Ocean Sci. Discuss., 9, C736–C750, 2012
www.ocean-sci-discuss.net/9/C736/2012/

Ocean Science Discussions

Interactive comment on "TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic" *by* P. Sakov et al.

P. Sakov et al.

pavel.sakov@nersc.no

Received and published: 17 July 2012

Response to interactive comment C471 by Anonymous Referee #2 to manuscript OS-2012-33 "TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic" by P. Sakov et al.

General Comments

This study aims to present a description and evaluation of the TOPAZ4 operational sea ice-ocean assimilation system for the Arctic and North Atlantic. However no details of how the system is run operationally are given with no demonstration of the skill of the operational system.

REPLY: We mention that the system is run operationally at met.no. We do not be-

lieve than more details on that are necessary. The skills of the operational forecasts are regularly updated on the MyOcean web page: http://myocean.met.no/ARC-MFC/V2Validation/index.html. We also stress that the innovation statistics shown is obtained for the forecast, before any assimilation of observations is conducted, and therefore demonstrates the forecasting skill of the system.

Rather, the authors use a multi-year reanalysis to evaluate the quality of the analysis system. However, they undermine the usefulness of the reanalysis by making numerous changes to the data assimilation methodology, in addition to the observations assimilated, throughout the reanalysis. Thus, making it neither straightforward to judge the quality of the analysis system nor the reanalysis itself.

REPLY: The paper discusses the "pilot" reanalysis. The purpose of the pilot reanalysis is to serve as a testbed for the main reanalysis. In our view, the pilot reanalysis provides comprehensive material for assessing the quality of the system. In comparison, an operational forecast system would have undergone many more changes in the same time, making its evaluation even more complicated.

The authors claim that this is the only "operational large-scale ocean data assimilation system that uses the ensemble Kalman filter". Does this not then merit a clear evaluation of the quality of the operational analyses? While I agree that reanalysis is a useful technique to evaluate an analysis system, a reanalysis should maintain the same analysis system throughout.

REPLY: We can not see how several changes in the system setup over the course of 6 years do preclude us from assessing the system. Even one year of reanalysis tells a lot to a data assimilation specialist about the quality of the system...

If the aim is to evaluate the impact of modifications to the analysis system this should be done in parallel experiments to show clearly the impact. Doing so progressively throughout the reanalysis makes it difficult to separate the impacts of the modifications from interannual variability, temporal inhomogeneity of observations, etc. While I am sensitive to the high numerical cost of running these systems this should not be used as an excuse for a lack of rigor.

REPLY: The high numerical costs, but also the limited duration of the MyOcean project and the limited human resources indeed prevented us from conducting a strict experimental plan, which we agree would have been more convincing. We believe however that the time allowed between each of the changes is long enough to conclude with confidence, based on the results presented in this paper whether each change was positive or not.

In addition to my concerns regarding the evaluation methodology noted above, I have a number of specific issues regarding the metrics provided. In particular, a number of basic evaluations are missing (see below), including spatial errors for the sea ice cover, which is presumably one of the main outputs of this system.

REPLY: See our reply to p. 22.

Also, the system description is only given with respect to the TOPAZ3 system presented in Bertino and Lisaeter (2008). However, Bertino and Lisaeter (2008) do not provide a detailed description of the system, nor do they present an evaluation of the system performance. Thus, in so far as I could discern, no clear evaluation of the TOPAZ analysis system has yet to be published in peer-reviewed literature.

REPLY: Section 2 provides the description of the mode; Section 3 provides description of the data assimilation. We believe that both these descriptions are rather comprehensive.

As such, I recommend that the authors retract the paper and consider rewriting it using a new set of experiments to clearly show the system performance and improvements in TOPAZ4. Specific experiments to show the benefits of the ensemble covariances and bias correction scheme would be of particular value.

REPLY: We believe that this paper shows the performance and improvements in

TOPAZ4. The benefits of ensemble covariances have already been exhibited in Lisaeter et al. 2003, Brusdal et al. 2003. These references are cited in the manuscript.

Specific Comments

1. Pg 1520, line 7: No demonstration of the spatial features of ocean circulation and sea ice cover are provided. Thus this claim is not supported.

REPLY: The quality of ocean circulation is supported by results on: - SLA and SST forecast innovation statistics (Fig. 5); - comparison of model SLA with drifter trajectories (Fig. 11); - comparison of mean monthly salinity in the Arctic at 0 m and 100 m with climatology (Fig. 13, 14); - comparison of temperature in the Fram Strait section with observations (Fig. 16); - comparison of the mean transport through Svinoy and Barents Sea Opening section with observations (Fig. 17); - indirectly - by the consistency of the bias estimates for MSSH and SST (Fig. 8-10). The quality of the sea ice is demonstrated by: - the forecast innovation statistics (Fig. 7); - comparison of ice thickness with observations (Fig. 12). An additional figure showing the agreement with the observed ice edge is now included.

2. Pg 1521, line 17: If this is the Arctic MFC shouldn't scales appropriate for the Arctic be mentioned?

REPLY: It is mentioned at page 1524, line 8.

3. Pg 1522, line 4: While I agree that time dependent state error covariances may help to better capture features along the ice edge, they are far from "essential". If this is an aim of the paper there should be a demonstration of the systematic impact of these covariances near the ice edge. Indeed, only a single example is given in the paper.

REPLY: We put up what we see as a typical and revealing example. We could easily expand this line of argument, but believe that the example in the paper is sufficient to make a valid point to a data assimilation specialist. A more dedicated discussion is given in Lisaeter et al. 2003 (cited).

4. Pg 1522, lines 6-20: This description is not clear. Figure 2 needs to be complemented by maps of the background state to show how the system is affected.

REPLY: We believe that this is not necessary. The position of the ice edge is the only feature from the background state worth showing for this example.

5. Pg 1524, line 28: The reference should be that of Dee et al. QJRMS, 2011. The ERA-Interim fields are provided on a 0.25 grid, however, the model used is on a T255 grid, giving roughly 79km resolution (see Dee et al, 2011).

REPLY: We do not discuss the ERA-interim reanalysis, and to avoid focusing on these details we removed the description of its resolution. We replaced the reference to Simmons et al (2007) by reference to Dee et al (2011).

6. Pg 1525, line 3-4: Provide details of how incoming radiation is calculated.

REPLY: The details occupy 8 pages in the report by Drange and Simonsen (1996), already cited for sea ice thermodynamics, the citation has been repeated there (also requested by Reviewer 1).

7. Pg 1526, section 3.1: This section is difficult to read as it only provides details concerning differences with respect to TOPAZ3 without any basic description of the system itself. To the best of my knowledge, no detailed description of the TOPAZ analysis system exists in the peer-reviewed literature. As such one should be provided here. This is especially important considering that TOPAZ is the only operational iceocean EnKF.

REPLY: It is not true that this section "only provides details concerning differences with respect to TOPAZ3". It provides an adequate description of the system, if indeed one gets familiar with such schemes and concepts as DEnKF, local analysis, and asynchronous data assimilation by looking into references.

8. Pg 1527, line24: Why was a factor of 2 chosen? Would 1.5 suffice? Again, given that this is the only operation ice-ocean EnKF it is important to justify clearly and demonstrate the impact of these choices.

REPLY: There are many factors like this in data assimilation. Why a localisation radius of 300 km was chosen? Will 500 km do better? Why the ensemble size of 100 was chosen? Could one get away with 50? Why the assimilation cycle of 1 week was chosen? And so on.

A trivial scalar case in a standard situation (obs. errors between 2 and 4 times larger than forecast errors) shows that the parameter only has a mild effect on the system and always on the safe side, as argued in the manuscript, as the parameter keeps the ensemble spread somewhat higher. We did not add more justifications in the paper since we do not expect the reader to be interested in them. Similarly, we have presented no scientific justification why the assimilation cycle is set to 1 week.

9. Pg 1528, line 25: How were these values chosen? How sensitive is the system to their values? How does this impact on the ice-ocean covariances? Some demonstration of the sensitivity to these choices should be given.

REPLY: The values were chosen based on the expected statistical properties of errors in near-surface atmospheric fields and have been used in previous studies as explained in the text. Sensitivity tests to these values are of limited interest at a time when atmospheric models are commonly running ensembles, so a specific study should evaluate the value of ensemble atmospheric input against random perturbations with semi-arbitrary statistical parameters. We believe that this would belong to a different paper, however. Note that Lisaeter et al. (2007) evaluate different perturbation systems in the context of assimilation of ice thickness data, which are not assimilated here. Our perturbation system is slightly more inclusive than the one recommended there.

@article{lis07, author = {Knut Arild Lis{\ae}ter and Geir Evensen and Seymour Laxon}, title = {"Assimilating synthetic CryoSat sea ice thickness in a coupled ice-ocean model"}, journal = JGR, year = {2007}, number = 112, pages = "C07023" }

10. Pg 1530, line 17: How are DFS and SRF useful? Fields are presented in Fig. 3 with no discussion of the values or how this is useful to demonstrate the quality of the

system. Moreover, if they are so useful why simply show an example rather than a timeseries or an average map, etc.

REPLY: If for example something went wrong in a given assimilation cycle, then examining DFS and SRF maps for this cycle provides information that helps to differentiate between scenarios when the system setup is incorrect, or when there is a problem with observations.

We added the following text: "DFS is a good indicator of potential rank problems. Ideally, one would like to keep it below about 20 for the ensemble size of 100; while values of around 50 would point at too small ensemble or too big localisation radius. SRF characterises the "strength" of data assimilation. “Strong” data assimilation implies a high degree of optimality of the system and should be avoided. Ideally, the magnitude of SRF should not much exceed 1. If SRF is consistently higher than that then perhaps a shorter cycle is needed to limit the growth of unstable modes."

11. Pg 1532, line 18: Why 0.7? How do the resulting representivity error values compare to the observed variances?

REPLY: This is an empirical value tuned to the system performance. For example, the value of 0.5 results in substantially more analysis artifacts.

12. Pg 1533, line 1: What is the value of the observation error?

REPLY: For OSTIA SST it is around 0.3 K (for a product with 1/20-degree resolution). We basically keep this order of error, but for super observations.

13. Pg 1534, line 23: Ice drift has been assimilation by Stark et al. JGR, 2008

REPLY: The paper by Stark et al. is indeed the first demonstrating assimilation of ice drift, however, contrarily to our approach, the Lagrangian vectors are considered as Eulerian measurements and the demonstration is only conducted on a limited time period. Stark's method assumed that the only source of errors on ice drift was stemming from the surface winds, which possibly lead to unrealistic corrections of surface winds

in areas where the ocean currents or sea ice rheology influence the drift. Also note that Stark's method has not been taken up in the FOAM system. After the question 26 we removed the claim on ice drift assimilation.

14. Pg 1535, line 25: If the bias estimation system doesn't correct biases then won't this impact on the quality of the analysis at the ice edge where strong non-linearities around the background state are present? For example, if the ice extent is systematically underrepresented (which appears to be the case for TOPAZ4 in summer) the error covariances will also highlight uncertainty (as in Fig 2) in the wrong location. A clear demonstration of the impact of the bias correction needs to be provided.

REPLY: Firstly, the lower mean ice concentration in summer does not necessarily transfer into smaller ice extent. (As our ice model tends to ice concentrations about 0 and 1, rather than, say, about 0.8.) Secondly, the system could not perform robustly when the ensemble spread in the ice edge did not embrace the observed ice edge position. We made a considerable effort to tune the perturbation system to achieve that, so that this does not happen systematically. Thirdly, since there are no SST or SLA observations for ice covered areas, the bias estimates do not affect ICEC. Therefore, we conclude that bias correction for SST and SLA has little (if any) impact on ICEC. This can be confirmed by examining the innovation statistics for ICEC in Fig. 7, which shows little change after switching on the bias correction.

15. Pg 1537, line 6: The agreement can hardly be considered "good" when the innovation standard deviation is only 50% of the RMSD.

REPLY: Ideally, the green line (observed innovation RMSD) should match the red dotted line (estimated innovation RMSD). There is a perfect match in the Tropical box, and acceptable matches in other boxes, with the worst match in the Gulf Stream box (about 15 cm innovation RMSD, 8-9 cm estimated RMSD). It is on a low side, but from our experience this is rather "good" than "bad". We would be interested to learn if the Reviewer knows examples of better match.

16. Pg 1537, line 22: It would be much more insightful to have run the assimilation of both sst products in parallel to isolate the impact of this change.

REPLY: Yes, it would. In fact, there are a number of similarly important parameters.

17. Pg 1538, line 5-10: This argument is not clear. If there is a seasonality in the mixed layer depth, why not evaluate this against Argo-based mixed layer depth products rather than just speculate?

REPLY: The seasonality of the mixed layer depths should be well known to all readers of the MyOcean special issue, and the paper would not benefit from additional plots. The "speculation" is just a plain common sense in our view.

18. Pg 1538, line 12: Section 3.2 does not indicate the date this starts. Why not start this at the beginning of the reanalysis?

REPLY: Because we observed a gradual spread reduction in SLA, particularly in the Gulf Stream and Gulf Stream Extension boxes, and particularly so at that moment of time. We had no idea how the ensemble spread would behave at the start of the reanalysis.

19. Pg 1538, line 15: I can't see any impact. Clarify.

REPLY: One can see a substantial increase in the ensemble spread in the Gulf Stream Extension and Gulf Stream boxes just after the vertical dashed line in Figure 5, and some increase in Nordic Seas and Gulf Stream extension boxes in Figure 6. We can see no more systematic reduction of spread after that.

20. Pg 1538, line 18: The previous paragraph commented on the changes in ensemble spread and here it says the spread is "relatively constant".

REPLY: Relatively constant, that is what it is. No ensemble collapse.

21. Pg 1538, line 19: I don't agree that there's "no tendency" towards ensemble collapse. Prior to the red vertical line in Fig 5 and 6 there seems to be a negative trend.

REPLY: The Reviewer could not "see any impact" in comment 19, but now sees a "negative trend" "prior to the red vertical line". Anyway, we indeed were tuning the system during the pilot reanalysis, and are quite happy now how it behaves in regard to the ensemble spread. The full reanalysis (ongoing so far for 17 years) confirms that the tuning was successful.

22. Pg 1539, line 9: Where is the bias? Given the biases noted in Fig 5,6 and 7, an indication of where the biases are located should be given, especially for sea ice. Given that this analysis system is used for the Arctic MFS in MyOcean I would have expected a focus on evaluating the sea ice cover in the Arctic. Moreover, Fig. 7 suggests some important seasonal biases in sea ice cover. A systematic demonstration of where and how these occur should be included. This is especially relevant given the assertions made in the introduction that the ensemble covariances are "essential" for coupled ice-ocean data assimilation.

REPLY: We have not initially included the results on the ice concentration in the paper because we judged that these results, although rather good, could be judged by a naive person as unimportant, because ice concentration is an assimilated variable. We have now added the results on ice concentration to Section 5.4 "Evaluation of ice fields".

Concerning the bias, it is not huge (generally < 0.1), and persists over a rather short period of time, except summer 2005. We have not investigated the spatial distribution of the bias in the pilot reanalysis, mainly because the results in the main reanalysis we switched to using ICEC product from met.no, and the results are considerably better (the RMSE is about 0.1, and bias under 0.05).

23. Pg 1540, line 11. This looks more like tightening of the Gulf Stream rather than shifting.

REPLY: A careful look at Fig. 9 (right) indicates both a tightening and a shifting, correcting the slope only would leave the blue line shifted to the North.

24. Pg 1541, line 27: This argument is misleading given that the rossby radius is smaller in the Arctic, and thus the system is not eddy-permitting in the Arctic even if it is for the Gulf Stream.

REPLY: The point of the model not being eddy-permitting at high latitude was already made clear p. 1524 l.5, and the statement p. 1541 l. 25 clearly concerns the Gulf Stream only. So the statement l. 27 cannot be mistaken as "the model resolves eddies in the Arctic". However, the statement in the conclusion was indeed misleading and has been modified (see p. 32).

25. Pg 1542, line 14: These results are not systematically related to EVP. There are many examples of EVP models able to produce thick ice along north of Greenland and the Canadian Archipelago.

REPLY: The EVP models generally simulate thicker ice North of Greenland and the Canadian Archipelago, but the thick ice remains too thin and thin ice too thick regardless of the forcing fields or thermodynamics used (see Johnson et al. JGR 2012 for all 6 AOMIP models). This points towards a signature from the viscous model, used by all AOMIP models, that does not fracture the thick multi-year ice as much as observed (see Kwok and Cunningham, GRL 2010), thus underestimating the lateral melting.

The sentence was modified as follows: "The ice is too thin in areas of thick ice and inversely, too thin in areas of thick ice, which is a common feature in models that use a viscous rheology \citep{joh12}."

@article{joh12, author = {Mark Johnson and Andrey Proshutinsky and Yevgeny Aksenov and An Nguyen and Ron Lindsay and Christian Haas and Jinlun Zhang and Nicolay Diansky and Ron Kwok and Wieslaw Mas\'lowski and Sirpa H\ÂÍakkinen and Igor Ashik and Beverly de Cuevas}, title = {{Evaluation of Arctic sea ice thickness simulated by Arctic Ocean Model Intercomparison Project models}}, journal = JGR, year = {2012}, volume = {117}, number = {C00D13}, pages = {1-21} }

@article{kwo10, author = {Ron Kwok and Cunningham, G. F.}, title = {{Contribution of melt in the Beaufort Sea to the decline in Arctic multiyear sea ice coverage: 1993-2009}}, journal = GRL, year = {2010}, volume = {37}, number = {L20501}, pages = {1-5} }

26. Pg 1542, line 22: Why are these diagnostics not shown? Fig. 11 shows drifters in the Gulf Stream and the text states that the Arctic Ocean is the main focus (Pg 1541, line 27). Moreover, the text states that this is the first system to assimilate ice drift. Should it not then demonstrate the result of this?

REPLY: Please find the results shown in Fig. 1 (courtesy of Denis Demchev, NIERSC, Ru). They have not been included by lack of space. The assimilation of CERSAT ice drift has a very small effect when one assumes an observation error of one pixel (DFS inferior to 0.1). So the impact of its assimilation is not obvious from these results and we understand that the claims of novelty need to be removed short of being supported by the results. A dedicated impact study should be undertaken when time and personnel can be made available for that purpose.

The passage "To the best of our knowledge, assimilation of ice drift in TOPAZ represents the first example of assimilating Lagrangian data in a realistic ocean model." p. 1534, l. 22 has been removed.

27. Pg 1542, Sec 5.4: What about the spatial distribution of sea ice concentration errors? Given that the Arctic is the focus of the system and the claims made regarding the need for ensemble error covariances a systematic demonstration of effects along the ice edge needs to be presented.

REPLY: We have added in the Fig. 12 (revised version) the comparison of ice concentrations to observations. One may note in particular that the ice edge fits quite well, except in Summers 2005 and 2006 as commented in the text. Also note how the sea ice edge recovers in the following years after the tuning described in the text.

28. Pg 1543, Sec 5.5: Again, given the Arctic focus a more detailed analysis of the representation of temperature and salinity in the Arctic needs to be included. With the International Polar Year, there are numerous in situ datasets to use, in addition to more basic evaluations of model drift and how assimilation corrects this (or not).

REPLY: The IPY temperature and salinity profiles have been assimilated, as mentioned in the article. The impact on salinity is shown in Fig 13 as the difference between 2007 and 2008 and compared to climatology to show that the assimilation has corrected a model bias. The same comparison with Temperature shows less visible differences but brings the same message. There was unfortunately too little time between the start of the IPY and the end of our reanalysis to evaluate any trends, but this will be done with the ongoing reanalysis (1991-2010).

29. Pg 1543, line 20: Showing a single year comparison is hardly sufficient given the large number of changes applied to the system during the reanalysis.

REPLY: These findings were confirmed in the course of the main (1991-2010) reanalysis.

30. Pg 1545, line 16: Given the initialization problem how well can the data assimilation constrain the system? This in itself is an interesting question that could have been addressed here with parallel experiments.

REPLY: In its current state, DA in TOPAZ makes a decent job in regard to the ice concentration and ice extent. It improves salinity and temperature, but is far from constraining these fields in the Arctic, where they currently mainly depend on the quality of the model. The assimilation is still able to constrain well the sea ice cover because the biases at depths are isolated from the ice by a cold halocline. There is no need for parallel experiments to understand that point.

31. Pg 1548, line 13: As noted above, this claim is unsupported by the demonstration provided. Only one example of the effect is given with no systematic evaluation of the

impact over time or a comparison with an analysis produced without it.

REPLY: See our reply to p. 3.

32. Pg 1550, line 28: The claim regarding the circulation in the Arctic should be revised as nothing is shown of the circulation in the Arctic itself and Fig. 15 suggests the system has difficulty even reproducing climatology.

REPLY: We agree that the phrasing was misleading here. The sentence has been changed as follows: "... we demonstrate that TOPAZ4 produces a realistic representation of the mesoscale ocean circulation in the North Atlantic, and a realistic representation of sea ice variability in the Arctic".

33. Pg 1550, line 29: What does "almost similar" mean?

REPLY: The real-time system assimilates real-time input (observations and atmospheric forcing) instead of delayed-mode data, it has also followed a longer spin-up and benefits from improvements that have been made available after the pilot reanalysis has been completed (a 4th order numerical scheme for advection of momentum). It is otherwise using exactly the same source code. We assume the reader does not need to know these details.

Technical Corrections

1. Pg 1531, line 12: should read "spun up"

REPLY: Thank you.

Interactive comment on Ocean Sci. Discuss., 9, 1519, 2012.

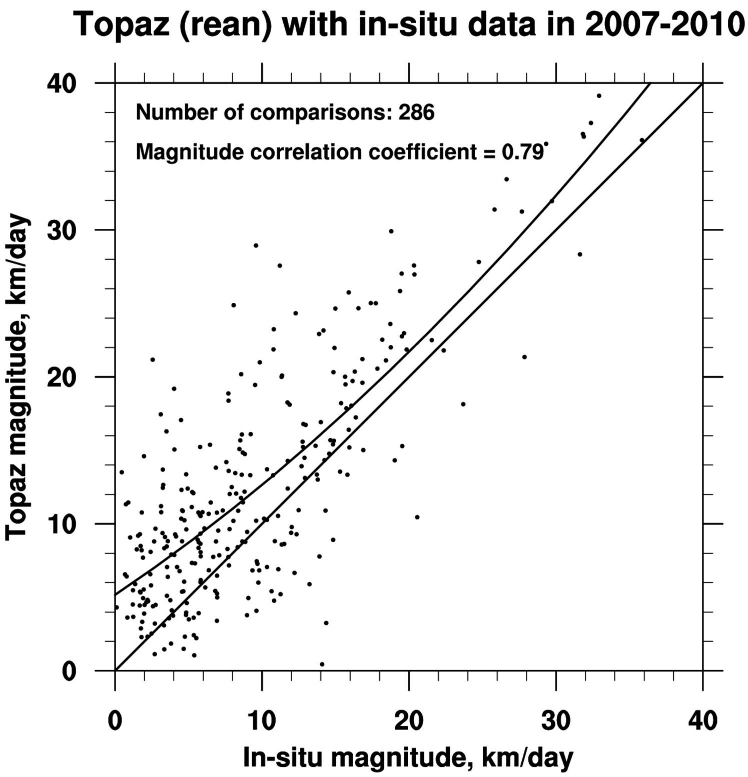

Fig. 1. Comparison of the ice drift in TOPAZ with the observed ice drift