastal communities in an unprecedented way. Economic activities such as resource extraction, maritime transportation, and tourism may benefit from these changes provided that risks, e.g. of sea ice hazards, can be managed. In this context, the performance of short-term to seasonal forecasts of sea ice conditions is of crucial importance (Eicken, 2013).

Forecasts of the sea ~~ice-ocean dynamics~~ ice and the ocean state are routinely performed by coupled sea ice-ocean models that are driven by prescribed atmospheric conditions. ~~Such forecastssuffer from uncertainty~~ In order to derive reliable forecasts, uncertainties in the model ~~'s~~ initial state, of the atmospheric boundary conditions, and ~~the parametrisation~~ in the parameterisations of physical processes ~~. Only observations can help to~~ need to be minimised. Observations can help reduce such uncertainties and, thus, improve the forecast quality. Recently Earth observation (EO) products of sea ice thickness (SIT) have been shown to provide particularly valuable constraints (Lisaeter et al., 2007; Yang et al., 2014; Day et al., 2014; Kauker et al., 2015; Xie et al., 2016). The ~~constraint from rawer EO products that~~ constraints from lower-level EO products (i.e. rawer products that more directly related to the actual measurement) that are used to derive SIT products may be even stronger, because ~~these rawer products~~ such products that conform more closely to the raw EO data are typically more accurate. ~~In~~ For the example of the CryoSat-2 SIT product (Ricker et al., 2014) retrieved ~~at~~ by the Alfred Wegener Institute (AWI) the uncertainty in the radar freeboard (RFB) product underlying their SIT retrieval is smaller by about two orders of magnitude compared to the derived ice thickness product (Figure 13). This difference is a consequence of ~~uncertainty in assumptions in particular on~~ the uncertainty associated in particular with snow and ice density and snow depth, which are used to retrieve SIT from RFB. For direct assimilation of RFB these variables can be ~~taken~~ extracted from the model into which the data are assimilated, but ~~still they are uncertain~~even in this approach significant uncertainty remains. Hence, the trade-off between assimilation of SIT or RFB requires a rigorous quantitative assessment. This is even more important, when the products are assimilated jointly with ~~products of further~~ variables such as snow depth (SND) that ~~bring in~~ introduce complementary information.

Such rigorous assessments can be performed in an efficient manner by the quantitative network design (QND) approach~~. QND allows~~ , allowing for an objective evaluation of ~~observation impact on~~ the added value of observations for a given aspect of a model simulation or forecast. The technique originates from seismology (Hardt and Scherbaum, 1994) and was first applied to the climate system by Rayner et al. (1996), who optimised the spatial distribution of in situ observations of atmospheric carbon dioxide to achieve minimum uncertainty in inferred surface fluxes. After an initial QND study that demonstrated the feasibility of the approach for remote sensing of the column-integrated atmospheric carbon dioxide concentration (Rayner and O'Brien, 2001) QND is now routinely applied in the design of CO2 space missions (e.g., Patra et al., 2003; Houweling et al., 2004; Crisp et al., 2004; Feng et al., 2009; Kadygrov et al., 2009; Kaminski et al., 2010; Hungershoefer et al., 2010; Rayner et al., 2014; Bovensmann et al., 2015). For the western Arctic domain, the QND approach ~~was successfully demonstrated through the evaluation of the combination of hypotheticalairborne altimeter/radar observations~~ has been successfully applied to evaluate the impact of (hypothetical) airborne measurements of SIT/SND in improving sea ice predictions (Kaminski et al., 2015). The study evaluated two idealised flight transects derived from NASA's Operation IceBridge airborne altimeter ice surveys in terms of their potential to improve ten-day to five-month forecasts of sea ice conditions, including for operational purposes.

The present study describes the implementation of the QND methodology into a system for Arctic mission benefit analysis (ArcMBA) and then applies the system to investigate the impact of a series of EO products ~~onto~~ on forecasts of snow and ice volume over three regions along the Northern Sea Route (NSR). It addresses products of SIT, SIFB, RFB, laser freeboard (LFB), and SND. The layout of the remainder of this article is as follows: Section 2 will describe the methodological aspects, including the QND approach, the coupled sea ice-ocean model, and the EO products. Section 3 will present the simulated sensitivities of target quantities and observation equivalents to the model's control vector that is composed of process parameters, initial and boundary conditions. Section 5 will present the QND assessments, followed by ~~their discussion~~ a discussion of these findings in Section 6. Finally, Section 7 provides a summary and conclusions.

2 Methods

2.1 Quantitative ~~Network Design~~network design

The QND methodology is presented by Kaminski and Rayner (2017), partly based on ~~algebra by~~ Tarantola (2005) and Rayner et al. (2016). For the sake of self-containedness we provide a shortened form of the presentation by Kaminski and Rayner (2017). As mentioned, the QND formalism performs a rigorous uncertainty propagation from the observations ~~via the control vector~~ to a target quantity of interest ~~through a dedicated modelling chain. Hence, it is worth recalling the four influence factors which produce~~ relying on the indirect link from the observations to the target variables established by a numerical model. We distinguish between four sources of uncertainty in a model simulation:

1. Uncertainty caused by the formulation of individual process representations and their numerical implementation (structural uncertainty).
2. Uncertainty in constants (process parameters) in the formulation of these processes (parametric uncertainty).
3. Uncertainty in external forcing/boundary values (such as surface winds or precipitation) driving the relevant processes.
4. Uncertainty in the state of the system at the beginning of the simulation (initial state uncertainty).

The first ~~factor~~ category reflects the implementation of the relevant processes into the model (code) while the others can be ~~understood as~~ represented by a set of input quantities controlling the behaviour of a simulation using the given model implementation. The QND procedure formalises the selection of these input quantities through the definition of a control vector, x. The choice of the control vector is a subjective element in the QND procedure. A good choice covers all input ~~factors~~ quantities with high uncertainty and high impact on simulated observations d_{mod} or target quantities y (Kaminski et al., 2012; Rayner et al., 2016).

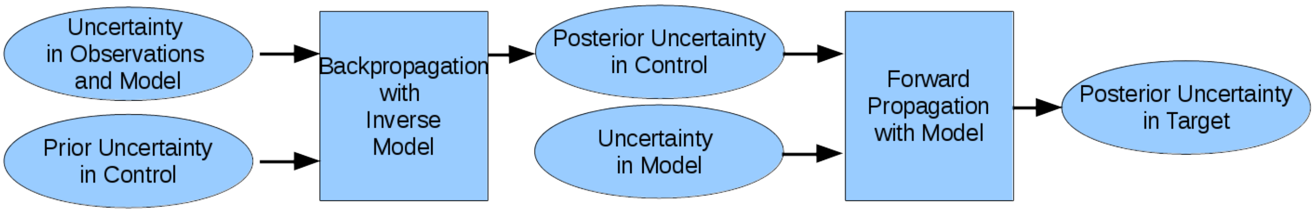


Figure 1. Data flow through two-step procedure of QND formalism. ~~Ovals~~ Oval boxes denote data, rectangular boxes denote processing. Figure taken from Kaminski and Rayner (2017).

The target quantity may be any quantity that can be extracted from a simulation with the underlying model ~~, i.e. any potential model output~~ (in the current study regional integrals of predicted sea ice and snow volumes, see Section 2.2), but also any component of the control vector, for example a process parameter such as the ~~albedo of the snow~~ snow albedo. In the general case, where the target quantity is not part of the control vector, the QND procedure operates in two steps (Figure 1). The first step (inversion step) uses the observational information to reduce the uncertainty in the control vector, i.e. from a prior to a posterior state of information~~, and the~~ . The second step (prognostic step) propagates the posterior uncertainty forward to the simulated target quantity.

~~In this procedure we take uncertainty into account by representing all variables, i.e. the prior and posterior control vectors as well as the observations, their equivalents simulated by the model, and the simulated target quantity by~~ Within the QND formalism, we present all involved quantities by probability density functions (PDFs). We typically assume a Gaussian form for the prior control vector and the observations, if necessary after a suitable transformation. The Gaussian PDFs' covariance matrices express the uncertainty in the respective quantities, i.e. $\mathbf{C}(x_0)$ and $\mathbf{C}(d_{\text{obs}})$ for the prior control vector and the observations. In the context of these PDFs we will use the term uncertainty to refer to its full covariance matrix in the case of a vector quantity, and in the case of a scalar quantity or a given vector component it refers to the square root of the entry on the diagonal of the full covariance matrix corresponding to that particular vector component. In the latter case the uncertainty refers to one standard deviation of the marginal PDF corresponding to that component, and we use the notation $\sigma(d_2)$ to denote, for example, the standard deviation of the second component of d.

For the first QND step we use ~~the model~~ a mapping M ~~as a mapping~~ from control variables onto equivalents of the observations. In our notation the observation operators that map the model state onto the individual data streams (see Kaminski and Mathieu (2017) and Section 2.5) are ~~absorbed~~ incorporated in M. Here we refer to M as model. Let us first consider the case of a linear model, for which we denote by $\mathbf{M}'$ the Jacobian matrix of M, i.e. the derivative of M with respect to x. In this case, the posterior control vector is described by a Gaussian PDF with ~~covariance~~ uncertainty $\mathbf{C}(x)$, ~~i.e. the uncertainty~~ which is given by

$$\mathbf{C}(x)^{-1} = \mathbf{M}'^{T}\mathbf{C}(d)^{-1}\mathbf{M}' + \mathbf{C}(x_0)^{-1} \tag{1}$$

where the data uncertainty $\mathbf{C}(d)$ ~~combines~~ is the combination of two contributions:

$$\mathbf{C}(d) = \mathbf{C}(d_{\text{obs}}) + \mathbf{C}(d_{\text{mod}}) \tag{2}$$

The term $\mathbf{C}(d_{\text{obs}})$ ~~with the uncertainty~~ expresses the uncertainty in the observations and $\mathbf{C}(d_{\text{mod}})$ the uncertainty in the simulated equivalents of the observations $M(x)$~~:~~

$$\cancel{\mathbf{C}(d)^2 = \mathbf{C}(d_{\text{obs}})^2 + \mathbf{C}(d_{\text{mod}})^2}$$

. The first term in Equation (1) expresses the ~~observational constraint~~ impact of the observations and the second term the ~~prior information content~~ impact of the prior information. In the non-linear case we use Equation (1) as an approximation of $\mathbf{C}(x)$.

~~In the secondstep, the Jacobian matrix $\mathbf{N}'$ of the model (now used as a~~ The mapping N involved in the second, the uncertainty propagation step, is the mapping from the control vector onto ~~target quantities and denoted by~~ a target quantity, y. The Jacobian matrix $\mathbf{N}'$ of the mapping N ~~)~~ is employed to ~~propagate the~~ approximate the propagation of the posterior uncertainty in the control vector $\mathbf{C}(x)$ forward to the uncertainty in a target quantity, $\sigma(y)$ ~~:~~ via

$$\sigma(y)^2 = \mathbf{N}'\mathbf{C}(x)\mathbf{N}'^{T} + \sigma(y_{\text{mod}})^2 \,. \tag{3}$$

If the model ~~was~~ were perfect, $\sigma(y_{\text{mod}})$ would be zero. In contrast, if the control variables were perfectly known, the first term on the ~~right hand~~ right-hand side would be zero. The terms $\mathbf{C}(d_{\text{mod}})$ in Equation (2) and $\sigma(y_{\text{mod}})$ in Equation (3) capture the structural uncertainty as well as the uncertainty in those process parameters, boundary and initial values that are not included in the control vector. These two terms typically rely on subjective estimates. When comparing the effect of different data sets in the same setup, $\sigma(y_{\text{mod}})$ acts as an offset (for the respective variance) in Equation (3)~~)~~. To sharpen the contrast between the products we remove it from the assessment and report two plausible estimates separately.

To conduct a valuable QND assessment, the requirement on the model is not that it simulates the target quantities and observations under investigation realistically, but ~~the requirement is~~ rather that it provides a realistic *sensitivity* of the target quantities and observations under investigation with respect to a change in the control vector. ~~If these sensitivities ,~~ (As a hypothetical example we can think of a perfect regional tracer model that is run with an offset in the initial or boundary conditions for a passive tracer. The simulated tracer concentration will carry this offset, but all sensitivities will be perfect.) If the sensitivities of the target quantities and observations (i.e. the Jacobians~~,~~) are realistic, but the simulation of target quantities and observations incorrect, we can always ~~make~~ obtain a valuable QND assessment with appropriate model uncertainty. The result of the assessment may then be that a particular data stream is not useful in constraining a particular target quantity given current modelling capabilities. ~~In this situation we could operate~~ Under such circumstances, the QND system could be operated with reduced model uncertainty to explore ~~which accuracy~~ the level of accuracy required of the model ~~is required~~ for a data stream to ~~be~~ serve as a useful constraint on a given target quantity. In particular when it comes to ~~new~~ newly available, unvalidated data streams and target quantities the accuracy of both, the simulation and the sensitivities, ~~are~~ is hard to assess. In the case of a model that ~~misses~~ does not capture relevant processes we may expect errors in both the simulation and the sensitivities, and consequently also in the QND assessment.

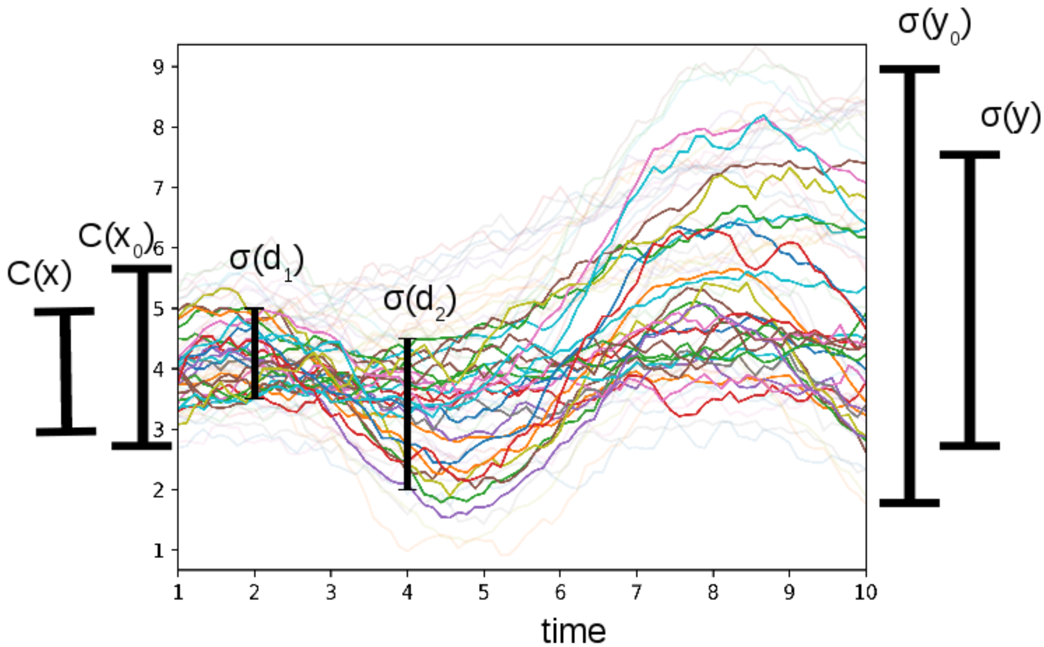


Figure 2. Schematic ~~Presentation~~ presentation of the QND procedure: Each coloured line illustrates a model trajectory that simulates for a given value of the control vector (x) counterparts of the observations (d_1 and d_2) and a target quantity (y). Through the model, the observations act as constraints on the control vector, which reduces its uncertainty from $C(x_0)$ to $C(x)$. This uncertainty reduction on the control vector translates into an uncertainty reduction in the target quantity from $\sigma(y_0)$ to $\sigma(y)$

.

Our performance metric is the (relative) reduction in posterior target uncertainty $\sigma(y)^2$ with respect to a reference. To compare against the case without any observations we ~~use~~compute, as the reference, the prior target uncertainty, $\sigma(y_0)$, via

$$\sigma(y_0)^2 = \mathbf{N}'\mathbf{C}(x_0)\mathbf{N}'^T + \sigma(y_{\text{mod}})^2 \,. \tag{4}$$

The uncertainty reduction with respect to the prior,

$$\frac{\sigma(y_0) - \sigma(y)}{\sigma(y_0)} = 1 - \frac{\sigma(y)}{\sigma(y_0)}, \tag{5}$$

quantifies the impact of the entire network. A schematic illustration of the approach with the prior and posterior uncertainty ranges is shown in Figure 2. The observations d_1 and d_2 render a range of trajectories unlikely, which in the first step leads to a reduction of uncertainty in the control vector (from $C(x_0)$ to $C(x)$) and in the second step to a reduction in the target uncertainty (from $\sigma(y_0)$ to $\sigma(y)$).

We note that (through Equation (1) and Equation (3)) the posterior target uncertainty solely depends on the prior and data uncertainties, the contribution of the model error to the uncertainty in the simulated target variable, $\sigma(y_{\text{mod}})$, as well as the observation and target Jacobians (quantifying the linearised model responses of the simulated observation equivalent and of the target quantities). ~~The QND formalism does not require real observations and can thus~~ Hence, the QND formalism can be employed to evaluate hypothetical candidate networks. Candidate networks are ~~defined by a set of observations~~ characterised by observational data type, location, sampling frequency and time, and data uncertainty but not the observational value. Here, we define a network as the complete set of the characterisation of observations, d, used to constrain the model. The term network is not meant to imply that the observations are of the same type or that their sampling is coordinated. For example, a network can combine different types of in situ and satellite observations.

In practice, for pre-defined target quantities and observations, model responses can be pre-computed and stored. A network composed of these pre-defined observations can then be evaluated in terms of the pre-defined target quantities without any further model runs. Only matrix algebra is required to combine the pre-computed sensitivities with the data ~~uncertainties~~uncertainty. This aspect is exploited in our ArcMBA system.

2.2 Target ~~Quantities~~quantities

For this study we selected target quantities that are particularly relevant for maritime transport, namely predicted sea ice volume (SIV) and snow volume (SNV) over three regions along the ~~Northern Sea Route (NSR)~~NSR. These three regions are displayed in Figure 3 and respectively denoted as “West Laptev Sea” (WLS), “Outer New Siberian Islands” (ONSI), and “East Siberian Sea” (ESS). We perform these predictions for May 28, 2015, a point in time at which there is still sufficient snow cover for our prediction to be relevant. These predictions are started on April 1 and are constrained by observational information until April 30, i.e. ~~we perform~~ the assimilation window in April is followed by a four-week prediction period (Figure 4).

2.3 ~~Model~~Sea ice-ocean model

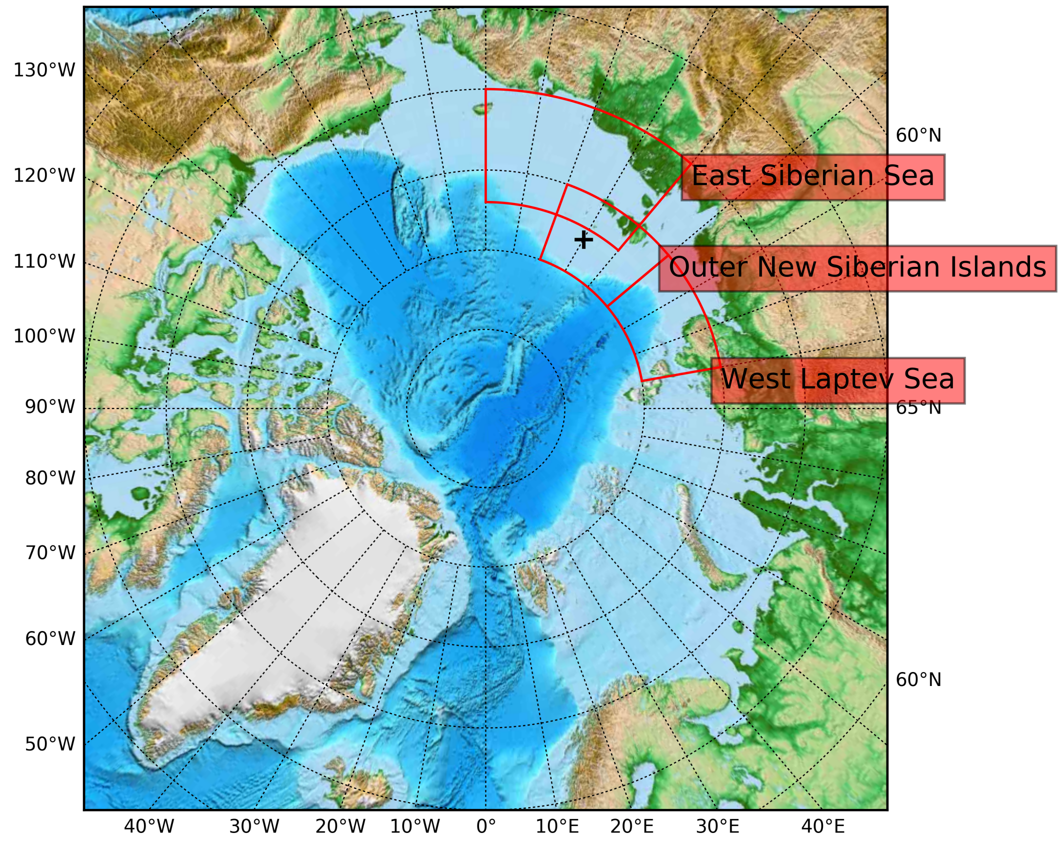


Figure 3. Target regions along the ~~Northern Sea Route~~NSR. Black cross indicates a location for further use in Figure 14.

~~The requirement on the dynamical~~ To simulate observation equivalents (M in Equation (1)) and target quantities (N in Equation (3)) we employ a coupled model of the ~~coupled~~ sea ice-ocean system~~is that it simulates in a realistic manner~~. The model is required to provide realistic simulations of the sensitivity of ~~the~~ observation equivalents and ~~the~~ target quantities to changes in the control variables. In the present study we use the Max-Planck-Institute Ocean Model ~~(MPIOM) (Jungclaus et al., 2012, 2013; Haak et al., 2003)~~(MPIOM, Jungclaus et al., 2012, 2013; Haak et al., 2003), i.e. the sea ice-ocean component of the Max-Planck-Institute Earth System Model ~~(MPI-ESM) (Giorgetta et al., 2013)~~(MPI-ESM, Giorgetta et al., 2013). MPI-ESM regularly provides climate projections for the Intergovernmental Panel on Climate Change (IPCC) in particular to the IPCC's 5th assessment report (Stocker et al., 2013) and the upcoming 6th assessment report (AR6) and within the seasonal

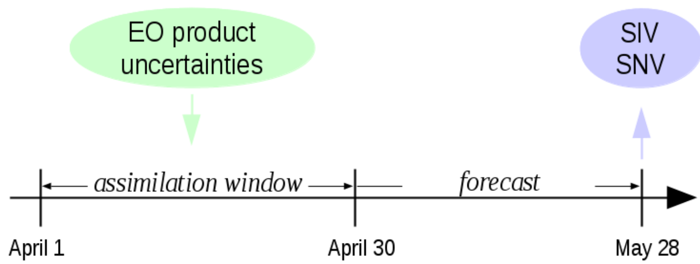


Figure 4. Time line of ~~experimental~~ assimilation and forecast setup.

to decadal prediction system (Müller et al., 2012). In the following we provide a brief description of the ~~model~~current model development status, largely following Jungclaus et al. (2006) and Niederdrenk (2013).

~~MPIOM is based on the primitive equations, a set of nonlinear differential equations that approximate the oceanic flow and are used in most oceanic models. They consist of three main sets of balance equations: A continuity equation representing the conservation of mass, the Navier-Stokes equations ensuring conservation of momentum, and a thermal energy equation relating the overall temperature of the system to heat sources and sinks. Diagnostic treatment of pressure and density is used to close the momentum balance. Density is taken to be a function of model pressure, temperature and salinity (UNESCO, 1983). Recent development of the model~~ Recent development of the ocean part of the model includes the treatment of horizontal discretisation which has undergone a transition from a staggered E-grid to an orthogonal curvilinear C-grid. The treatment of subgridscale mixing has been improved ~~by~~ through the inclusion of a new formulation of bottom boundary layer slope convection, an isopycnal diffusion scheme, and a Gent and McWilliams style eddy-induced mixing parameterisation ~~. Along-isopycnic~~ (Gent and McWilliams, 1990). Along-isopycnal diffusion is formulated following Redi (1982) and Griffies (1998). Isopycnal tracer mixing by unresolved eddies is parameterised following Gent et al. (1995). For the vertical eddy viscosity and diffusion the Richardson number–dependent scheme of Pacanowski and Philander (1981) is used. An additional wind mixing proportional to the cube of the 10-m wind speed (decaying exponentially with depth) compensates for too low turbulent mixing close to the surface. Static instabilities are removed through enhanced vertical diffusion.

A viscous–plastic rheology (Hibler, 1979) is used for the sea ice dynamics. ~~The thermodynamics is~~ Sea ice thermodynamics are formulated using a Semtner (1976) zero-layer model relating changes in sea ice thickness to a balance of radiant, turbulent, and oceanic heat fluxes. In the zero-layer model the conductive heat flux within the sea ice/snow layer is assumed to be directly proportional to the temperature gradient across the sea ice/snow layer and inversely proportional to the thickness of that layer, i.e. the sea ice does not store heat. The effect of snow accumulation on sea ice is included, along with snow–ice transformation when the snow/ice interface sinks below the sea level because of snow loading (flooding). The effect of ice formation and melting is accounted for within the model assuming a sea ice salinity of 5 psu.

MPIOM allows for an arbitrary placement of the model's poles on an orthogonal curvilinear grid. In the setup used here (taken from Niederdrenk (2013); Mikolajewicz et al. (2015); Niederdrenk et al. (2016)) the poles are located over Russia and North America (~~as can be seen in~~ Figure 5). Placement over land avoids numerical singularities that for poles over the ocean would be caused by the convergence of the meridians, and the non-diametric placement allows to reach high resolution (average of about 15 km) ~~of~~ in the Arctic. ~~In the following we will call the model in that configuration~~ This setup achieves a spatial resolution as high as that of the EO products we analyse (in fact over the target regions the model resolution is higher) without major computational constraints, which allows an evaluation of the full spatial information content provided by the respective EO products. Here, we will refer to this particular model configuration as Arctic MPIOM.

As forcing data at the ocean's surface, the model needs heat, freshwater, and momentum. These data are taken from ECMWF's ERA-Interim reanalysis (Dee et al., 2011). ERA-Interim is a global atmospheric reanalysis (of the period from 1979 to present) that is produced by a 2006 release of the Integrated Forecasting System (IFS – version Cy31r2) and applies a 4-dimensional variational analysis with a 12-hour analysis window. The spatial resolution of the data set is approximately 80 km (T255 spectral) on 60 vertical levels from the surface up to 0.1 hPa. ERA-interim surface variables ~~to~~ that force Arctic MPIOM are 2-meter temperature, 2-meter dew point temperature (surrogate of 2-meter specific humidity – not ~~delivered~~ provided by ECMWF), 10-meter zonal and meridional wind velocity (to calculate the wind speed), total cloud cover and the following fluxes (~~delivered~~ provided in accumulated form over the 12-hourly forecast window): surface downward solar radiation, surface downward thermal radiation, total precipitation, zonal and meridional wind stress. Land runoff into the ocean is taken from the German Ocean Model Intercomparison Project (OMIP, Röske, 2001).

~~In this study, all model experiments will be started~~ For the computation of the jacobians M' and N' (introduced in Section 2.1) that is described in Section 3 we run Arctic MPIOM from a restart file for April 1, 2015. This restart file is in turn generated from a hindcast run of Arctic MPIOM ~~which~~ that is initialised on ~~1.1.1979 (start time of ERA-Interim).~~ January 1, 1979. This initialisation is based on a set of observations that consists of a topography data set ~~(ETOPO5 5-minute gridded elevation data (NOAA, 1988))~~(ETOPO5 5-minute gridded elevation data, NOAA, 1988), and a hydrographic climatological data set (Polar science center Hydrographic Climatology, PHC3; Steele et al., 2001) containing potential temperature and salinity. The ocean is assumed to be at rest. Sea ice is assumed to be present if the sea surface temperature falls below the freezing temperature of sea water. 100% ice cover and a sea ice thickness of 2m is assumed where sea ice is present and sea ice is assumed to be at rest. From this initial state the model is integrated with the ERA-Interim surface forcing until ~~31.3.2015~~March 31, 2015 (the beginning of our assimilation window). While a ~~34~~ 36 year integration is certainly too short to spin up the deep ocean, it is sufficient for the purpose of this study, because the spinup time of sea ice and the upper ocean (depth above about 500m) is generally assumed to be only a few decades.

For a successful QND assessment it is essential that MPIOM provides realistic sensitivities of the observation equivalent and the target quantities to the changes in the control vector (Equation (1) and Equation (3)). However, observations are not available to validate these sensitivities. The only validation of MPIOM possible is against observations of the state of the sea ice and ocean. In the following we present comparisons with selected observation based products first for the hindcasting period, and then for the assimilation window and the forecasting period.

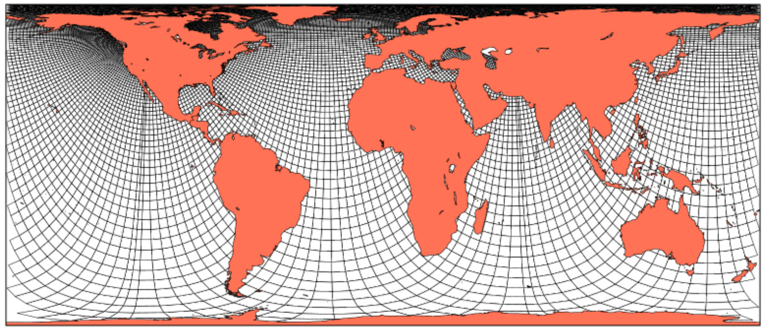


Figure 5. Model grid, ~~global (upper panel) and over the Arctic (lower panel); Upper panel shows~~ mesh indicates groups of 4 by 4 grid cells~~and lower panel groups of 2 by 2 grid cells~~.

The hindcast with Arctic MPIOM has been validated against remotely sensed ice concentration from the reprocessed ~~OSI SAF~~ Ocean and Sea Ice Satellite Application Facility (OSI SAF) sea ice concentration product (Eastwood et al., 2015) and against a combination of ~~in-situ~~ in situ and remotely sensed ice thickness observations. ~~In-situ~~ In situ observations of sea ice thickness still have a high uncertainty and each data source has its own strengths and weaknesses. As of today the most reliable ~~information about~~ pan-Arctic sea ice thickness ~~stems~~ data set is derived from a combination of various sources of ~~in-situ observation~~ in situ observations and remotely sensed satellite sea ice thickness products ~~by Lindsay and Schweiger (2015)~~(Lindsay and Sch

The reprocessed OSI SAF sea ice concentration product is available daily on a 10 km spatial grid and includes spatially and temporally varying uncertainty estimates. For an assessment of the performance of the Arctic MPIOM, the sea ice concentration has been compared to the long-term means of the March, June, and September monthly means for the period 1990 to 2008 (Figure 6). In March (panel d) and June (panel e) only ~~small~~ relatively small scale misfits to the OSI SAF ice concentration are found but they can reach up to 50% (here and in the following we use the term “misfit” for the model-observation difference). The sea ice margin in the Nordic Seas and Barents Sea is captured well. The anomalies apparent in March correspond to the results of a study performed with the MPIOM version of the Max-Planck-Institute’s Earth System model MPI-ESM-LR (Notz et al., 2013), for which the MPIOM was forced with the same atmospheric forcing data set as used in our study (ERAinterim) (see panel f of their Figure 3).

In September large misfits to the OSI SAF sea ice concentration are obtained (Figure 6 panel f). Especially over the Eurasian basin the model’s sea ice margin is located too far north but also over the central Arctic the model ~~is underestimating~~ underestimates the sea ice concentration. In our target ~~reions~~ regions the misfit remains relatively small. The aforementioned analysis by Notz et al. (2013) shows similar misfits (see panel f of their Figure 4) to a different sea ice concentration data set, namely NSIDC-CDR (National Snow and Ice Data Center Climate Data Record).

An evaluation of the hindcast simulation with Arctic MPIOM with respect to the modelled SIT is much more difficult, because the observation-based products exhibit large uncertainties reflecting the corrections imposed by the respective measurement principle. For example, ~~Electro-magnetic~~ electro-magnetic Air-EM measurements detect the air-snow interface, and

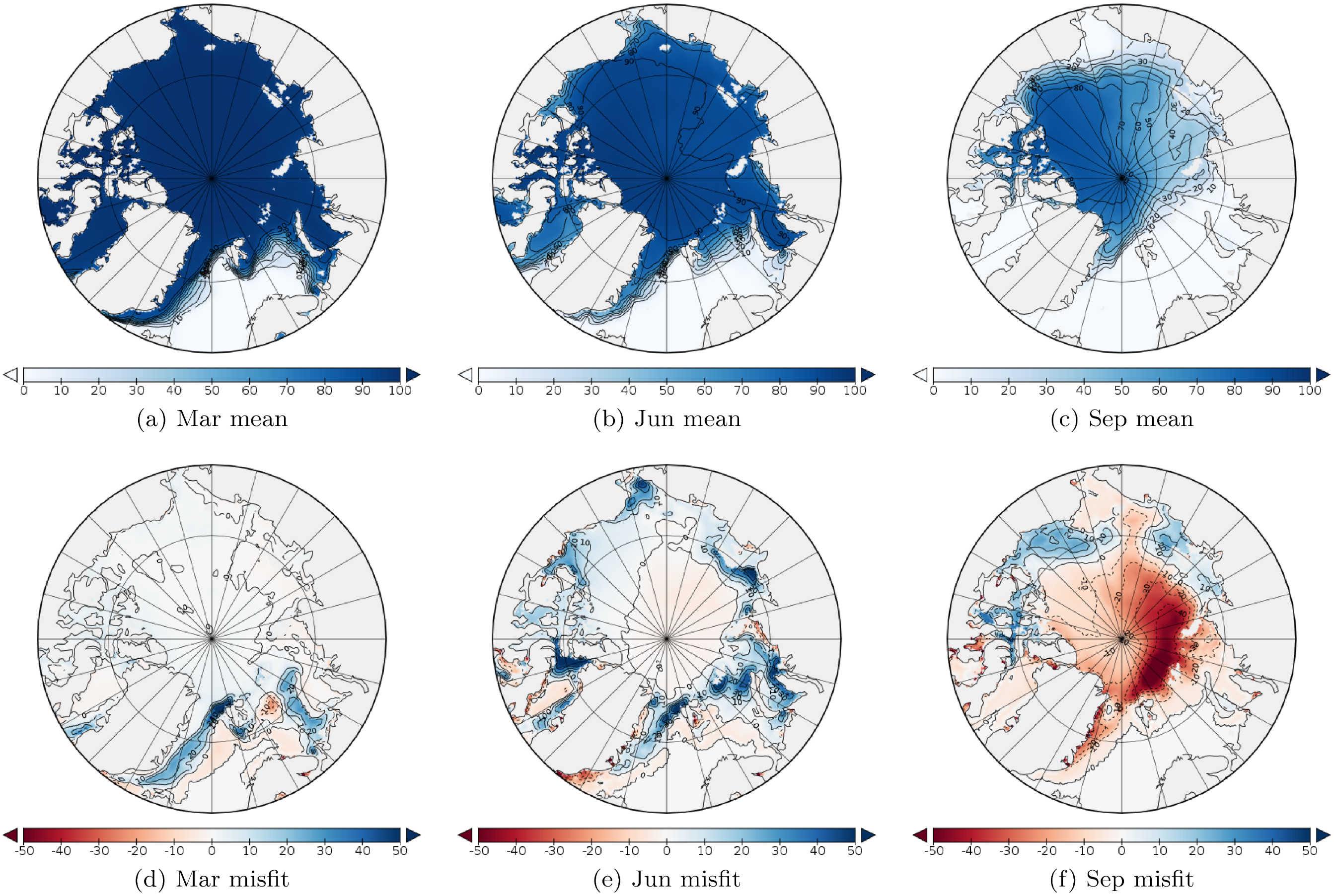


Figure 6. The long-term mean sea ice concentration [%] of the Arctic MPIOM for 1990 to 2008 for March, June and September (panel a to c) and the misfit to the OSI SAF sea ice concentration (panel d to f). In panels d to f, red colours indicate underestimation and blue colours overestimation of sea ice concentration in the model.

not the interface between snow and sea ice, introducing significant errors in the SIT estimates that are corrected by assumptions or measurements of snow depth. Moored and submarine ULS measurements have to be corrected for the first return echo. Differences in the observed and measured spatial scales further complicate the comparison. The aforementioned study of Lindsay and Schweiger (2015) synthesises all available ~~in-situ~~ in situ and remotely sensed satellite SIT products in an ice thickness regression procedure (ITRP) for the time period 2000 to 2012. Low order spatial and temporal polynomials are fitted to the available sea ice thickness measurements. The resulting sea ice thickness regression product describes the evolution in the central Arctic and is linear in time plus a quadratic time-dependent component, i.e. it does not contain year-to-year variability. Uncertainty ranges are deduced from the uncertainty of the individual regression coefficients. The year-to-year variability is reflected in this uncertainty. Lindsay and Schweiger (2015) ~~could~~ show for example that the ICESat ice thickness product from the Jet Propulsion Laboratory (ICESat-JPL, Kwok and Cunningham (2008)), which is widely used for model validation, had a large positive bias.

Here we compare the modelled long-term mean (2000 to 2012) sea ice thickness of the Arctic MPIOM ~~model experiment~~ hindcast to the ITRP sea ice thickness for the two-months periods February/March and October/November. We selected these two-month periods, because the availability of the ICESat satellite product ensures a high data coverage in the ITRP. The long-term mean sea ice thickness of the ~~the~~ Arctic MPIOM hindcast simulation for February/March and October/November is depicted in Figure 7 (panel a and panel b) together with the misfit to the ITRP ice thickness (panel c and panel d). A prominent feature is a strong underestimation of the Arctic MPIOM sea ice thickness north and west of Fram Strait and in the strait itself. In the regions of interest ~~of~~ for our QND study ~~, in~~ (the areas around the ~~Northern Sea Route,~~ NSR) the misfit is moderate in February/March (overestimation of about 25%) with the exception around the ~~East~~ New Siberian Islands where the misfit can reach more ~~then~~ than 1 meter (overestimation of about 50%). In October/November the misfit is very moderate in these areas except for Bering Strait where Arctic MPIOM underestimates the sea ice thickness by more ~~then~~ than 50cm.

~~As we base our QND experiments on simulations from April 1 to May 28, we~~ Next we address Arctic MPIOM performance over our assimilation and forecasting period (see Figure 4). We show the April mean and the May 28 mean of the modelled SIT and the misfit of the April mean thickness to that retrieved from CryoSat-2 (Figure 8). For a comparison of CryoSat-2 thickness to in situ observations we refer to Haas et al. (2017). The misfit to the CryoSat-2 ice thickness in April 2015 is similar to the misfit to the ITRP shown in Figure 7: a strong underestimation north of the Canadian Archipelago and ~~North and West~~ north and west of Fram Strait and a moderate overestimation in the area of the target quantities of about or less ~~then~~ than $50cm$ (about 25% relative error). Figure 9 depicts the April mean and the May 28 mean of the modelled snow depth and the misfit to the modified Warren climatology (Warren et al., 1999) that is used in the CryoSat-2 retrieval (see Section 2.5). The main challenge for sea ice thickness retrieval with satellite altimeters is the parameterisation of snow depth on sea ice, which is still not measured routinely. The current CryoSat-2 retrieval uses a modified snow climatology that addresses shortcomings of the Warren et al. (1999) climatology that was based largely on data from drifting stations mainly on multi-year sea ice collected over the past decades, and hence is not reflective of a much younger, more seasonal Arctic ice cover. Given the increased fraction of first-year ice in the Arctic Ocean, the approach proposed by Kurtz and Farrell (2011) is used and the climatological snow depth values used in the retrieval are multiplied over first-year ice by a factor of 0.5. Note that on May 28 ~~parts of~~ over the target regions ~~are almost snow free already~~ a large fraction of snow cover has already melted. The misfit to the modified Warren climatology in the target area East Siberian Sea is on the order of about $10cm$ (50% relative error) but much less for the other target areas.

Overall, the misfits of the Arctic MPIOM are acceptable in particular for our target regions along the ~~Northern Sea Route~~ NSR (Figure 3) and are comparable to misfits found in sea ice-ocean model intercomparison projects (e.g, Chevallier et al. (2017)).

2.4 Control ~~Vector~~vector

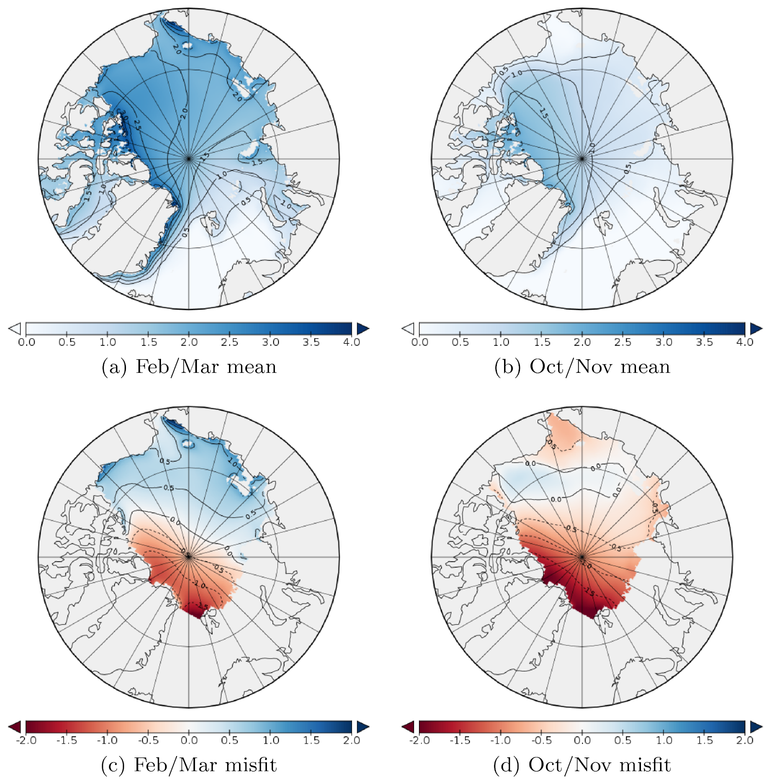


Figure 7.

The long-term mean (2000 to 2012) of the simulated sea ice thickness [m] for the two-month periods February and March and October and November (panel a and b) and the misfit (model – ~~observation~~observations) to the ITRP ~~.~~ (panel c and d). In panels c and d, red colours indicate underestimation and blue colours overestimation of sea ice thickness in the model.

~~The definition~~ Criteria for the choice of the control vector ~~and the specification on prioruncertainty follow Kaminski et al. (2015): The components and their prior uncertainty are~~ are presented in Section 2.1. The specification of prior, both mean (x_0) and uncertainty ($C(x_0)$), follows Kaminski et al. (2015), and is listed in Table 1. The largest possible control vector in our modelling system is the superset of initial and surface boundary conditions as well as all parameters in the process formulations, including the observation operators. As described in Section 3, the Jacobian computation requires an extra run for each additional component of the control vector. To keep our ArcMBA system numerically efficient, two and three-dimensional fields are partitioned into regions. More precisely, we divide the Arctic domain into nine regions (shown in Figure 10). In each of these regions we add a scalar perturbation to each of the forcing fields (indicated in ~~Table 1~~ Table 1 by “f” in the type column~~–~~); the perturbation is applied for the entire simulation time~~)~~. Likewise we add a scalar perturbation to six initial fields ~~(indicated in~~ indicated in Table 1 by “i” in the type column~~)~~. For the ocean temperature and salinity the size of the perturbation is reduced with increasing depth (and zero below ~~500m~~500 m). Finally we have selected 29 process parameters from the sea ice–ocean

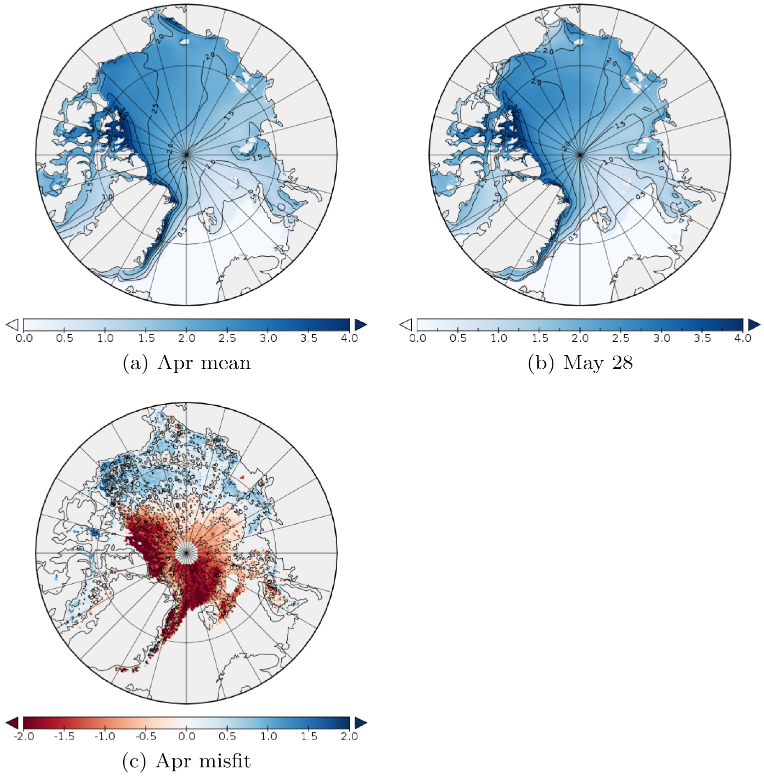


Figure 8. The a) modelled mean April 2015 sea ice thickness [m], b) the modelled sea ice thickness on May 28 2015, and c) the mean April 2015 misfit of the modelled sea ice thickness relative to the CryoSat-2 sea ice thickness. In panel c, red colours indicate underestimation and blue colours overestimation of sea ice thickness in the model.

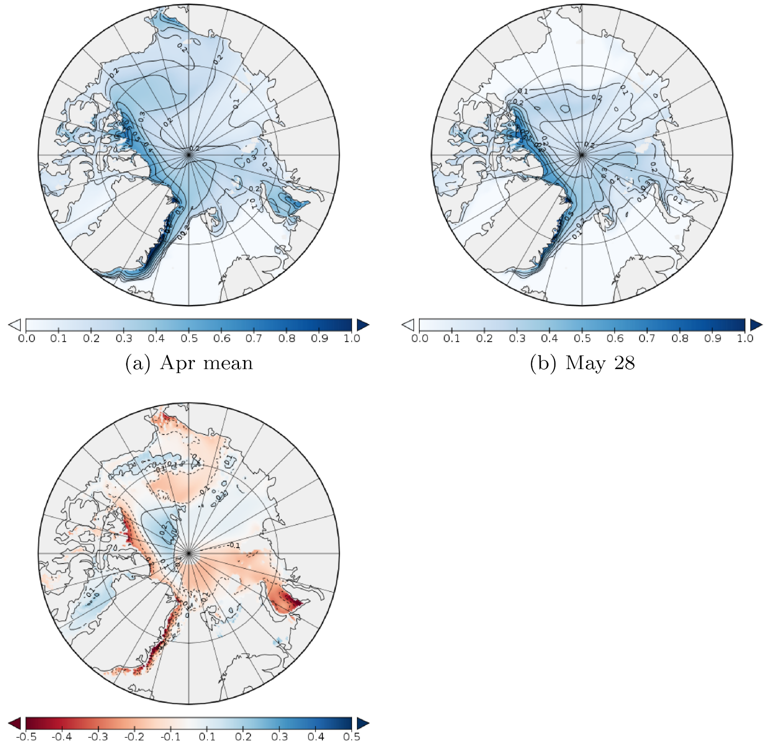
model plus two parameters from the observation operators for freeboard products (see Section 2.5 for details). This procedure ~~resulted~~ results in a total of 157 control variables. We assume the prior uncertainty to have diagonal form, i.e. there are no correlations among the prior ~~uncertainty~~ uncertainties relating to different components of the control vector. The diagonal entries are the square of the prior standard deviation. For process parameters this standard deviation is estimated from the range of values typically used within the modelling community. The standard deviation for the components of the initial state is based on a model simulation over the past 37 years and computed for the 37 member ensemble corresponding to all states on the same day of the year. Likewise the standard deviation of the surface boundary conditions is computed for the 37 member ensemble corresponding to the April-October means of the respective year.

2.5 Data ~~Sets~~ sets and ~~Observation Operators~~observation operators

The study evaluates three data sets retrieved by the AWI (Ricker et al., 2014) from observations provided by the CryoSat-2 mission, two data sets characterising hypothetical LFB products, and two data sets characterising hypothetical SND products.

Table 1. Control variables. Column 1 lists the quantities in the control vector; column 2 gives the abbreviation for each quantity; column 3 indicates whether the quantity is an atmospheric boundary field (forcing, i.e. f), an initial field (i), or a process parameter (p); column 4 gives the name of each quantity; column 5 indicates (the standard deviation of) the prior uncertainty and the corresponding units (unless unitless) and provides the magnitude of the parameter value in parenthesis, where applicable; and column 6 identifies the position of the quantity in the control vector – for initial and boundary values (which are differentiated by region) this position refers to the first region, while the following components of the control vector then cover regions 2 to 9.

Index #	Name	Type	Meaning	Prior uncertainty (value)	Start
1	hiccp	p	(alias pstar) ice strength (devided by density)	$15(20)\,[N\,m^{-2}\,kg^{-1}]$	1
2	hibcc	p	(alias cstar) ice strength depend. on ice conc.	$5.0(20.0)$	2
3	hicce	p	(alias eccen) squared yield curve axis ratio	$0.5(2.0)$	3
4	rleadclose1	p	extra lead closing (Notz et al., 2013)	$0.2(0.25)$	4
5	rleadclose2	p	extra lead closing (Notz et al., 2013)	$1.0(3.0)$	5
6	rleadclose3	p	extra lead closing (Notz et al., 2013)	$1.0(2.0)$	6
7	h_0	p	lead closing	$1.0(0.5)\,[m]$	7
8	hmin	p	mimimal ice thickness	$0.04(0.05)\,[m]$	8
9	armin	p	minimal ice compactness	$0.15(0.15)$	9
10	hsntoice	p	limit on flooding	$0.45(0.45)$	10
11	sice	p	salinity in sea ice	$2.0(5.0)[psu]$	11
12	albi	p	freezing ice albedo	$0.1(0.75)$	12
13	albm	p	melting ice albedo	$0.1(0.70)$	13
14	albsn	p	freezing snow albedo	$0.1(0.85)$	14
15	albsnm	p	melting snow albedo	$0.1(0.70)$	15
16	rhoice	p	density of sea ice	$20(910)[kg/m^3]$	16
17	rhosn	p	density of snow	$20(330)[kg/m^3]$	17
18	cw	p	ocean drag coeff.	$2.0\times10^{-3}(4.5\times10^{-3})$	18
19	av0	p	coeff vertical viscosity	~~$1.\times10^{-4}(2.\times10^{-4})$~~ $1.\times10^{-4}(2.\times10^{-4})[m^2/s]$	19
20	dv0	p	coeff vertical ~~diffusitiv~~ diffusitivity	~~$1.\times10^{-4}(2.\times10^{-4})$~~ $1.\times10^{-4}(2.\times10^{-4})[m^2/s]$	20
21	aback	p	background coeff vertical viscosity	~~$3.\times10^{-5}(5.\times10^{-5})$~~ $3.\times10^{-5}(5.\times10^{-5})[m^2/s]$	21
22	dback	p	background coeff vertical ~~diffusivity~~ diffusitivity	~~$1.\times10^{-5}(1.05\times10^{-5})$~~ $1.\times10^{-5}(1.05\times10^{-5})[m^2/s]$	22
23	cwt	p	vertical wind mixing parameter tracers	~~$2.0\times10^{-4}(3.5\times10^{-4})$~~ $2.0\times10^{-4}(3.5\times10^{-4})[m^2/s]$	23
24	cwa	p	vertical wind mixing parameter momentum	~~$0.4\times10^{-3}(0.75\times10^{-3})$~~ $0.4\times10^{-3}(0.75\times10^{-3})[m^2/s]$	24
25	cstabeps	p	vertical wind mixing stability parameter	$0.03(0.06)$	25
26	cdvocon	p	coefficient for enhanced vertical diffusivity	$0.1(0.15)$	26
27	bofric	p	linear bottom friction	~~$2.\times10^{-4}(3.\times10^{-4})$~~ $2.\times10^{-4}(3.\times10^{-4})[m^2/s]$	27
28	rayfric	p	quadratic bottom friction	~~$0.5\times10^{-3}(1.\times10^{-3})$~~ $0.5\times10^{-3}(1.\times10^{-3})[m^2/s]$	28
29	$jerlov_a$	p	jerlov atten - ocean-water types	$0.04(0.08)$	29
30	$jerlov_b$	p	jerlov bluefrac - ocean-water types	$0.20(0.36)$	30
31	albw	p	open water albedo	$0.05(0.1)$	31
32	sit	i	initial ice thickness	$0.5\,[m]$	32
33	siconc	i	initial ice concentration	0.1	41
34	sicsno	i	initial snow thickness	$0.2\,[m]$	50
35	thetao	i	initial ocean temperature	$0.5\,[K]$ (vertically decreasing)	59
36	so	i	initial salinity	$0.5\,[psu]$ (vertically decreasing)	68
37	zos	i	sea level elevation	$0.08\,[m]$	77
38	cloud	f	cloud cover	0.07	86
39	prec	f	total precipitation	$0.4\times10^{-8}\,[m\,s^{-1}]$	95
40	swrad	f	solar downward radiation	$6.\,[W\,m^{-2}]$	104
41	tdew	f	dew ~~pointe~~ point temperature	$1.1\,[K]$	113
42	tem	f	2m air temperature	$1.2\,[K]$	122
43	wind10	f	10m scalar wind speed	$0.6\,[m\,s^{-1}]$	131
44	wix	f	zonal wind stress x component	$0.02\,[N\,m^2]$	140
45	wiy	f	meridional wind stress y component	$0.02\,[N\,m^2]$	149



(a) Apr mean (b) May 28

(c) Apr misfit

Figure 9. The a) modelled mean April 2015 snow depth [m], b) the modelled snow depth on May 28 2015, and c) the mean April 2015 misfit of the modelled snow depth relative to the modified Warren climatology used in the CryoSat-2 sea ice thickness retrieval. In panel c, red colours indicate underestimation and blue colours overestimation of snow depth in the model.

~~In the following~~ Below, we describe these data sets and the simulation of their model equivalents, i.e. the respective observation operators that provide the ~~link~~ links from the model's state variables ~~(Kaminski and Mathieu, 2017) .~~ to the respective data sets (Kaminski and Mathieu, 2017) . Recall that the (combination of) data set(s) enters the QND algorithm through its uncertainty $C(d)$ and that the observation operator is incorporated in the model M (see Section 2.1).

The three products derived by AWI from CryoSat-2 are SIT (h_i), SIFB (f_i), and RFB (f_r). Their definition is illustrated in Figure 11 together with that of LFB (f_l).

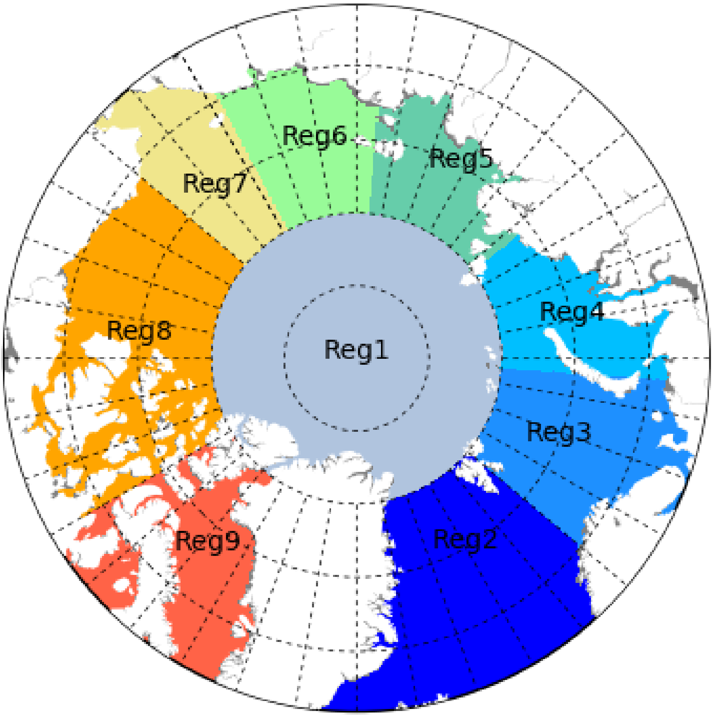


Figure 10. Sub-regions for spatial differentiation of initial and boundary values in the control vector. 1 (light plum): central Arctic; 2 (dark blue): North Atlantic; 3 (blue) Barents Sea; 4 (light blue) Kara Sea; 5 (green) Laptev Sea, 6 (light green) East Siberian Sea; 7 (yellow): Bering Strait/Chukchi Sea; 8 (orange): Beaufort Sea; 9 (red): Baffin Bay.

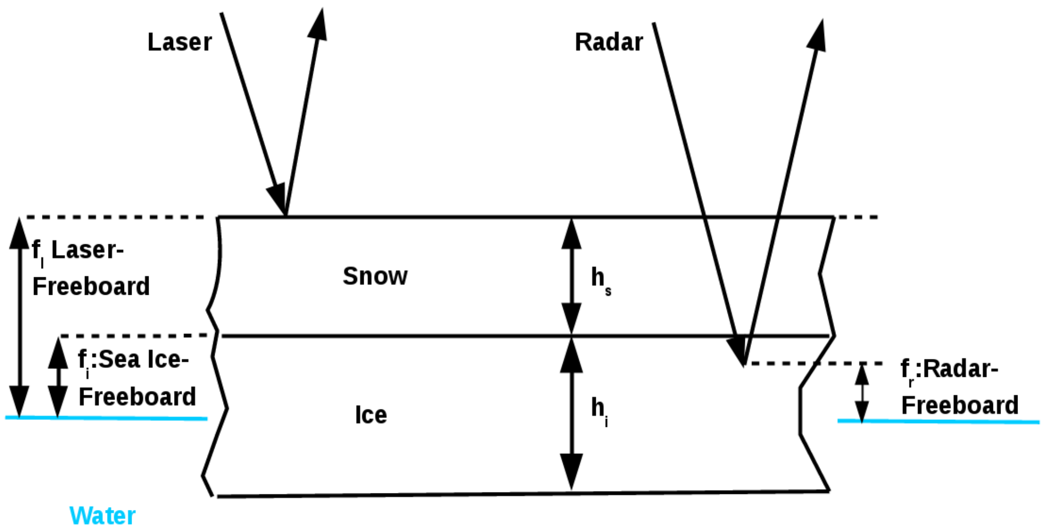


Figure 11. Schematic illustration of sea ice thickness and different freeboard variables.

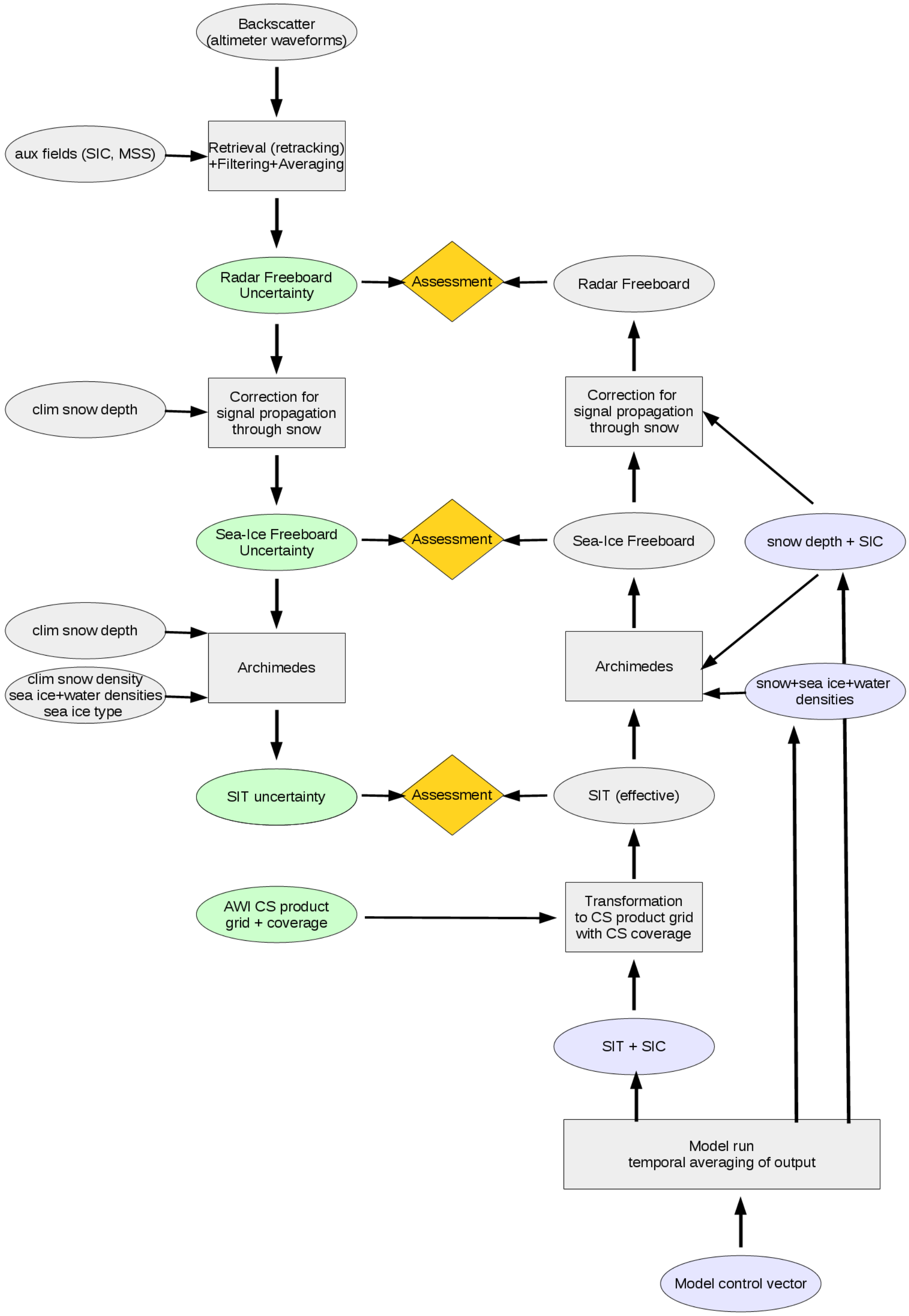


Figure 12. Overview on the processing chain for CryoSat-2 product retrievals (~~left hand~~ left-hand side) and the chain for modelling product equivalents (~~right hand~~ right-hand side). Oval boxes denote data and rectangular boxes processing steps. Green colour ~~indicates~~ emphasises remote sensing products and violet colour model variables. Yellow diamonds mark the assessment of the EO products with the QND algorithm. MSS: mean sea surface height.

The retrieval chain is described in detail by Ricker et al. (2014) and Hendricks et al. (2016). Recall that for each product, in order to run an assessment, we need the spatio-temporal coverage as well as the uncertainty ranges. The left-hand side of Figure 12 summarises the main steps in the retrieval chain, starting with the rawest (lowest-level) product (RFB) on top. When descending from RFB via SIFB to SIT each step adds further assumptions, which contribute to the product uncertainty. The other element required to evaluate a given product is the observational Jacobian, i.e. the sensitivity of the model simulation to a change in the control vector. The right-hand side of the graph illustrates how ~~this Jacobian is derived from the Jacobians of the~~ the equivalents of the respective products are simulated from the relevant model variables, which are denoted in violet colour. On this side of the graph, the complexity increases from bottom to top, i.e. from SIT via SIFB to RFB. For example, in the assessment of the SIT product, the uncertainty in quantities needed to apply the Archimedes' principle (including that of snow depth derived from climatology) is contained in the retrieval product, whereas the observation operator that extracts the product equivalent from the model is relatively simple (Archimedes' principle is described, for example, by Guerrier and Horley (1970)). We note that, while retrieved SIT is the effective SIT ($h_{i,eff}$), i.e. refers to the average over the ice-covered area of a grid cell, simulated SIT refers to the grid-cell average, i.e. for the Jacobian calculation it has to be divided by the simulated sea ice concentration (SIC, denoted by c):

$$h_{i,eff} = h_i/c. \tag{6}$$

Likewise for snow depth:

$$h_{s,eff} = h_s/c. \tag{7}$$

At the level of RFB, by contrast, it is the observation operator that includes inter alias, on the modelling ~~side~~branch, the application of Archimedes' principle ~~,~~ for which it requires simulated snow depth and the densities of snow (ρ_s), sea ice (ρ_i), and water (ρ_w), while the retrieval product is relatively raw. In particular ~~the~~ this retrieval product is not affected by uncertainties ~~caused by assumptions on~~ due to assumptions concerning the snow depth, ρ_s, ρ_i, and ρ_w.

The observation operators for f_i, for f_r, and for f_l are:

$$\begin{aligned} f_i &= h_i/c - (\rho_i h_i/c + \rho_s h_s/c)/\rho_w \\ &= (1-\rho_i/\rho_w)h_i/c - (\rho_s/\rho_w)h_s/c \qquad (8) \\ f_r &= f_i - 0.22h_s/c \\ &= (1-\rho_i/\rho_w)h_i/c - (0.22+\rho_s/\rho_w)h_s/c \qquad (9) \\ f_l &= f_i + h_s/c \\ &= (1-\rho_i/\rho_w)h_i/c + (1-\rho_s/\rho_w)h_s/c. \qquad (10) \end{aligned}$$

The term $-0.22h_s/c$ in Equation (9) adds to the simulated f_i the correction for the ~~signal propagation through snow~~ different propagation speed of the radar signal in snow compared to air, which is ~~contained in f_r~~ affecting f_r (Hendricks et al., 2016).

This is the reason why f_r is located below f_i in Figure 11. We note that, in these three observation operators, f_i, f_r, and f_l have the same sensitivity to h_i, but sensitivities to h_s and c differ. The sea ice component of the MPIOM uses constant densities of snow, sea ice, and water. As simulated freeboard is relatively sensitive to densities of snow and sea ice, we have, however, included these quantities as parameters of the observation operator in the control vector (see Section 2.4). For ρ_I=910.0 kg/m^3, ρ_S=330.0 kg/m^3, ρ_w=1025.0 kg/m^3, the sensitivity of f_i, f_r, and f_l to a change in h_i/c is $a = 0.112$, and the respective sensitivities to a change in h_s/c are $b = -0.322$, $b = -0.542$, and $b = 0.678$.

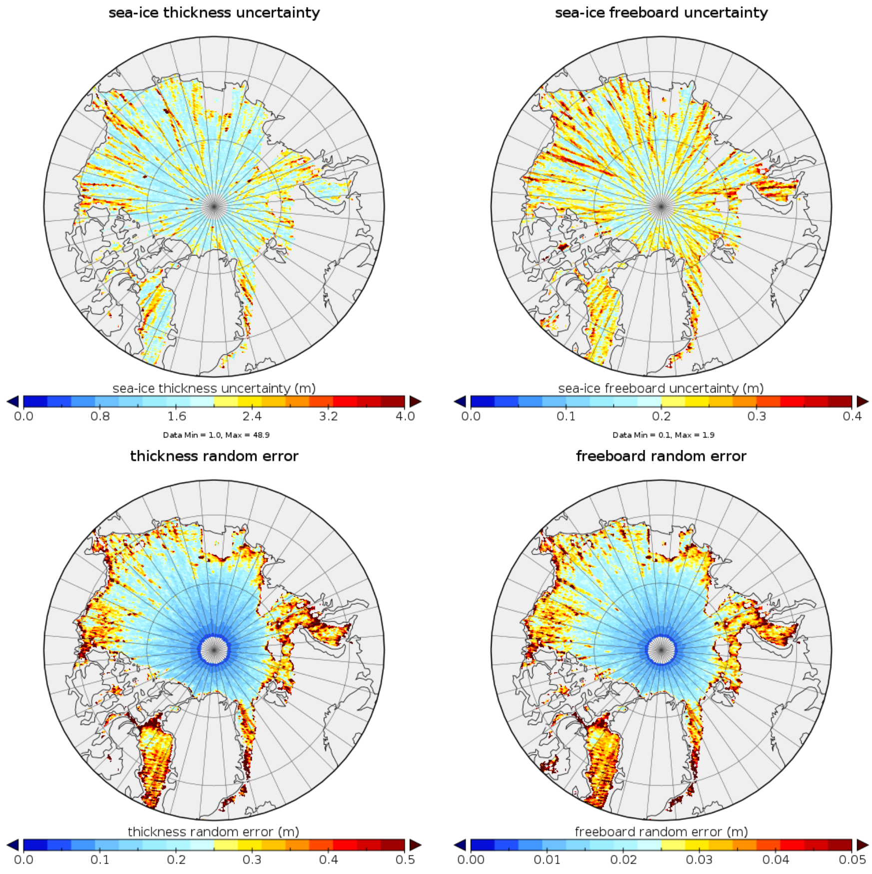


Figure 13. Uncertainty ranges of CryoSat-2 products: SIT (left), SIFB (right), total uncertainty (top), random component (bottom) ~~.~~ for April 2015.

The CryoSat-2 product files ~~provided by AWI~~ used in this study directly contain monthly SIT and SIFB on the ~~EASE~~ Equal-Area Scalable Earth Grid (EASE) 2.0 grid, respectively with random (based on standard uncertainty propagation) and total (random plus systematic) per-pixel uncertainty ranges (for details see Hendricks et al., 2016, and references therein). Figure 13 shows product uncertainties for April 2015. In our assessments we use the total uncertainties for the SIT and SIFB products, and for the RFB product the random uncertainty component of the SIFB product. ~~We assume uncertainties are~~ Recall, that we assume uncertainties to be uncorrelated in space.

For our hypothetical monthly LFB products, we assume a coverage of the northern hemisphere with a retrieved value over each cell of the EASE 2.0 grid with SIC above 0.7, in analogy to the threshold used in the CryoSat-2 retrieval (Hendricks et al., 2016). We explore two assumptions ~~on~~ with respect to the uncertainty of the products, a mission with a high accuracy (uniform uncertainty of 0.02 m) and a mission with low accuracy (uniform uncertainty of 0.20 m). In both cases uncertainties are uncorrelated in space.

For our hypothetical monthly mean ~~snow depth (SND)~~ SND products, we also assume a coverage of the northern hemisphere with a retrieved value over each cell of the EASE 2.0 grid with SIC above 0.7. As for LFB we explore two assumptions on the uncertainty of the products, a mission with a high accuracy (uniform uncertainty of 0.02 m) and a mission with low accuracy (uniform uncertainty of 0.15 m). In both cases uncertainties are uncorrelated in space.

Table 2 provides an overview on the products we assess. For later use~~is also lists , ~~, it also lists for each product and ~~three selected~~ the three control regions the number of sampled EASE 2.0 grid cells and the corresponding regional average uncertainties. Finally, it also shows the uncertainties on the spatial average of the sampled variable over all sampled EASE 2.0 grid cells based on the assumption of uncorrelated observational uncertainty.

Table 2. Overview on data sets, the # of sampled EASE 2.0 grids in control regions 5-7 (columns 2-4), the respective average uncertainties (columns 5-7), the uncertainty of the product aggregated over all sampled EASE 2.0 grid cells.

	n			average uncertainty			aggregated uncertainty
Product	5	6	7	5	6	7	[m]
SIT	937	1425	1377	1.86	1.95	1.94	0.0181
SIFB	937	1425	1377	0.21	0.20	0.21	0.00188
RFB	937	1425	1377	0.029	0.024	0.027	0.000364
LFB low accuracy	1104	1500	1429	0.20	0.20	0.20	0.00145
LFB high accuracy	1104	1500	1429	0.02	0.02	0.02	0.000145
SND low accuracy	1104	1500	1429	0.15	0.15	0.15	0.00108
SND high accuracy	1104	1500	1429	0.02	0.02	0.02	0.000145

3 Target and ~~Observation~~ observational Jacobians

We compute an observational Jacobian $\mathbf{M}'$ for each of the observational products we assess. For a given product, the observational Jacobian is computed in two steps. The first step performs the following actions: a reference run is performed using the prior control vector x, the input variables to the observation operator are stored over the observational period, aggregated to the model grid and the observation operator is applied to derive the observation equivalent $M(x)$ on the space-time grid of the observational product. In the second step, for each component of the control vector the following procedure is applied: A sensitivity run is performed with a control vector $x + p_i$ that is identical to the prior control vector but with the i component changed by a perturbation ~~ϵ~~ ϵ_i, and an observation equivalent $M(x + p_i)$ is computed in the same way as for the reference run. The Jacobian column is then computed as ~~$\sigma_i(M(x+p_i) - M(x))/\epsilon$~~ $\sigma_i(M(x+p_i) - M(x))/\epsilon_i$ where σ_i is the prior uncertainty of x_i. As a consequence of the normalisation by the prior uncertainty, each row in the Jacobian has the same unit as the respective observation. For a given product, column i of the corresponding observational Jacobian quantifies the sensitivity of the model-simulated equivalent to that product with respect to a ~~1-sigma~~ change of the i component of the control vector x_i by one standard deviation (see Table 1 for the value).

For any given product the dimension of the observational Jacobian is the product of the dimension of the control space and the grid size of the observational product. As an example, Figure 14 displays the row of the Jacobians for ~~a~~ April means of SIT, SIFB, RFB, LFB, and ~~snow depth (SND)~~ SND over a single point in space indicated by the black dot (and by the ~~yellow~~ black cross on Figure 3)~~)~~.

The SIT sensitivity is dominated by the model's initial SIT in control region 6 (black bars in Figure 14 and enlarged in Figure 15) but shows also considerable sensitivities to the initial SIC, the initial SND, the initial ocean temperature (TEMP) and the zonal wind stress (WIX). The negative sensitivity to SIC in that region is caused by two mechanisms. The first mechanism is expressed by Equation (6): The observation $h_{i,eff}$ is the effective SIT (thickness averaged over the ice-covered grid cell) and is reduced if the initial SIC is increased (and vice versa) because the model conserves the total sea ice volume. The second mechanism is related to ~~the formulation of~~ sea ice growth ~~in the model, which can grow~~ which depends on the open water fraction, i.e. more (less) sea ice can grow if the SIC is reduced (increased). The small negative sensitivity of SIT to SND is caused by the strong ~~insolation~~ insulation effect of snow, which hampers the growth of sea ice (or fosters the growth if SND is reduced).

The physical process behind the small negative sensitivities on the initial ocean temperature needs no further explanation; we recall, however, that, in the presence of sea ice, the control variable relates to a temperature change below the second model layer (~~in~~ at 17m depth). The negative sensitivity ~~on~~ with respect to the zonal wind stress (WIX) mirrors less advection of thick sea ice ~~stemming~~ originating from the Beaufort Gyre. WIX is positive for eastward wind stress. A positive perturbation on WIX is most distinct in region 6 (but also evident in regions 7 and 8) and slows down the Beaufort Gyre which advects less sea ice into the target region (sea ice behaves, at least in April and May, to a large extent like a rigid body, i.e. the impact in regions 7 and 8 acts almost instantaneously on the target regions) resulting in a negative sensitivity. The SIT sensitivities on model parameters (Figure 14 and enlarged in Figure 15) are very small compared to the sensitivities on the initial state or the atmospheric boundary conditions, as the short integration time (we sample the April mean of a model simulation starting on April 1) restricts the impact of the parameters.

The various freeboard products exhibit high sensitivity to initial SIT and SND (orange, red, and green bars in Figure 14). As SIT enters all freeboard observation operators in the same way (Section 2.5), the freeboard sensitivity ~~on~~ to April mean SIT is equal for all products, which also renders their sensitivity to initial SIT almost equal. The LFB sensitivity on the initial SND is positive (LFB is the freeboard at the top of the snow layer) while the sensitivity of the RFB and SIFB is negative because an increased SND will reduce the ~~radar~~ RFB and SIFB through the increased weight on the ice floe ~~.~~ (see Figure 11). Due to the definition of the observation operator for RFB (Equation (9)) its sensitivity to initial SND is larger than that of the SIFB (Section 2.5). The sensitivity of the freeboard products with respect to the parameters of the sea ice and ocean model is low. The impact of the sea ice density on the respective observation operators (Equation (8) to Equation (10)) is high, though, while ~~the sensitivity on~~ sensitivity with respect to the snow density is much lower (because the sea ice thickness is much larger than the SND at the observational point). The SND shows only considerable sensitivity to the initial SND in control region 6 and some minor positive sensitivity ~~on~~ with respect to the precipitation in the same region.

Likewise we computed target Jacobians $\mathbf{N}'$ for each of the six target quantities (SIV and SNV each over 3 regions) described in Section 2.2. Each target quantity is a scalar and thus the Jacobian has one entry for each component of the control vector. As an example Figure 16 displays the Jacobians for SIV and SNV over the Outer New Siberian Islands (ONSI) region. The first point to note is that sensitivities of regional SIV and SNV to the control vector differ, so an observation must constrain different components of the control vector to perform well for one or the other.

SIV over the ONSI region is highly sensitive to initial SIT over control regions 5 and 6 (Figure 17) which at least partly overlap with the target area. As the SIT observation and due to the same mechanisms discussed above, the SIV target quantity also exhibits a negative sensitivity to the initial SIC, SND, and zonal wind stress. It is interesting to note that SIV is also sensitive to initial and boundary conditions over more remote control regions. For example, it exhibits a positive sensitivity to the initial SIT in the control regions 1 and 7 from which thick sea ice is advected into the target region during the period from April 1 to May 28. This also explains the negative sensitivity to the zonal wind stress in region 7 and the meridional wind stress in region 1: For high enough concentration the sea ice almost behaves as an incompressible fluid allowing even for a sensitivity to wind stress changes in very remote control regions, e.g. the negative sensitivity to the zonal wind stress in region 8. The positive sensitivity to the zonal wind stress in region 1 (with thick ice) may be less obvious, as it follows the deflection of ice drift by about 20° to the right. The largest SIV sensitivity to model parameters (Figure 17) is found for the snow albedo of freezing conditions (albsn), but still that sensitivity is low compared ~~compared~~ to the sensitivity with respect to the initial state and atmospheric boundary conditions.

SNV shows particularly high sensitivity to the initial SND but also considerable sensitivity with respect to the precipitation and air temperature ~~(region 6).~~ in region 6. The largest model parameter sensitivity is found for the snow albedo for melting conditions: Increasing the snow albedo will reduce the melting.

4 ~~Sea ice and snow volume uncertainty~~Experimental setup

Based on the products shown in Table 2, we conducted assessments for the 15 cases listed in rows 4-18 of Table 3. ~~Row three ("prior") shows a reference case without observations, i.e. it shows the uncertainty in the target quantities that result from the prior uncertainty in the control vector. Then for each of the~~ These 15 cases cover all combinations of the five SIT/freeboard products described in Section 2.5:

1. SIT,
2. SIFB,

3. RFB,
4. hypothetical low accuracy LFB, and
5. hypothetical high accuracy LFB ~~there are three assessments, which cover the following cases:~~

and the following three assessments variants:

1. product evaluated individually,
2. product evaluated together with a hypothetical low accuracy SND product, and
3. product evaluated together with a hypothetical high accuracy SND product.

The reference for these assessments is a case without observations. Row three ("prior") shows the uncertainties in the target quantities that result from the prior uncertainty in the control vector.

5 Sea ice and snow volume uncertainty

As explained in Section 2.1 the uncertainty component from the model error $\sigma(y_{\mathrm{mod}})$ in Equation (3) covers the residual uncertainty that remains with an optimal control vector, i.e. it reflects uncertainty from uncertain aspects not included in the model error and structural uncertainty reflecting wrong or missing process formulations. $\sigma(y_{\mathrm{mod}})$ is model-dependent and is probably the most subjective component in the prior and posterior uncertainties. $\sigma(y_{\mathrm{mod}})$ acts as an offset (for the respective variance) for all cases, and reduces the contrast between the cases. As ~~the focus in our assessments lies~~ our assessments focus on the differences between the cases, we exclude it from the target uncertainties in rows 3-18 of Table 3 and provide estimates in separate rows. To illustrate the subjective nature of this estimate and possible ranges, we derive two crude estimates (last two rows). The first estimate (denoted by $\sigma_{\text{mod, absolute}}$ and listed in the last but one row) assumes a model that perfectly simulates the same ice-covered area of all three regions as our model and that, over this area, achieves an uncertainty of 0.2 m for SIT and of 0.1 m for SND. The second estimate (denoted by $\sigma_{\text{mod, relative}}$~~and listed in the last row~~) assumes a model that simulates the same SIV and SNV as our model with an uncertainty of 10% for SIV and 30% for SNV. We use a higher uncertainty for SNV because it has a stronger dependence on the surface forcing (mainly precipitation), for which the temporal and small-scale spatial structures are not resolved in the control vector.

Figure 18 shows the uncertainty reduction with respect to the prior case as defined in Equation (5) for both SIV and SNV and all three target regions. A value of 100% means that the product has resolved all uncertainty in the respective target quantity, while a value of 0% means that the product was not useful to improve the forecast of the target quantity. We first discuss the single product assessments, i.e. without additional use of a hypothetical snow product. For all three regions, the SIT has considerably better performance for SIV than for SNV. Between SIV and SNV the only difference consists in the target Jacobians, $\mathbf{N}'$. For example for target region ONSI, Figure 16 shows particularly high sensitivity of SIV to initial SIT and of SNV to initial SND in control regions ~~6 and 5.~~ 5 and 6. Hence, to constrain SIV (SNV) over that target region a product has

Table 3. Prior and posterior uncertainties of sea ice volume (SIV, columns 4-6) and snow volume (SNV, columns 7-9) respectively for three regions in km^3. Column 1 indicates observation, column 2 indicates uncertainty range ("product" refers to uncertainty specification provided with product), column 3 indicates uncertainty range of additional hypothetical snow product ("–" means no snow product is used). In each of columns 4-9 the lowest uncertainty range is highlighted in bold face font. The two bottom rows give estimates for the uncertainty due to model error, i.e. the residual uncertainty with optimal control vector.

			SIV			SNV		
Observation	σ[m]	$\sigma(h_s)$[m]	WLS	ONSI	ESS	WLS	ONSI	ESS
Prior	-	-	136.5	131.6	289.6	62.3	63.3	110.1
SIT	product	-	28.7	34.3	94.4	59.5	61.3	107.9
SIT	product	0.15	19.8	22.4	62.6	11.0	11.8	21.4
SIT	product	0.02	12.4	10.4	24.1	2.4	2.5	4.5
Sea Ice Freeboard	product	-	86.4	84.1	203.4	40.4	39.8	75.2
Sea Ice Freeboard	product	0.15	21.5	25.0	67.7	11.0	11.8	21.4
Sea Ice Freeboard	product	0.02	12.6	11.0	25.3	2.4	2.5	4.5
Radar Freeboard	product	-	51.3	39.2	93.8	16.4	14.2	26.0
Radar Freeboard	product	0.15	8.8	10.9	34.7	8.0	8.3	16.6
Radar Freeboard	product	0.02	3.0	3.8	12.4	2.2	2.3	4.4
Laser Freeboard	0.20	-	81.0	67.0	143.9	17.7	17.1	30.8
Laser Freeboard	0.20	0.15	20.4	22.1	57.8	9.0	9.6	17.7
Laser Freeboard	0.20	0.02	12.2	10.7	24.8	2.3	2.4	4.5
Laser Freeboard	0.02	-	11.5	9.0	20.0	2.5	2.3	4.2
Laser Freeboard	0.02	0.15	6.6	6.0	14.6	1.9	2.0	3.7
Laser Freeboard	0.02	0.02	**2.4**	**2.7**	**8.3**	**1.3**	**1.4**	**2.6**
$\sigma_{\text{mod, absolute}}$	-	-	30.3	36.2	73.5	15.1	18.1	36.8
$\sigma_{\text{mod, relative}}$	-	-	48.7	70.8	165.9	10.2	11.4	5.3

to constrain primarily initial SIT (SND) over these two control regions. Figure 14 shows that, indeed, SIT provides a much stronger constraint on initial SIT than on initial SND. ~~By~~ In contrast to SIT, SIFB has similar performance for SIV and SNV, over all target regions (Figure 18). Compared to SIT, SIFB shows a much lower sensitivity to initial SIT but a higher sensitivity to initial SND (Figure 14 - the sign of the sensitivity is irrelevant in this consideration), and thus a more balanced performance for SIV and SNV than the SIT product. RFB and the two hypothetical LFB products achieve a better performance for SNV than for SIV. The only difference between the RFB and SIFB Jacobians is the larger impact of h_s/c for RFB, as a consequence of the correction for the signal propagation through snow (see Section 2.5). ~~This is the reason~~Hence, why RFB shows a better performance for SNV than for SIV, while SIFB had about equal performance for SIV and SNV. LFB has the same sensitivity to initial SIT as RFB but an even larger sensitivity to initial SND. Consequently, for the low accuracy LFB product, the imbalance between the performance for SIV and SNV is even higher than for the RFB product. ~~The~~ This imbalance is lower for the high accuracy LFB product~~has so good performance on SIV already,~~ , because this product already performs excellently on SIV such that there is not much scope for ~~yet better~~ further increases in performance on SNV.

So far we have discussed ~~for a given product the~~ differences in performance for SIV and SNV for a given product. Next we address performance differences between products. ~~The first thing to note is that the step~~ First, we note that switching from SIT to SIFB drastically reduces the performance for SIV. As explained in Section 2.5, on the ~~retrieval~~ left-hand side of Figure 12 ~~the step~~ (retrieval branch) switching from SIFB to SIT applies Archimedes' principle, with uncertain assumptions primarily on the input variables snow and ice density and snow depth, which yield an increase in product uncertainty by about an order of magnitude (Figure 13 and Table 2). On the ~~modelling~~ right-hand side of Figure 12 ~~the step~~ (modelling branch) switching from SIT to SIFB is dealing with uncertainty on the same input variables (snow and ice densities and snow depth), which renders the simulation of SIFB more uncertain than that of SIT. In the model, the uncertainty in these variables is determined by the prior uncertainty of the control vector, either directly (snow and ice densities) or indirectly (snow depth) through their model-simulated dependency on the control vector. It appears that the increase in uncertainty, when going from SIT to SIFB on the modelling ~~side~~branch, overcompensates for the reduction in uncertainty on the retrieval side, when going back from SIT to SIFB. In other words, on the modelling ~~side~~branch, the assumptions on uncertain input appear more conservative than those on the retrieval ~~side~~branch. On the retrieval ~~side~~ branch going (backwards) from SIFB to RFB consists in a reduction of product uncertainty by about another order of magnitude, as the retrieval of RFB does not require information on snow depth. Even with this further reduction of product uncertainty, the performance of RFB is inferior to that of SIT for SIV over WLS and ONSI, and only just superior for SIV over ESS.

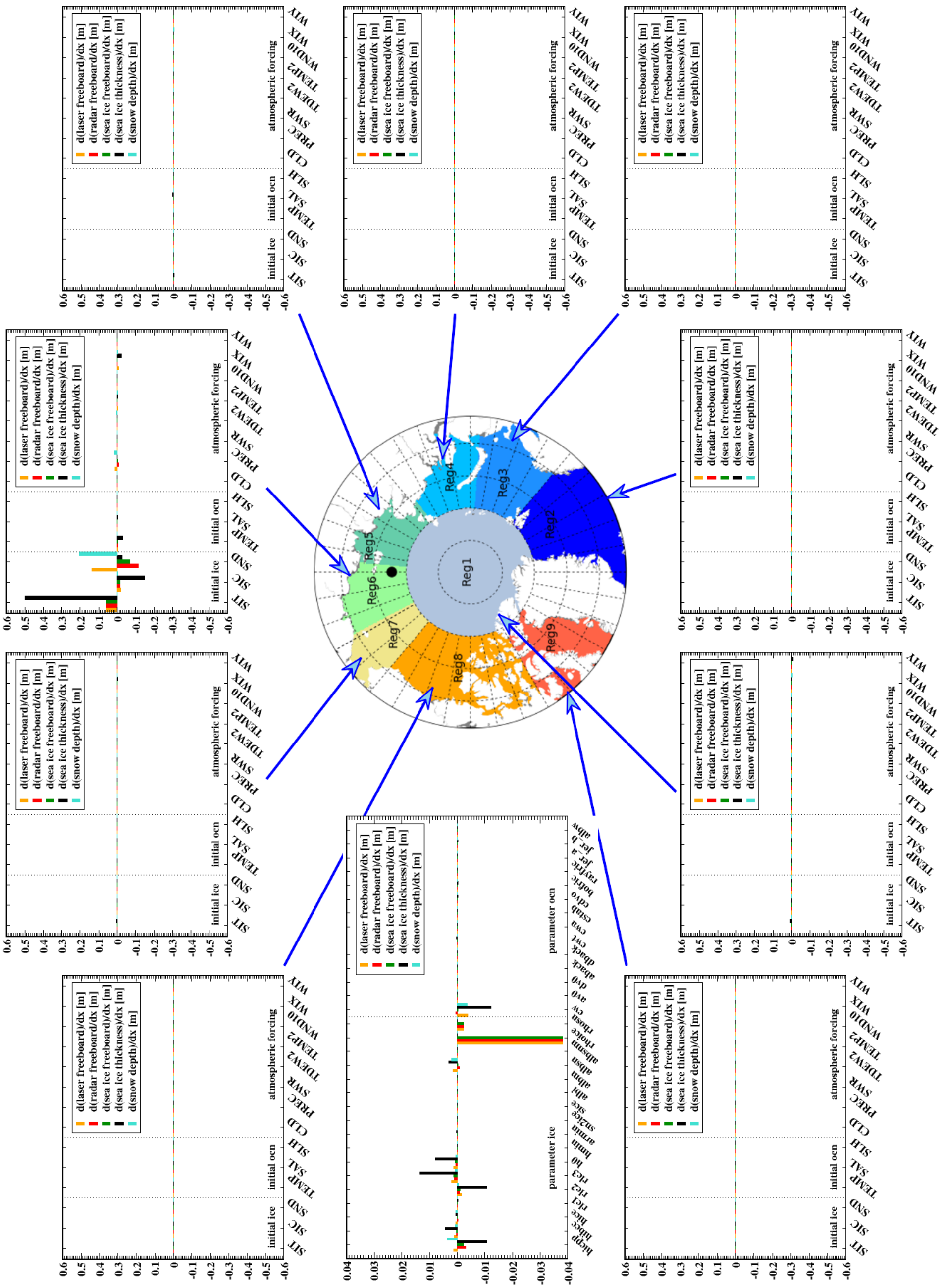
Differences between target regions in the performance of the same product are the result of a complex interplay of the Jacobians $\mathbf{N}'$ for the target regions and the product's constraint on the control vector quantified by $\mathbf{C}(x)$ (see Equation (3)). For each of the target regions a different (combination) of control regions is most relevant: For WLS this is control region 5 (not shown), for ONSI control ~~region~~ regions 5 and 6 (Figure 16 and enlarged in Figure 17) and for ESS on control ~~region~~ regions 6 and 7 (not shown). The ability of a product to constrain a particular control region is determined by the combination of the observational Jacobian of the product and the product uncertainty (see Equation (1)).

~~One is always tempted~~ It is tempting to explain regional performance differences ~~in a simple way, just from~~ simplistically by linking them to differences in observational coverage and uncertainty. Technically~~this means to replace~~ , such an explanation corresponds to replacing our observational Jacobian $\mathbf{M}'$ that is based on model dynamics ~~by~~ with a drastically simplified representation~~based on the assumptions~~. Such a simplistic approach would imply that only observations over a given control region ~~do constrain that~~ constrain that same region (and ~~no otherregion~~none other), and that the observational Jacobian for each product and control variable is spatially uniform. The ~~constraint~~ constraints of a product on a control region would then be proportional to the square root of the number of samples n of that region and to the reciprocal of the average observational uncertainty $\overline{\sigma}$ over the region. Table 2 shows both impact factors for the most relevant control regions, i.e. 5-7. For RFB and compared to region 6, the relevant quantity $\sqrt{n}/\overline{\sigma}$ is about 41% lower in region 5 and 12% lower in region 7. This is at least quantitatively in line with the performance decrease for RFB and SIV from ONSI (most relevant ~~regions:~~ in region 6 and to smaller extent in 5) to ESS (most relevant ~~regions:~~ in region 6 and to smaller extent in 7) to WLS (most relevant ~~region:~~ in region 5). But the performance ranking for RFB and SNV is different, i.e. the ~~simple explanation already fails~~simplistic approach does not hold. Also for SIT, the differences in $\sqrt{n}/\overline{\sigma}$ between the three control regions are smaller and fail to explain the performance decrease from WLS to ONSI to ESS. Such calculations demonstrate the limits of a performance assessment that is only based on observational coverage and uncertainty, while neglecting the model dynamics.

The two hypothetical LFB products have a slightly better spatial coverage of the most relevant control regions than the products derived from CryoSat-2 and use uniform data uncertainties that span the range from 2 cm (high accuracy LFB) to 20 cm (low accuracy LFB). ~~We need to recall here~~ Recall that the specified data uncertainty combines ~~()~~ the observational uncertainty (i.e. product uncertainty) with the residual model uncertainty due to structural errors and uncertain contributions not accounted for in the control vector (Equation (2)). Only the high accuracy LFB can clearly outperform all CryoSat-2 products for both SIV and SNV and over all three regions, while the low accuracy LFB is between that of SIFB and RFB.

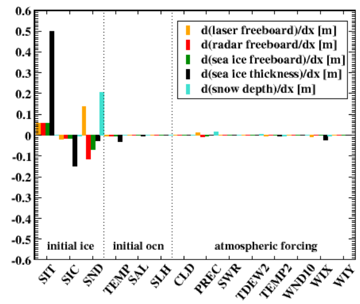
Next we discuss the effect of combining either of these five products with the two hypothetical SND products. The difference in the ~~(sample row of the)~~ respective product Jacobians shown in Figure 14 suggests complementarity of SND to the SIT and ~~FB~~ freeboard products. Indeed, the combination with SND considerably increases the performance of all SIT/freeboard products for SIV and SNV and over all regions. Most striking is the ~~lift of the~~ improved SIT performance for SNV. The combination with SND results in similar performance for SIT and SIFB, slightly better performance of low accuracy LFB, yet slightly better performance for RFB and the best performance for the high accuracy LFB. The assessment for SIV and in combination with low accuracy SND yields the same performance ranking of products, with slightly larger differences between products. Combining with the high accuracy SND product instead of the low accuracy SND product yields a performance gain for all products and for SIV and SNV over all regions.

Between the two LFB products, the increase in accuracy yields a considerable performance gain for SIV and SND over all regions, when assessed individually and in combination with SND. ~~The~~ Over all regions the combination of the high accuracy LFB with the low accuracy SND performs better for SIV ~~and SNV over all regions~~than the combination of the low accuracy LFB with the high accuracy SND. For SNV the two combinations are similar in performance.



~~Jacobian rows~~

Figure 14. The sensitivities of the respective EO product to the control vector ("observational Jacobians") for ~~a~~ April means of ~~SIT~~LFB (orange bars), ~~SIFB,~~ RFB (red bars), SIFB (green bars), SIT (black bars) and ~~LFB~~ SND (cyan bars) over a single point indicated by the black dot (and by ~~yellow~~ black cross on Figure 3). The observational Jacobians with respect to the process parameters are shown in the left middle panel. The other panels show the observational Jacobians with respect to the initial and forcing fields (see Table 1 for an explanation of the abbreviations).



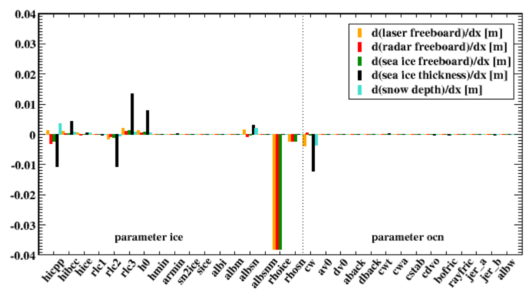


Figure 15. An excerpt of Figure 14 of the observational Jacobian (top) in target region 6 ("East Siberian Sea") and (bottom) for the model parameters.

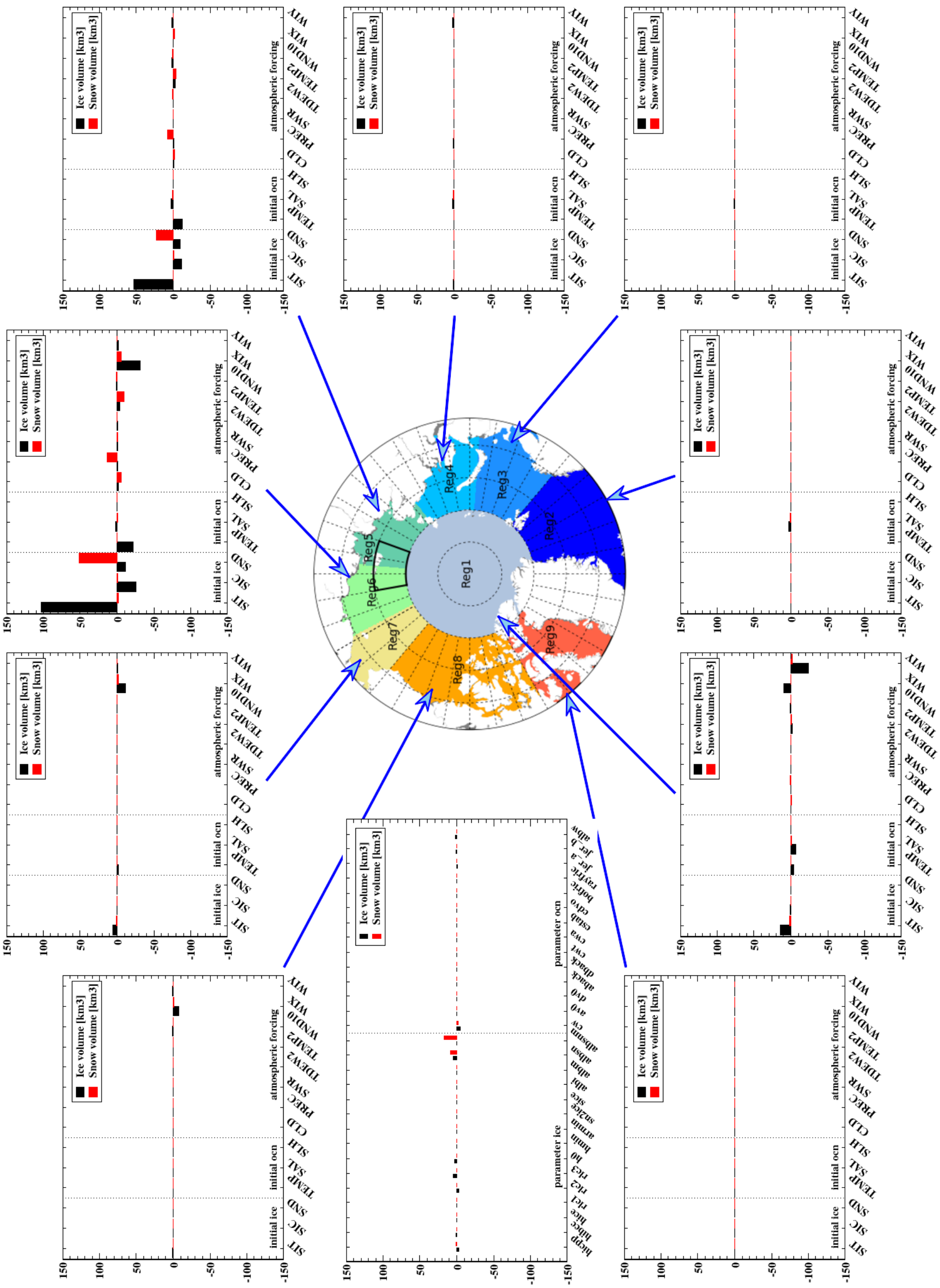


Figure 16. ~~Jacobians~~ As Figure 14 but for the sensitivities of the sea ice (SIV) and snow (SNV) volume over the target region Outer New Siberian Islands (ONSI) on the control vector ("target Jacobians").

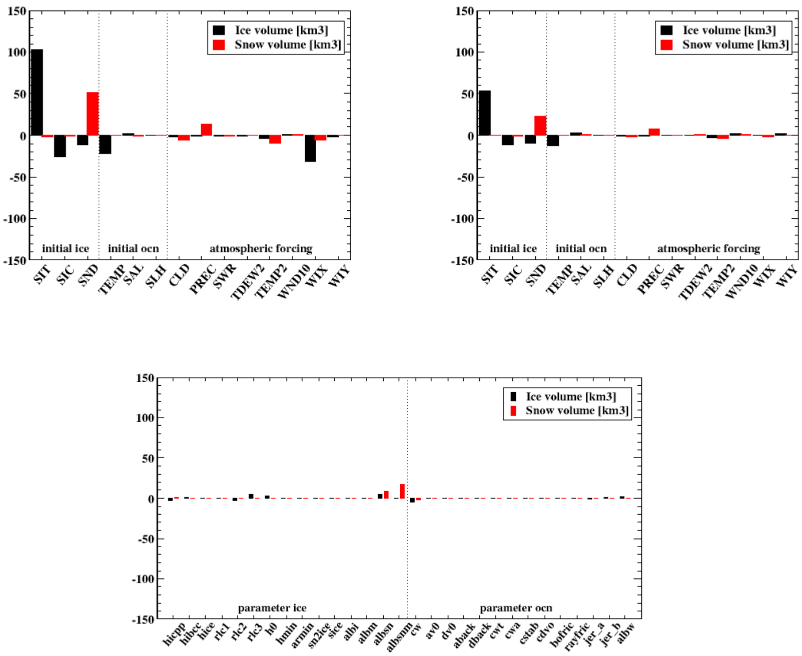


Figure 17. An excerpt of Figure 16 for the target Jacobian (top) in target region 6 (“East Siberian Sea” - left) and region 5 (“Laptev Sea” - right) and (bottom) for the model parameters.

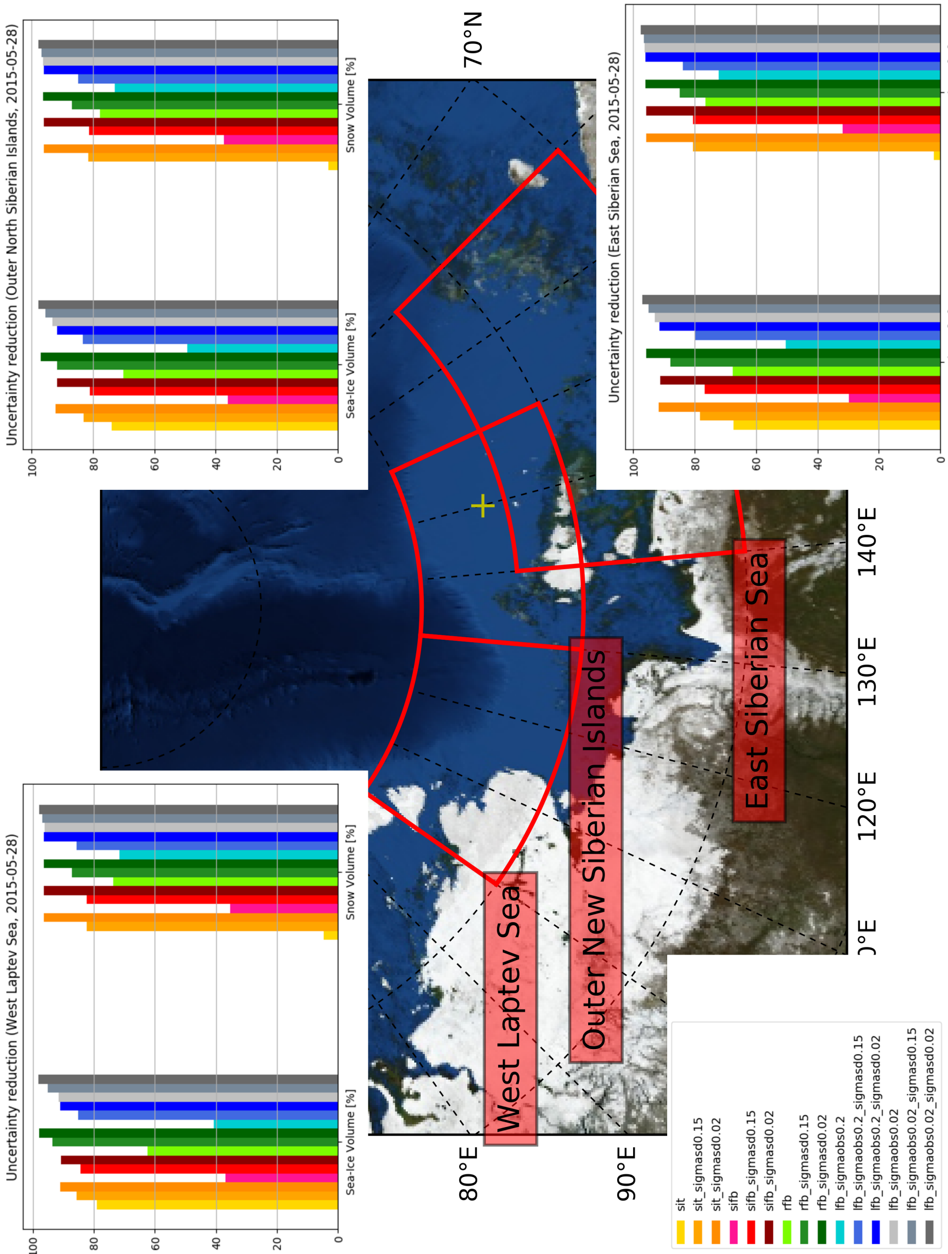


Figure 18. Uncertainty reduction for sea ice (SIV) and snow (SNV) volume over target regions when using observational constraints. The colour of the bars represents the different observational constraints. Yellowish: SIT and SIT in combination with a hypothetical snow depth product with two different uncertainties ($15cm$ and $2cm$), reddish: as SIT but for SIFB, greenish: as SIT but for RFB, bluish: as SIT but for hypothetical LFB with $20cm$ uncertainty, grayish: as SIT but for hypothetical LFB with $2cm$ uncertainty.

6 Discussion

There are a number of factors in the setup of our ArcMBA system that impact our assessments. One of them is the model that is required to realistically compute the sensitivities (Jacobians) of the target quantities and of the observation equivalents to changes in the control vector. As detailed in Section 2.3, the MPIOM has a state of the art representation of processes, compares reasonably ~~well~~ with a range of observations (Notz et al., 2013; Kaminski et al., 2017), and the setup of MPIOM we are using has a spatial resolution below the grid size of the observations and well below the size of the target regions. The model thus appears appropriate for our study and the ArcMBA system in general. Nevertheless, through the Jacobians the results depend on the model, and it would be useful to confirm the robustness of the assessments through the use of a second model, or even an ensemble of models.

The study has investigated the performance of four-week forecasts in May 2015. ~~It would be interesting to analyse if the relative performance of the products varies from year to year and with the length of the forecasting period, and how the products perform for different~~ The impact of an observation is likely to depend on the state of the Arctic sea-ice ocean system. The robustness of the ArcMBA assessments can thus be increased through extension of the system for an ensemble of ice and ocean conditions representing different forecasting times (for example 2, 7-10, and 90 days and also 0 days, i.e. an analysis), different seasons, different typical years (potentially also including conditions of very low ice cover), and different target regions and variables, e.g. SIC.

In our setup, the control vector has 157 components. In particular within any of our 9 control regions we do not resolve changes in the spatial patterns of the initial conditions nor in the spatio-temporal patterns of the forcing data. This means that we are ignoring uncertain aspects in the inputs that determine our simulation, which results in so-called aggregation errors (Trampert and Snieder, 1996; Kaminski et al., 2001) and renders the ArcMBA assessments of the product impacts too optimistic. As the target quantities are integrals over large regions, we expect, however, that our control regions can capture most of the uncertainty. Also the set of reasonable surface forcings is in practise limited by physical relations between variables, in space ~~,~~ and in time. Similar restrictions apply to the initial state. Further, we use the same control vector for all cases, so that the relative performance with respect to the prior (uncertainty reduction) and among products is more reliable. Nevertheless it appears useful to explore extended control vectors, for example with decreased sizes of the control regions, in particular in areas with high impact on observations or target quantities.

Another factor that impacts our assessments are correlations in the data uncertainty. These uncertainty correlations are difficult to estimate. We used zero correlation for each of the products, which is certainly the most optimistic assumption and yields the best performance. As we made this assumption consistently for all products, the relative performance between the products is less affected. To illustrate the implications of uncorrelated uncertainty in the products, we have computed the resulting uncertainty in the respective average of each observable over all sampled grid cells (last column of Table 2). This yields for the April 2015 mean of SIT about ~~2cm~~2 cm, of SIFB about ~~2mm~~2 mm, of RFB about ~~0.4mm~~ 0.4 mm and (using the respective high-accuracy versions) of LFB and SND about ~~0.1mm~~0.1 mm.

The effect of uncertainty correlation on the assessments can be demonstrated also by the following simplified calculation: If we partition our product grid into groups of n by n pixels, and assume perfect uncertainty correlation and the same Jacobian for each observation within a given group, then we decrease the first term in Equation (1) by a factor of n^2. This case then yields the same results as a case with an uncertainty that is uncorrelated and increased by a factor of n. This means we can interpret the impact of the low resolution LFB product (uncorrelated uncertainty of 20 cm) as the impact of a high resolution LFB product (with 10 times lower uncertainty, i.e. 2 cm) in which the uncertainty within each 10 by 10 group of pixels is completely correlated. Likewise for the SND product and (roughly) 6 by 9 groups of pixels, because (15 cm/2 cm)^2 is about 6×9. One reason for spatial uncertainty correlation would be a sensor footprint that exceeds the size of a 25km EASE grid cell. Likewise, for sensors with footprints considerably smaller than a 25km EASE grid cell, the procedure for upscaling from the sampled fraction of a grid cell to a grid-cell average could suffer from systematic errors that affect large scales in the same way, which would result in large-scale uncertainty correlations.

Our hypothetical products (LFB and SND) observe every pixel with SIC above 0.7. This is optimistic but, at least for snow, not totally unrealistic, depending on the mission concept. Recalling that the data uncertainty has to include also an uncertainty from model error, the value of 0.02 m for the high accuracy products (without spatial correlation) is extreme and unrealistic (as it is already a challenging requirement on the observational uncertainty) but still useful as a sanity check ~~.~~ for the methodology. We note that, even for the assessment of an individual product, the posterior uncertainty on the target quantities is not a simple linear function of the product uncertainty, because of the contribution from the prior term in Equation (1). This means, for example, the posterior uncertainties achieved by a hypothetical LFB product with 0.11 cm uncertainty will not be the average of the posterior uncertainties achieved by our two hypothetical SND products with respective uncertainties of 0.02 m and 0.20 m. For combined assessment of multiple products the relation between the uncertainty of a single product and the posterior target uncertainty is yet more complex.

The uncertainty specified with the SIT product is higher than the uncertainty specified with the SIFB products derived from CryoSat-2. This increase reflects the inclusion of uncertainty input quantities for the application of Archimedes' principle, in particular of climatological snow depth. In the assessment of SIFB~~Archimides~~, Archimedes' principle is applied in the observation operator, where the input quantities including snow depth are taken from the model. The fact that the impact of SIFB on SIV is lower than ~~than~~ that of SIT indicates that the assumptions on the uncertainty of input quantities for the application of Archimedes' principle is more conservative on the modelling ~~side~~ branch than those that were made on the retrieval ~~side~~ branch (yielding the respective product uncertainties). More conservative assumptions on the retrieval ~~side~~ branch would yield higher uncertainty in the SIT product. We also note that ~~the use of the RFB product that~~ biases may be reduced when using the RFB product instead, as it does not rely on an external snow depth climatology but uses a consistently simulated snow depth~~may reduce biases~~.

7 Summary and conclusions

The ArcMBA tool was used to assess the impact of a series of EO products on the quality of four-week forecasts of SIV and SNV over three regions along the ~~Northern Sea Route~~ NSR in May 2015. The tool is built around the MPIOM, a coupled model of the sea ice-ocean system extended by observation operators that link the simulated variables to equivalents of SIT, SIFB, RFB, LFB, and SND products.

On the basis of the per-pixel uncertainty ranges that are provided with the CryoSat-2 SIT, SIFB, and RFB products, the SIT ~~achieves a much better performance for SIV than the SIFB product, while the performance of RFB is more similar to that of SIT. For SNV~~and RFB products achieve the best performance for the target SIV. For the SNV target, the performance of SIT is only low, the performance of SIFB higher and the performance of RFB yet higher. A hypothetical LFB product with low accuracy (20 cm uncertainty) lies in performance between SIFB and RFB for both SIV and SNV. A reduction in the uncertainty of the LFB product to 2 cm yields a significant increase in performance.

Combining either of the SIT/freeboard products with a hypothetical SND product achieves a significant performance increase. The uncertainty in the SND product matters: A higher accuracy product achieves an extra performance gain.

The provision of spatial and temporal uncertainty correlations with the EO products would be beneficial not only for assessments within systems like the ArcMBA, but also for assimilation of the products. For example, complete uncertainty correlation within each group of 10 by 10 pixels (with uniform Jacobians within each group) is equivalent to an uncertainty increase by a factor of 10.

The ArcMBA can be extended to cover further EO products and further target variables. In the setup used here the model can simulate a range of sea ice-ocean variables in addition to those considered in the present study (e.g. ice drift, mixed layer depth, freshwater, sea surface salinity, sea surface temperature, or circulation). Switching to a more comprehensive model configuration would enable the investigation of yet further variables. For example the model can be operated with its biogeochemistry module HAMOCC (Ilyina et al., 2013) or in a mode coupled to an atmospheric general circulation model.

The study has investigated the performance of four-week forecasts in May 2015. It would be interesting to analyse how the relative performance of the products varies from year to year, with the length of the forecasting period, for other target regions, and with a different sea ice-ocean model.

As the QND approach can evaluate a (group of) hypothetical product(s) - characterised by their space-time coverage and uncertainty ranges - it is generally suited to assess the benefit of filling a given observational gap. It provides answers to hypothetical questions such as: *"Provided we could derive a product of a given variable with a given accuracy, at a given sampling frequency and spatial coverage. What is the added value of this additional observation for the quality of sea ice forecasts (as quantified by uncertainty reduction in a set of predicted target quantities)?"* ArcMBA assessments of a set of such products (each filling an observational gap) can help to establish a priority list.

The ArcMBA system is an ideal framework to assist the formulation of mission requirements or the development of EO products. ~~Through~~ In an end-to-end simulation it can translate product specifications in terms of spatio-temporal resolution and coverage, accuracy, and precision into a range of performance metrics. Alternatively, it can translate requirements on forecast performance into requirements on the respective observables. As demonstrated in the present study, the joint assessment of products from (constellations of) multiple satellites is one of the particular strengths of the ArcMBA approach. This type of assessment can be performed for higher-level products (e.g. SIT or SIC) but also for rawer products (e.g. freeboard or brightness temperature).

Acknowledgements. This work was funded by the European Space Agency.

This is a contribution to the Year of Polar Prediction (YOPP), a flagship activity of the Polar Prediction Project (PPP), initiated by the World Weather Research Programme (WWRP) of the World Meteorological Organsation (WMO).

References

Bovensmann, H., Bösch, H., Brunner, D., Ciais, P., Crisp, D., Dolman, H., Hayman, G., Houweling, S., and Lichtenberg, L.: Report for mission selection: CarbonSat - An earth explorer to observe greenhouse gases, Tech. rep., Noordwijk, The Netherlands, http://nora.nerc.ac.uk/514012/, freely available via Official URL link., 2015.

Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Zuo, H., Balmaseda, M., Chang, Y.-S., Ferry, N., Garric, G., Haines, K., Keeley, S., Kovach, R. M., Kuragano, T., Masina, S., Tang, Y., Tsujino, H., and Wang, X.: Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project, Climate Dynamics, 49, 1107–1136, https://doi.org/10.1007/s00382-016-2985-y, https://doi.org/10.1007/s00382-016-2985-y, 2017.

Crisp, D., Atlas, R., Breon, F.-M., Brown, L., Burrows, J., Ciais, P., Connor, B., Doney, S., Fung, I., Jacob, D., Miller, C., O'Brien, D., Pawson, S., Randerson, J., Rayner, P., Salawitch, R., Sander, S., Sen, B., Stephens, G., Tans, P., Toon, G., Wennberg, P., Wofsy, S., Yung, Y., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission. Trace Constituents in the Troposphere and Lower Stratosphere, Advances in Space Research, 34, 700 – 709, https://doi.org/http://dx.doi.org/10.1016/j.asr.2003.08.062, http://www.sciencedirect.com/science/article/pii/S0273117704003539, 2004.

Day, J. J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, Geophysical Research Letters, 41, 7566–7575, https://doi.org/10.1002/2014GL061694, http://dx.doi.org/10.1002/2014GL061694, 2014.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Eastwood, S., Jenssen, M., Lavergne, T., Sorensen, A., and Tonboe: EUMETSAT Ocean and Sea Ice Satellite Application Facility. Global sea ice concentration reprocessing product (v1.2), Product user manual, Technical Report, Norwegian and Danish Meteorological Institutes, Oslo, Norway and Copenhagen, Denmark, 14, 2079–2087, 2015.

Eicken, H.: Arctic sea ice needs better forecasts, Nature, 497, 431–433, https://doi.org/10.1038/497431a, 2013.

Feng, L., Palmer, P. I., Bösch, H., and Dance, S.: Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman Filter, Atmospheric Chemistry and Physics, 9, 2619–2633, https://doi.org/10.5194/acp-9-2619-2009, http://www.atmos-chem-phys.net/9/2619/2009/, 2009.

Gent, P. and McWilliams, J.: Isopycnal mixing in ocean circulation models, J. Phys. Oceanogr., 20, 150–155, 1990.

Gent, P., Willebrand, J., McDougall, T., and McWilliams, J.: Parameterizing eddy-induced tracer transport in ocean circulation models, J. Phys. Oceanogr., 25, 463–474, 1995.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, Journal of Advances in Modeling Earth Systems, 5, 572–597, https://doi.org/10.1002/jame.20038, http://dx.doi.org/10.1002/jame.20038, 2013.

Griffies, S. M.: The Gent-McWilliams skew flux, J. Phys. Oceanogr., 28, 831–841, 1998.

Guerrier, D. and Horley, F.: Archimedes: Archimedes' Principle and the Law of Flotation, Discovering with the Scientists Series, Blond & Briggs, https://books.google.de/books?id=FOa_AAAACAAJ, 1970.

Haak, H., Jungclaus, J., Mikolajewicz, U., and Latif, M.: Formation and propagation of great salinity anomalies, Geophysical Research Letters, 30, n/a–n/a, https://doi.org/10.1029/2003GL017065, http://dx.doi.org/10.1029/2003GL017065, 1473, 2003.

Haas, C., Beckers, J., King, J., Silis, A., Stroeve, J., Wilkinson, J., Notenboom, B., Schweiger, A., and Hendricks, S.: Ice and Snow Thickness Variability and Change in the High Arctic Ocean Observed by In Situ Measurements, Geophysical Research Letters, 44, 10,462–10,469, https://doi.org/10.1002/2017GL075434, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL075434, 2017.

Hardt, M. and Scherbaum, F.: The Design of Optimum Networks for Aftershock Recordings, Geophys. J. Int., 117, 716–726, 1994.

Hendricks, S., Ricker, R., and Helm, V.: User Guide - AWI CryoSat-2 Sea Ice Thickness Data Product (v1.2), 2016.

Hibler, W.: A dynamic thermodynamic sea ice model, Journal Geophysical Research, 9, 815–846, 1979.

Houweling, S., Breon, F.-M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., and Ciais, P.: Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, Atmos. Chem. Phys., 4, 523–538, 2004.

Hungershoefer, K., Breon, F.-M., Peylin, P., Chevallier, F., Rayner, P., Klonecki, A., Houweling, S., and Marshall, J.: Evaluation of various observing systems for the global monitoring of CO_2 surface fluxes, Atmospheric Chemistry and Physics, 10, 10 503–10 520, https://doi.org/10.5194/acp-10-10503-2010, http://www.atmos-chem-phys.net/10/10503/2010/, 2010.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Núñez-Riboni, I.: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations, Journal of Advances in Modeling Earth Systems, 5, 287–315, https://doi.org/10.1029/2012MS000178, http://dx.doi.org/10.1029/2012MS000178, 2013.

Jungclaus, J., Giorgetta, M., Reick, C., Legutke, S., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Fischer, N., Glushak, K., Gayler, V., Haak, H., Hollweg, H., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Müller, W., Notz, D., Pohlmann, T., Raddatz, T., Rast, S., Roeckner, E., Salzmann, M., Schmidt, H., Schnur, R., Segschneider, J., Six, K., Stockhause, M., Wegner, J., Widmann, H., Wieners, K., Claussen, M., Marotzke, J., and Stevens, B.: CMIP5 simulations of the Max Planck Institute for Meteorology (MPI-M) based on the MPI-ESM-P model: The 1pctCO2 experiment, served by ESGF, WDCC at DKRZ, https://doi.org/10.1594/WDCC/CMIP5.MXEPc1, 2012.

Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J.-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM, Journal of Climate, 19, 3952–3972, https://doi.org/10.1175/JCLI3827.1, https://doi.org/10.1175/JCLI3827.1, 2006.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J.: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth system model, J. Adv. Model. Earth Syst., 5, 422–446, https://doi.org/10.1002/jame.20023, 2013.

Kadygrov, N., Maksyutov, S., Eguchi, N., Aoki, T., Nakazawa, T., Yokota, T., and Inoue, G.: Role of simulated GOSAT total column CO2 observations in surface CO2 flux uncertainty reduction, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, https://doi.org/10.1029/2008JD011597, http://dx.doi.org/10.1029/2008JD011597, 2009.

Kaminski, T. and Mathieu, P.-P.: Reviews and syntheses: Flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle, Biogeosciences, 14, 2343–2357, https://doi.org/10.5194/bg-14-2343-2017, http://www.biogeosciences.net/14/2343/2017/, 2017.

Kaminski, T. and Rayner, P. J.: Assisting the Evolution of the Observing System for the Carbon Cycle through Quantitative Network Design, Biogeosciences Discussions, 2017, 1–22, https://doi.org/10.5194/bg-2017-168, http://www.biogeosciences-discuss.net/bg-2017-168/, 2017.

Kaminski, T., Rayner, P., Heimann, M., and Enting, I.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4703, 2001.

Kaminski, T., Scholze, M., and Houweling, S.: Quantifying the Benefit of A-SCOPE Data for Reducing Uncertainties in Terrestrial Carbon Fluxes in CCDAS, Tellus B, https://doi.org/http://dx.doi.org/10.1111/j.1600-0889.2010.00483.x, 2010.

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, Atmospheric Chemistry and Physics, 12, 7867–7879, https://doi.org/10.5194/acp-12-7867-2012, https://www.atmos-chem-phys.net/12/7867/2012/, 2012.

Kaminski, T., Kauker, F., Eicken, H., and Karcher, M.: Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies, The Cryosphere, 9, 1721–1733, https://doi.org/10.5194/tc-9-1721-2015, http://www.the-cryosphere.net/9/1721/2015/, 2015.

Kaminski, T., Kauker, F., Toudal Pedersen, L., Voßbeck, M., Haak, H., Niederdrenk, L., Hendricks, S., Ricker, R., Karcher, M., Eicken, H., and Gråbak, O.: Arctic Mission Benefit Analysis: Impact of Sea Ice Thickness, Freeboard, and Snow Depth Products on Sea Ice Forecast Performance, The Cryosphere Discussions, 2017, 1–39, https://doi.org/10.5194/tc-2017-249, https://www.the-cryosphere-discuss.net/tc-2017-249/, 2017.

Kauker, F., Kaminski, T., Ricker, R., Toudal-Pedersen, L., Dybkjaer, G., Melsheimer, C., Eastwood, S., Sumata, H., Karcher, M., and Gerdes, R.: Seasonal sea ice predictions for the Arctic based on assimilation of remotely sensed observations, The Cryosphere Discussions, 9, 5521–5554, https://doi.org/10.5194/tcd-9-5521-2015, http://www.the-cryosphere-discuss.net/9/5521/2015/, 2015.

Kurtz, N. and Farrell, S.: Large-scale surveys of snow depth on Arctic sea ice from Operation IceBridge, Geophysical Research Letters, 38, 10,462–10,469, https://doi.org/10.1029/2011GL049216, 2011.

Kwok, R. and Cunningham, G. F.: ICESat over Arctic sea ice: Estimation of snow depth and ice thickness, Journal of Geophysical Research: Oceans, 113, n/a–n/a, https://doi.org/10.1029/2008JC004753, http://dx.doi.org/10.1029/2008JC004753, c08010, 2008.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, https://doi.org/10.5194/tc-9-269-2015, https://www.the-cryosphere.net/9/269/2015/, 2015.

Lisaeter, K. A., Evensen, G., and Laxon, S.: Assimilating synthetic CryoSat sea ice thickness in a coupled ice-ocean model, Journal of Geophysical Research: Oceans, 112, n/a–n/a, https://doi.org/10.1029/2006JC003786, http://dx.doi.org/10.1029/2006JC003786, c07023, 2007.

Mikolajewicz, U., Sein, D., Jacob, D., Königk, T., Podzun, R., and Semmler, T.: Simulating Arctic sea ice variability with a coupled regional atmosphere-ocean-sea ice model, Meteorol. Zeitschrift, 14, 793–800, 2015.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, Geophysical Research Letters, 39, n/a–n/a, https://doi.org/10.1029/2012GL053326, http://dx.doi.org/10.1029/2012GL053326, l22707, 2012.

Niederdrenk, A.: The Arctic hydrologic cycle and its variability in a regional coupled climate model, PhD Thesis, Unversity Hamburg, pp. 1–186, 2013.

Niederdrenk, A. L., Sein, D. V., and Mikolajewicz, U.: Interannual variability of the Arctic freshwater cycle in the second half of the twentieth century in a regionally coupled climate model, Climate Dynamics, 47, 3883–3900, https://doi.org/10.1007/s00382-016-3047-1, https://doi.org/10.1007/s00382-016-3047-1, 2016.

NOAA: Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth, NOAA, National Geophysical Data Center, Boulder, Colorado, 1988.

Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., and Marotzke, J.: Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth system model, Journal of Advances in Modeling Earth Systems, 5, 173–194, https://doi.org/10.1002/jame.20016, http://dx.doi.org/10.1002/jame.20016, 2013.

Pacanowski, R. and Philander, S.: Parameterization of vertical mixing in numerical-models of tropical oceans, J. Phys. Oceanogr., 11, 1443–1451, 1981.

Patra, P. K., Maksyutov, S., Sasano, Y., Nakajima, H., Inoue, G., and Nakazawa, T.: An evaluation of CO2 observations with Solar Occultation FTS for Inclined-Orbit Satellite sensor for surface source inversion, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, https://doi.org/10.1029/2003JD003661, http://dx.doi.org/10.1029/2003JD003661, 2003.

Rayner, P., Michalak, A. M., and Chevallier, F.: Fundamentals of Data Assimilation, Geoscientific Model Development Discussions, 2016, 1–21, https://doi.org/10.5194/gmd-2016-148, http://www.geosci-model-dev-discuss.net/gmd-2016-148/, 2016.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Let., 28, 175–178, 2001.

Rayner, P. J., Enting, I. G., and Trudinger, C. M.: Optimizing the CO_2 Observing Network for Constraining Sources and Sinks, Tellus, 48B, 433–444, 1996.

Rayner, P. J., Utembe, S. R., and Crowell, S.: Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study, Atmospheric Measurement Techniques Discussions, 7, 1367–1392, https://doi.org/10.5194/amtd-7-1367-2014, http://www.atmos-meas-tech-discuss.net/7/1367/2014/, 2014.

Redi, M. H.: Oceanic isopycnal mixing by coordinate rotation, J. Phys. Oceanogr., 12, 1154–1158, 1982.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M.: Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, The Cryosphere, 8, 1607–1622, https://doi.org/10.5194/tc-8-1607-2014, 2014.

Röske, F.: An atlas of surface fluxes based on the ECMWF reanalysis—A climatological data set to force global ocean general circulation models, Tech. Rep. 323, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 2001.

Semtner, A.: A Model for the Thermodynamic Growth of Sea Ice in Numerical Investigastions of Climate, Journal pf Physical Oceanography, 6, 379–389, 1976.

Steele, M., Morley, R., and W. Ermold, W.: PHC: A global ocean hydrography with a high quality Arctic Ocean, J. Climate, 14, 2079–2087, 2001.

Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.M. (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, https://doi.org/10.1017/CBO9781107415324, 2013.

Tarantola, A.: Inverse Problem Theory and methods for model parameter estimation, SIAM, Philadelphia, 2005.

Trampert, J. and Snieder, R.: Model Estimations Biased by Truncated Expansions: Possible Artifacts in Seismic Tomography, Science, 271, 1257–1260, 1996.

UNESCO: Algorithms for computation of fundamental properties of seawater, UNESCO Technical Papers in Marine Science, 44, 1983.

Warren, S., Rigor, I., Untersteiner, N., Radionov, V., Bryaz-gin, N., Aleksandrov, Y., and Colony, R.: Snow depth on Arctic sea ice, J. Clim., 12, 1814–1829, 1999.

Xie, J., Counillon, F., Bertino, L., Tian-Kunze, X., and Kaleschke, L.: Benefits of assimilating thin sea ice thickness from SMOS into the TOPAZ system, The Cryosphere, 10, 2745–2761, https://doi.org/10.5194/tc-10-2745-2016, https://www.the-cryosphere.net/10/2745/2016/, 2016.

Yang, Q., Losa, S. N., Losch, M., Tian-Kunze, X., Nerger, L., Liu, J., Kaleschke, L., and Zhang, Z.: Assimilating SMOS sea ice thickness into a coupled ice-ocean model using a local SEIK filter, Journal of Geophysical Research: Oceans, 119, 6680–6692, https://doi.org/10.1002/2014JC009963, http://dx.doi.org/10.1002/2014JC009963, 2014.

Data and Code Availability

Data Sets are available ~~upon~~ upon request to the corresponding author.