Geosci. Model Dev. Discuss., 6, C2984–C2991, 2014
www.geosci-model-dev-discuss.net/6/C2984/2014/

Geoscientific Model Development Discussions

Interactive comment on "Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts" *by* E. W. Blockley et al.

E. W. Blockley et al.

ed.blockley@metoffice.gov.uk

Received and published: 8 April 2014

The authors would like to thank Reviewer #2 for their review of this work and for taking the time to read through the paper in such detail.
Below is a response to the review whilst comments and the changes to the manuscript can be found in a GMD style paper uploaded as supplimentary material.

Use of 48-hour assimilation window
There seems to be some confusion relating to the 48-hour window that is used operationally and the 24-hour daily cycling used in the reanalyses. Given that Reviewer

#1 also had some questions relating to this it is evident that the explanation provided needs to be made clearer.

When we run FOAM operationally in near-real-time we start from T-48h and produce our 'best estimate' analysis for the (T-48h,T-24h] period. The model state at T-24h is then saved for the next day as our best estimate of the ocean state. An update run is then performed for the period (T-24h,T+00h] before we start the 7-day forecast. The impact of running a second analysis period therefore is to allow more observations to be used in the generation of out best-guess initial conditions each day owing to the late arrival of some observations. As mentioned in Section 2.4.2 (now Appendix A) this led to a RMS error reduction of approx 5-6% globally in the operational near-real-time system when it was implemented. To make this clearer in the paper I have modified the explanation of the 48-hour window in Section 2.3 to better distinguish between the 'best estimate' analysis and the 'update run'.

We do not use this approach for the reanalyses because they are run in delayed time (typically > 6 months after real-time) and so have access to more observations than the operational system because late arrivals will be included. These reanalyses are not an identical copy of the operational suite and are used in a calibration sense to understand the potential errors in the operational system.

I have therefore expanded the last sentence at the end of the 1st paragraph of Section 3 to explain that we do not need to run 2 separate assimilation cycles in the reanalyses because the observations are extracted in delayed-time and are therefore more plentiful than those used operationally.

Re-structure of Section 2
Details of the v10 -> v11 upgrade have now been moved into an appendix (Appendix

A) and Section 2.4 has been simplified in this regard.
However we do not feel that the NEMO/NEMOVAR details in Section 2.1 and 2.2 should be moved to appendices because we think, as this is essentially a system description, these are an important focus of the paper.

Collins et al. reference

I don't think Collins et al. 2006 support the affirmation that "the Atlantic meridional overturning circulation at 26.5 N" is "important for the initialisation of the coupled seasonal forecasts". Its focus is on interannual to decadal forecasts. Yes you are correct this is more aimed at inter-annual to decadal.
I was trying to say 2 things here; 1. that the improved representation of the mesoscale is beneficial when we use FOAM to initialise our coupled seasonal forecasts and 2. that an improved representation of the AMOC leads to improvements on longer timescales.
This sentence has been modified and the Collins et al. reference replaced with 2 additional references: Barnston et al., (2012) who discuss the importance of an improved initialisation for seasonal forecasts and Cunningham et al. (2013) who present some recent observational results demonstrating the potential importance of the AMOC in controlling sub-surface temperature anomalies in the sub-tropical Atlantic.

Sea ice

Could the authors be more specific? What are the inconstancies implied by LIM2?
The inconsistencies implied by using LIM2 are simply that it is not the same model as is used in the other Met Office forecasting systems (seasonal, decadal and climate) which use the CICE sea ice model with 5 thickness categories. The goal of the Met Office is to develop a consistent, seamless approach to forecasting across all time scales which is described further by way of the addition of Brown et al., (2012) to the

references. Aside from the seamless agenda, consistency is particularly important for the GloSea5 seasonal forecasting system which is initialised using FOAM ice analyses each day. We therefore require FOAM and GloSea to be as consistent as possible to reduce the chance of coupled initialisation shock.

It is not clear whether there is some balance relationship between sea-ice concentration and the other state variables (none of that is in Weaver et al, 2005). Could the authors clarify this?
Sea ice concentration has been implemented as an unbalanced variable in the linearised balance relationships and so is not balanced with respect to the other state variables. This is described in Walters et al. (2014) and I have added a sentence to Section 2.2 to make this clearer.

Is there a constraint that the ice thickness is positive within the assimilation scheme, or is this ensured by the model?
Although the model would prevent this happening the assimilation does not actually make any reduction to the category mean ice thickness and so this is not actually possible.

Persistent warm bias at 100m

The formulation "NEMOVAR fails to fully constrain a persistent model bias" is a bit specious... this implicitly says that OCNASM succeeded reducing it in v11. What is the bias of the v11 equivalent free run?
This is a persistent bias because, as well as being present in the free-running NEMO model at v12, it is also apparent in the v11 free run (not shown). The difference between v12 and v11 is that the NEMOVAR assimilation has not managed to constrain this bias as effectively as the old OCNASM system.

This issue is related to the difficulties associated with fitting the relatively sparse sub-surface observations with the short correlation length-scales employed by NEMOVAR. It is hoped that the adoption of dual length-scales will improve things in the ocean interior and development of this is underway.
This issue is mentioned at the end of Section 4.1.3 and in the Summary.

NEMOVAR constraining mesoscale eddies better
It is claimed several times in the paper that the NEMOVAR assimilation scheme is more suitable for constraining the eddy variability, but it is not fully demonstrated in the paper in my opinion.
The evidence for these statements is based upon the gridded data comparisons in Section 4.3 supported by the findings of Waters et al. (2014).

However we are inclined to agree with you when you say you say that you don't feel that this has been fully demonstrated. The approach that you suggest seems very sensible and so we have adopted this to add some clarity to this issue. To do this the extra-tropical ocean (between 23-66 latitude) was partitioned into high and low variability regimes based on the standard deviation of SLA observations for the full 2-year assessment period. Statistics for SLA and near-surface currents were then calculated for each of these high/low variability regimes for both the v11 and v12 systems. We then calculated the relative improvement (as a percentage reduction in RMS error) of v12 over v11 for each of variables and regimes.

Results show that the improvement in SLA for v12 is a factor of 10 higher in high variability regimes than for low variability regimes. The same is true for near-surface velocities which show a 2-fold reduction in RMS error in high variability compared with low variability regions.

These percentage improvements can be found in the new Table 3 and are the process is introduced and discussed in Section 4.1.4

This section does not present forecast statistics of sea level anomalies. I wonder why, because sea level is a useful indicator of the upper ocean dynamics.
Sea level anomaly forecast statistics have not been calculated because, owing to a problem with our archiving, we do not have all the data required to perform the validation. Unfortunately the altimeter bias file used each day to calculate the SSH from the SLA was not archived correctly.
The only way to generate these statistics would be to re-run the trials in their entirety, which is not possible at present. This issue has been corrected and so these results will be included in future work.

The problem of overfitting has already been addressed within the NEMOVAR system in Daget et al., 2009. They propose a diagnostic that could be used here.
The error variances will be recalculated as part of dual length-scale changes and this will be a useful metric to diagnose over-fitting as part of this change.

Remark about Figure 7b and d: the forecast lead-time information is not obvious to understand. I suggest a better caption text, or a different representation.
The caption of Fig. 7 has been modified to include more explanation of the forecast profiles in Fig. 7b, d.

Increments in the tropics
I don't understand the argument about increments in the tropics. It is said line 17 of the same page that tests "including the use of a second-order velocity balance" are under way. This implies that this balance is not applied in the present system, and therefore

that the velocity increments should be zero in the tropicalregions. If this is right, how could velocity increments indicate anything as said line 9-10?
The balance in the present system is geostrophic and so velocity increments will be effectively zero near the equator. We are only talking about tracer increments here and so in the text I have replaced 'increments' with 'tracer increments' to make this clearer.

Figure font sizes
Reviewer #1 also commented on the relatively small font size used for the figures and so we acknowledge that this will need to be increased.

The font sizes used in the figures are either 18 or 20pt but it is the figure scaling employed by the typesetting that causes them to be smaller than the journal text. Ideally these should align with the font-size used for the figure caption. When the final typesetting is done we shall make sure, in conjunction with the journal, that the fonts used in the figures are clear and in keeping with GMD guidelines.

We have not done so as part of this response because the figure scaling used for this (single column) GMDD discussion document will most likely be different from that used in the (dual column) final version - meaning that any changes we made now may very well need to be redone at the typesetting stage.

The following typos have also been corrected:
p6239, line 5: typo “salinity”
p6254, line 11: typo “will be upgraded”
p6255, line 24: typo “ocean”

Please also note the supplement to this comment:
http://www.geosci-model-dev-discuss.net/6/C2984/2014/gmdd-6-C2984-2014-supplement.pdf

Interactive comment on Geosci. Model Dev. Discuss., 6, 6219, 2013.