Geosci. Model Dev. Discuss., 6, C2363–C2369, 2014
www.geosci-model-dev-discuss.net/6/C2363/2014/


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.

**G. B. Brassington (Referee)**

g.brassington@bom.gov.au

Received and published: 15 January 2014

The article is publishable as a presentation of the performance of the FOAM v12 system which is operational at the UK Met Office and its progress compared with the previous system. As a general statement verification of the performance of ocean analyses and forecasts is challenged by the lack of independent observations. The FOAM v12 results are mixed with some clear challenges for the forecasts of T/S and some encouraging results for surface currents.

The presentation style does include several vague statements that could be better supported by an objective measure however it is in the main logically structured and



readable. Several figures have very small fontsizes and could be improved.

Several suggestions are made to improve or clarify the presentation of results generally of a minor nature. If these can be addressed I would recommend this article for publication.

Two caveats:

Note that Water et al., 2013 (submitted) was not reviewed. Key details would be relevant to interpretation of these results.

I am not an expert on sea-ice and have deferred from a review of these sections. Just read with interest.

Abstract

a. The FOAM system is described in the abstract and early section as inclusive of multiple models of varying resolutions. However the validation only focuses on the global system. In some cases this is a disadvantage as some of the conclusions suggest that eddy-permitting resolution might be responsible for some biases. As the focus is on the performance of the global system I suggest removing mention of the other systems from the abstract and retain in the introduction.

b. The v12 system includes many significant changes. The control experiment chosen was to perform a free running model. It might have been more instructive to perform a data assimilative run with fewer of the changes to provide some comparison on what changes might be responsible for the positive and negative results.

Introduction

It is worthwhile noting the negative impact of adopting the bulk formulae in terms of the mis-match in SST boundary conditions between the ocean model and the NWP. Particularly a system that does not beat persistence.

Section 2.1

the instability is a numerical one. Explicit diffusion schemes must satisfy criterion of the form K * delta t / delta xˆ2 < 1. Therefore the adjustment in K would need to scale with the square in change in the grid spacing. The problem is not just due to unrealistically high diffusion.

Observations assimilated

are only OGDR's used for altimetry, what data volume and coverage is assimilated on average from each platform in each of the two 24 hr analysis windows.

Section 2.3 - a later finding is that there is a large response from the initialisation to the first 24h forecast. As mentioned by the authors it is indicative of both overfitting and lack of independence of the verifying observations. Some comments below relate to clarifying the source of this:

a. Why is T-54h required for the atmospheric forcing?

b. What is the resolution of the atmospheric model used in the hindcasts? Could you add this to Table 1

c. How much impact does step 6 have positive or negative? Are these observations overfitted compared with the other 65%. Can you compute the forecast and analysis innovations for these observations in the [T-48h,T-24h] and determine if they are statistically different from the other 65%.

d. Is there any systematic bias associated with the observations available for the [T-24h,T+00h] analysis? i.e., are they based on the same small number of platforms. Are any of these observations included in the verification analysis.

e. Does Water et al., 2013 show whether FGAT matters over a 24h time window?

Section 4.1

a. "reanalysis observations are filtered. . ." do you mean sorted to a common subset?



b. "The large improvements at high latitudes can be mainly attributed to the . . ." Can you quantify this statement with a metric?

c. Typo "Topical Atlantic"

Figure 1

The scale for the plots is being set by the free run. Figure 1d, 1e and 1f are being impacted by this where it is difficult to see the performance of v11 vs v12. I suggest that the x-axis be reduced and where the free model goes beyond the limit place a bracketed value above the line to indicate its value. A similar approach could be adopted in Figure 2.

Figure 3

It would be instructive to see the tropical Pacific as an additional 2 panels to shed further light on what the distribution of error is in this region.

Salinity profiles

a. The paper indicates that there is a general deterioration in the performance of v12 in the ocean interior. This requires some further discussion. In some cases the systematic errors have increased but this is not always the case (figure 2). Does Water et al offer some clues on how altimetry is projecting onto salinity. Previous FOAM systems have converted altimetry into synthetic profiles. Is this still the case? It is not clear in the system description.

b. The large bias in the North Atlantic is associated with a large cool bias in the temperature. Whilst the Mediterranean is associated with a warm bias. Is this consistent with the precipitation hypothesis? i.e., is the North Atlantic bias associated with winter cooling while the Mediterranean associated with a net warming. The North Atlantic temperature profile bias is the same sign as the SST. However for the Mediterranean these are apposing sign. Section 4.2 points out a change to the vertical mixing scheme, what role has this change played?

c. Add a comment on the Arctic, which is the only region with a positive bias in the free run. Does the precipitation include snow?

Sea surface height

NEMO is a volume conserving model. The long term drift in the model needs to be explained by a change in volume rather than one in terms of steric expansion such as low density water. A volume conserving model does adjust sea level for steric effects but this is done without any net change to the global volume.

Near-surface velocities

a. Remove 2nd paragraph, repetition of the 1st.

b. It is worth emphasising that this verification is based on independent observations. How many observations are being used in each basin? Why is the Indian Ocean excluded?

c. For the Tropical Pacific both the meridional and zonal components seem to be inferior to v11 which is not consistent with the text.

d. Final comment related to data assimilation introducing additional kinetic energy. Is this the DA u,v fields or the initialisation shock of the T/S/eta state? Do you track KE during the initialisation?

Figure 5

a. Increase fontsize – difficult to read in this scaled down version. Also use the full column width.

b. It is instructive to see biases in each component. However, it is also useful to perform the analysis on the total vector. For example, Kundu, 1976, JPO using a complex correlation.

Forecast validation



Interpolated climatology? Linear interpolation?

SST

The in situ observations are not independent. The same drifting buoys are propagating with the fluid parcel and errors are correlated. This will be a contributing factor to the skill in persistence.

Temperature profiles

a. The deterioration in the forecasts for profiles requires some additional attention. The in situ observations in this case are independent indicating that away from the assimilated observations the RMSE over the global ocean is comparable to v11. There is a clear indication that the profiles are fitting information local to the profile that is not present in the remainder of the model.

b. It would be instructive to compare the power spectrum of the analysis and the forecasts. Is there larger power in the high wavenumbers as speculated which is subsequently dissipated through the forecast period?

Salinity profiles

It is worth noting that the hypothesis of over mixing for the temperature biases does not seem to be present in the salinity results.

Figure 7 – increase fontsize of axis labels

Section 4.3

a. there is a lot of material introduced for a single case study example. If supplementary material is permitted it would be desirable to show at least two other examples.

b. The example shown would benefit from the calculation of spatial correlations and included in the text to quantify the improvement of v12.

Summary

a. The opening statement of the 2nd paragraph must state that the results are mixed. There are clear advantages when the observation density is high but for regions with sparse observations the performance has deteriorated. Some of the results also need to caveat the level of independence of the observing system.

b. In the discussion there is some confusion between the results for the analysis and the forecasts in respect to the profiles

c. The results for the surface velocities are encouraging. Independent observations (a point to be emphasised) and clear improvements for v12.

d. There is no information/diagnostics presented on the initialisation shock such as global KE. Given the rapid deterioration in the T/S profiles in the forecast compared with v11 these diagnostics would be instructive.

---

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