Responses to reviewers' comments

We would like to thank the reviewers for their careful readings of our manuscript and the detailed comments. The major changes in the revision are the following:

- Added comparison with the independent sea ice thickness from buoy observations from autonomous ice mass balance (IMB; http://imb.erdc.dren.mil) shown in Fig. 7 of the revised manuscript.
- Equations from (1) to (5) are modified or changed their orders in the text.
- Modifications of figures of 1, 4, and 5 as the suggestions.

The detailed responses are listed one by one as following:

Reviewer #1:

Page 3, line 22: "Measurements of thick sea ice draft. . ." should read "Measurements of thick sea ice freeboard. . ." as altimeters measure freeboard, not draft.

Reply: Thank you, this is now corrected.

Page 5, line 2: the authors state the thickness of TOPAZ was validated over the period 1991-2013 using ICESat and IceBridge. ICESat was from 2003-2009, and IceBridge started after that. What data was used starting in 1991 to validate the model? They re- fer to an unpublished manuscript submitted by Xie (2016) several times. I find this trou- bling, as it is not peer-reviewed and cannot be referred to. It also does not say where is submitted. I suggest a different reference be used throughout the manuscript.

Reply: We have clarifed that. Xie et al. (2016) was submitted to *Ocean Science*, and is available online as doi:10.5194/os-2016-38. In this paper, the reanalysis for the period 1991-2013 is validated with in situ data and satellite data; namely ICESat (2003-2008), IceBridge (2009-2011), and other in situ data for the period 1993-2005 (Lindsay, 2013).

Page 5, line 4, the authors state "While the spatial pattern and regression compare reasonably well, large biases exist" What regression and spatial pattern are they talking about? Are the biases positive or negative? After talking about TOPAZ validation, they state inaccuracy in the ice thickness is a drawback. . .. More detail needs to be added to this section.

Reply: Thanks. This is corrected in the revision.

“In the Arctic reanalysis, the daily sea ice thickness of TOPAZ has been validated for the period 1991-2013 compared to different types of available observations (Xie et al., 2016). TOPAZ shows good agreement with the spatial distribution of ice thickness in ICESat data (available between 2003 and 2008) with a spatial correlation 0.74 in spring and 0.84 in autumn. However, TOPAZ shows a clear overestimation of ice thickness in the Beaufort Sea and an underestimation in the other areas of the Arctic.”

Page 5, line 31: Is the ice model a multi-category model or one layer? This is important because it will come into play when assimilating ice thickness.

Reply: The sea ice model has only one layer (two category, ice or no ice). This is clarified in the model description.

“The NERSC-HYCOM model is coupled to a one thickness category sea-ice model for which the ice thermodynamics are described in Drange and Simonsen (1996) and the ice dynamics are based on the elastic-viscous-plastic rheology described in Hunke and Dukowicz (1997) and with a modification from Bouillon et al. (2013).”

Page 6, equation 3 and lines thereafter: The P^a covariance is described, but not used. I’m not sure what P^a is supposed to be. Also, on line 24 the authors state “the extra term is quadratic and positive.” What term is extra, the second term?

Reply: P^a is the residual error covariance, posterior to the assimilation of data. It is indeed not explicitly required in the paper. We have revised this part of the manuscript to clarify the particularities of the DEnKF.

Page 7, line 8: Another reference to unpublished/unaccepted manuscript for validation. This publication needs to be accepted first, or include the details from that validation in this paper.

Reply: The paper by Xie et al. (2016) is currently under review for ocean science but available in ocean science discussion.

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013, Ocean Sci. Discuss., doi:10.5194/os-2016-38, in review, 2016.

Page 7, line 12: where are the SMOS-Ice products available? A reference or website should be included. Is it available in near real time for operational centers?

Reply: The information is now added.

Page 7 line 9-13: "They are provided by Hamburg University at the website of https://icdc.zmaw.de/1/daten/cryosphere/l3c-smos-sit.html (Kaleschke et al., 2012; Tian-Kunze et al., 2014). SMOS sea ice thickness maps are provided at daily frequency from October 2010 and are available in near-real time during the cold season."

Page 7, line 15: What does "TOPAZ equivalent" mean? Is there some spatial averaging or processing to match the observation location for comparison?

Reply: Sorry this was unclear. The sentence is replaced by "The TOPAZ ice thicknesses shown in Fig.2 are at the same locations and times as the observations."

Page 7, line 17: The term 'RMSD' is used, but not defined until Eq. (6) on page 9. RMSD should be defined here.

Reply: Thanks. The definition of the bias and the RMSD are moved up.

Page 8, line 16: This is the first time the term "innovations" is used. I presume this is referring to the 2nd term on the RHS of equation (1). This should be spelled out.

Reply: Thanks. The definition of "innovations" is added in Section 2.2.

Page 8, equation (4): y_smos is not defined. I guess this is the SMOS observation thickness. Also, is sea ice volume assimilated into the sea ice model? Is there only one ice category in their model to assimilate volume? If more than one category, how do the authors decide what category to assimilate the thickness?

Reply: The definition of y_{smos} is added. Our sea ice model only has one category of ice thickness.

Page 8, line 31. The authors state they implement an upper limit on observation standard deviation of 5 meters. This seem like a large standard deviation value (12.5

times the max observation value of 0.4 m) given they are only assimilating SMOS observa- tions up to 0.4 meters. Why was 5 meters chosen for the standard deviation limit? Are there SMOS observations with 0.5 meters with a 5 meter standard deviation? Figure 2 suggests standard deviations less than 2 for SMOS < 0.4 meters.

Reply: A maximum observation error of 5 meters is set by default by the data provider for saturated values. These measurements are qualitative and cannot be assimilated without algorithmic developments. Other measurements may have uncertainties higher or lower than the observation values but there is nothing in the data assimilation framework that prevents from assimilating them. We show below the uncertainties of the observations as function of the observed thickness of SMOS-Ice in one month of March and November 2014 shown in the Fig. A. In March, the uncertainty is possibly close to 5 meter if the thickness thicker than 0.5 m. Meanwhile, for the thin sea ice (<0.4 m), the related uncertainty may be 10 times the observation value.

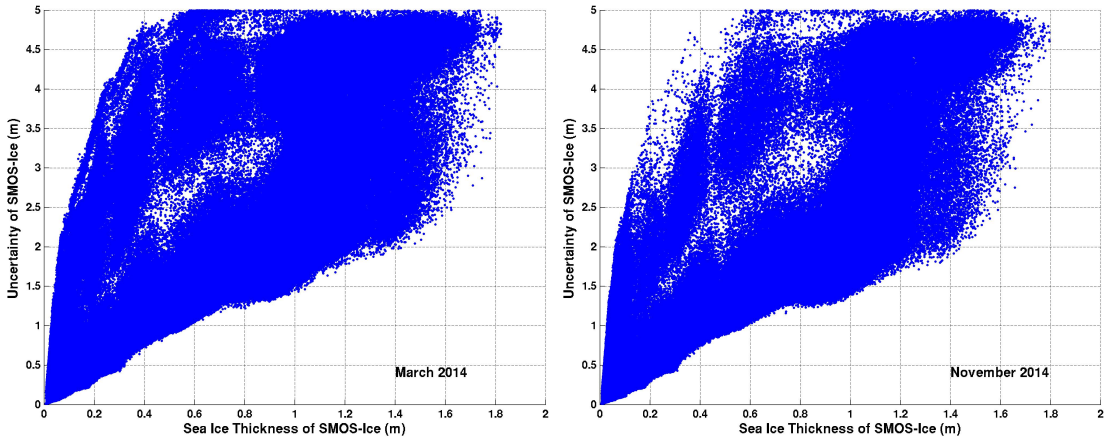

Fig. A The uncertainty of the observation as function of the observed thickness from the SMOS-Ice in March (left) and November (right) of 2014.

The text is modified as such p. 10, l. 9: “with an upper limit of 5 m beyond which the observations are assumed fully saturated”.

Page 8, line 25: why use the symbol TSLA? SLA is used for along-track sea level anomaly on page 4 and 6.

Reply: Thanks. The inconsistencies are corrected, and we only use SLA in the paper.

Page 9, line 10: Sea ice thickness SMOS-Ice is stated to be in Table 1, but it is not listed there.

Reply: Thanks. The statement is changed now.

Page 9, equations 5 and 6: What is 'H'? It is used as Bilinear operator and Obs error in previous equations.

Reply: H is the observations operator, which computes the model equivalent of the observations, it is used for spatial interpolation with a bilinear operator as in equation (1), and the concerned statements are changed.

Page 9, line 20: I find it interesting that there are minimal observations in the Beaufort during March. Can the authors expand on why this is? Is the ice too thick for SMOS at this time? Or not there?

Reply: During March the observed sea-ice thicknesses are mostly thicker than 1 m in the Beaufort Sea, and the thin sea ice (< 0.4 m) appears around the Mackenzie estuary region, shown as the blue shading in the left of Fig. B. It implies the observed thicknesses have been rejected in the OSE runs. In addition, in the right panel of Fig. B the maximal uncertainties (about 5 m) occupy most of the Beaufort Sea, which may relate with the overestimation of the observation uncertainties in the version 2.1 of SMOS-Ice.

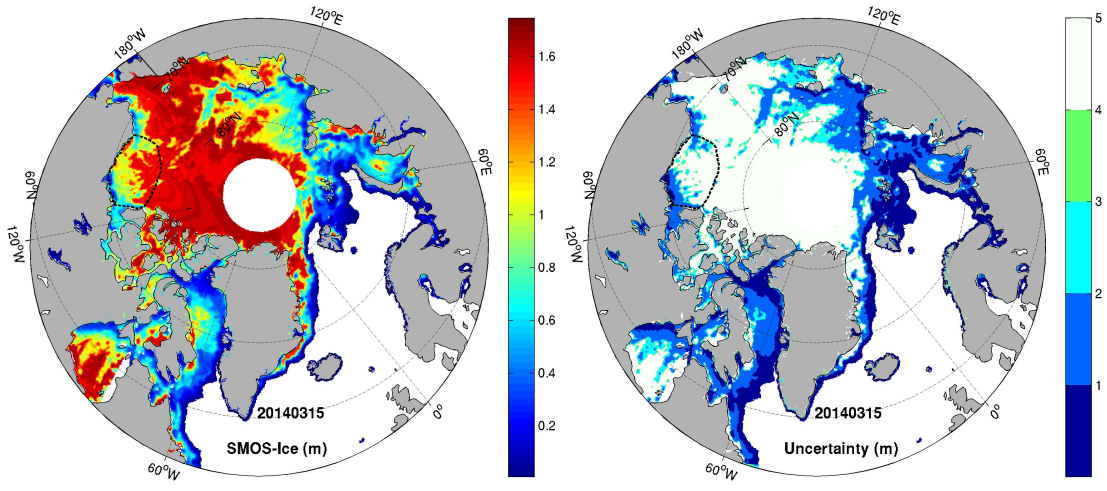

Fig. B Snapshot of Sea ice thickness (left) and its observation uncertainty (right) from the SMOS-Ice data in 15th March 2014. The dashed black line represents the domain of the Beaufort Sea.

Page 10: line 15: The terminology "highlighted with marked lines" is confusing to me. I think it would be clearer if the authors state something like "the averaged thickness of thin sea-ice . . . are shown with marked lines in the panels of Figure 6".

Reply: Thanks. It is replaced in the revision.

Page 11, equation (7): does 'tr' mean trace? This, as well as all terms, should be defined to ensure clarity.

Reply: Yes. The definition is added.

Page 11, line 25: On this line I think the authors are referring to equation (7).

Reply: This sentence was unclear and it has been revised now.

Page 11, equation (8): This looks more like a root mean square (RMS) than a mean. Is there a reason why the authors decided to use a RMS here? Also, what are the subscripts i and j?

Reply: Thanks for this comment. We use the mean DFS to replace the RMS of DFS in the revision, and also update the related figures and statements. In addition, *j* represents the *j*'th type of the observations, *i* accounts the times of data assimilation in the equation (8).

Page 12: equation (9), what is subscript j? j'th observation set?

Reply: Yes, *j* represents the *j*'th type of the observation data-set assimilated in the TOPAZ system. The related statements in the revision are corrected.

Page 12: line 6: Is there a reference for the ice-tethered profile data? Is this ice or ocean profile data?

Reply: Yes, the related reference is added, and they are ocean profile data below the sea ice.

Page 13: line 30: I would say the blended sea ice thickness has been "tested with" the U.S. Navy Arctic Cap Nowcast/Forecast system. The term "implemented" implies the blending is currently being used operationally, which is not the case.

Reply: Thanks. The statement has been modified as recommended, and is changed as following:

"Incidentally, the U.S Navy Arctic Cap Nowcast/Forecast System (ACNFS) is currently testing the assimilation of a combined sea ice thickness product where the sea ice thickness is blended from SMOS-Ice and CryoSat2 based on each satellite retrieval error (personal communication from David Hebert)."

Figure 6: If I understand correctly, the blue line with triangles is the test run where ice thickness is assimilated once a week. There does not seem to be any evidence of the assimilation. I would expect the blue triangle line to get closer to the SMOS line at the assimilation interval. Why does this not occur?

Reply: The blue line is indeed from the test run. The reduction of error is not obvious unless one constructs a weekly cycle of the error but it is not necessarily visible if the error growth is slow. This is not only the case for SMOS but as well for other ocean and sea ice data types (see for example the skills of the operational runs on http://myocean.met.no/ARC-MFC/V2Validation/timeSeriesResults/index.html). The data assimilation updates are however visible as a little flashing if the daily outputs are animated, as can be done using the visualization service on http://marine.copernicus.eu. We are considering positively that the error time series does not bear the scars of assimilation steps as long as the error remains consistently below those of the official run. This means that data assimilation is keeping the test run constantly under control and that the forecast errors are unlikely to grow suddenly after assimilation.

Technical Corrections

In many places throughout the document (e.g., page 2 lines 13 and 18, page 3 line 16, page 4 line 19,) words are run together without proper spacing. I don't know if this is an artifact of the submission formatting or something else, but the word and sentence spacing needs to be verified.

Reply: Thanks. The related formatting is verified.

Page 3, line 29: "Cryostat-2" should be "CryoSat2"

Reply: Thanks. They are modified by CryoSat-2.

Page 4, line 13: Yang reference is (2015), but in bibliography it is 2014. Please check on the year of this publication.

Reply: Thanks. This mistake is corrected.

Page 4, line 14: Define LSEIK.

Reply: Yes, the definition is added.

Page 4, line 16: should read "This study is a follow up and assesses. . ."

Reply: Thanks. “The present study follows up the work from Yang et al. (2014) but uses a different model and assesses.”

Page 4, line 16: What is this a follow-up of? Yang (2015)?

Reply: The rephrasing should have clarified the issue.

Page 4, line 34: should read “. . . and does not apply post processing. . .”

Reply: Thanks. It is modified.

Page 8 and after: no need to bold “official run” and “test run”. I find this distracting in the rest of the document.

Reply: Thanks. They are modified.

Page 12, line 25: need units after 0.4 (should read thinner than 0.4 m)

Reply: Thanks. It is corrected.

Page 12, line 32: Yang reference is (2015), but in bibliography it is 2014. Please check on the year of this publication.

Reply: Thanks. This mistake is corrected.

Figure 1: the words of the regions are hard to see (especially Kara and Beaufort). I suggest putting a white background for these words.

Reply: Thanks for this suggestion. It is modified.

Figure 3: Don’t need the word “resp” when doing comparisons. Usually when compar- isons are done in this manner you just state the contrasting item. The green (red) line represents the mean bias for March (November) of each year.

Reply: Thanks. They are deleted.

Figure 4: Bottom row. It is hard to see the orange line. I suggest choosing a separate color not in the colorbar. Same comment about “resp” as above. Say Top Row, Middle Row, Bottom Row.

Reply: Thanks. The suggestions are taken in the revision.

Figure 5: The vertical axis changes on each plot. Please plot each item with the same vertical axis.

Reply: Thanks. The figure is modified as suggested.

Figures 6, 7, 8: same comment about "resp" and bold test run, official run.

Reply: Thanks. They are modified as suggested.

Reviewer #2:

(1) Page 2, line 6: "winter season" should be "cold season".

Reply: We agree with the reviewer, and also change throughout the manuscript.

(2) Page 2, line 13, the full name of TOPAZ should be given in the Abstract.

Reply: The origin of the name TOPAZ is from a European project ***(Towards) an Operational Prediction system for the North Atlantic European coastal Zones.*** However time has passed and we now consider TOPAZ as a brand name and no longer as an acronym.

(3) Page 2, line 20, should "contents" be corrected to "contains" ?

Reply: Thanks. It is replaced by "contains".

(4) Page 2, line 28, the "Keywords" should be revised, e.g., a lot of readers do not know "OSE" and "DFS".

Reply: These words are replaced by "Observing System Experiment" and "Degrees of Freedom for Signal".

(5) Page 3, line 22, "draft" should be "freeboard".

Reply: Thanks. This mistake is corrected by freeboard.

(6) Page 4, line 13: Yang et al. (2015) should be (2014).

Reply: Thanks. The reference is corrected by Yang et al. (2014).

(7) Page 4, line 14: "LSEIK" should be defined.

Reply: It is defined by with the Localized Singular Evolutive Interpolated Kalman filter (LSEIK, *ref*. Nerger et al., 2005).

(8) Page 4, line 26, "Xie et al., 2016" is frequently referred in this MS, this should be corrected, as it has not been accepted, the authors even have not tell us the journal they submitted to.

Reply: Xie et al. (2016) was submitted to *Ocean Science discussion* at the moment, and is available online as doi:10.5194/os-2016-38. http://www.ocean-sci-discuss.net/os-2016-38/

(9) Page 7, line 15, is "TOPAZ equivalent ice thickness" "TOPAZ model mean ice thick- ness"?

Reply: This statement is modified as "The TOPAZ ice thicknesses shown in Fig.2 are at the same locations and times as the observations."

(10) Page 7, line 17, "RMSD" is not defined here.

Reply: Thanks. This definition is added.

(11) Page 7, line 29, you only assimilate the SMOS data less than 0.40 m, why not 0.50 m? As you referred, "the penetration depth into sea ice is about 0.5 m". Although you mentioned that "the effect of ice melting may lead to a saturation thickness of less than 0.4 m", but for this paper, you run the experiments in the cold season, basically there is no melting in the sea ice surface. If you increase the upper limit, more SMOS observation data is available, thus stronger influence/correction to the TOPAZ system is expected. In Yang et al. (2014), they use an upper limit of 1.0 m.

Reply: It is correct that by raising the threshold to 1 m we would increase the influence of the observation. However the observation error becomes very large above 0.4 m, so we do not expect that we are loosing much information (see also Fig. A in the answer to Reviewer #1). The main motivation for rejecting the observation above this threshold is that there is an obvious bias between model and observation beyond this threshold. Data assimilation with bias is problematic because the correction of the bias may be transferred to other variables via the multivariate updates of the scheme. We have therefore taken a cautious approach and decided not to use the data > 0.4 m for the moment. The word "multivariate" is added on p. 8, l. 8

(12) Page 8, line 3, 4, 7: "thick" should be "thickness"?

Reply: Here, it means the sea-ice thickness simulated by the model is too thick relative to the SMOS-Ice data. We want to keep the indication of the sign of the bias (too thick instead of too thin).

(13) Page 8, line 19, y_{smos} is not defined.

Reply: Thanks. The definition is added in the revision.

(14) Page 9, line 10: "SMOS-Ice" is forgotten in Table 1.

Reply: Table 1 lists the observations assimilated in the present TOPAZ system. This is clarified in the revision.

(15) Page 11, line 1: In the Beaufort Sea, there are some sea ice draft measure- ments from Beaufort Gyre Exploration Project (BGEP) by upward-looking sonar (ULS) moorings located in the Beaufort Sea (http://www.whoi.edu/beaufortgyre). Also, there are some sea ice thickness data obtained from autonomous ice mass balance (IMB; http://imb.erdc.dren.mil). I would suggest the authors to use these data as the inde- pendent ice thickness observations in the evaluation of their model results.

Reply: Thanks. The two buoys from the IMB have been used to validate the sea ice thickness as the Fig. 7 in the revision. As the buoys are far away from assimilated observation, the impact is small. Still there is a slight improvement.

(16) Page 12, line 32, an "a" is missing before "slight". (17) Page 13, line 24, should be "In addition".

Reply: Thanks. It is added in the revision.