**Ocean Science
Discussions**

Interactive comment on "Evaluation of real time and future global monitoring and forecasting systems at Mercator Océan" *by* J.-M. Lellouche et al.

A. Sellar (Referee)

alistair.sellar@metoffice.gov.uk

1 Overall assessment

The authors describe an operational ocean forecast system, and present some results of validation of hindcasts of this system. The evaluation of the hindcasts is comprehensive and the conclusions are broadly justifued by the evidence. The figures, while numerous, are well chosen and demonstrate the breadth of the validation peformed.

I would like to see some clarification regarding the correspondence between these "forecast" results in section 4.2 and the accuracy of the operational forecasts (I expand

C285

on this point below). Finally, the quality of the writing is variable and needs to be improved before publication: many sentences are over-complicated or confusing.

2 Major issues

In the section on forecast accuracy (4.2), the statistics come from the data assimilation innovations. For readers who are not familiar with data assimilation, the authors should explain why the innovations give an indication of forecast quality. They should also include the caveats on this, i.e. briefly describe the differences between the "forecast" used as the DA background, and an operational forecast. For example, the surface fluxes are presumably from NWP analyses rather than NWP forecasts.

Also, it is conventional when presenting forecast scores to indicate the forecast lead time. Presumably in this work these are averaged over all 7 days of the run, in which case the results are indicative of the average performance over the first 7 days. This should be explained in section 4.2, and figures 13 to 17 should include some mention of the lead time they refer to, if it is possible to do this without making the captions too long.

In section 4.4 I am a little confused about the drift in the IRG_DEV system. The figures show that there is a difference between the model 2011 in 2007, and that the model is not biased relative to the observations in 2011. Therefore the model is either biased in 2007, or there is a similar change in the observations between these 2 years (presumably the former).

3 Specific comments

3.1 Text

I highlighted some of the sentences which I found confusing, but there are more than I have listed here.

- p1126 line 2: "It is declined in different configurations" I don't know what this means.

- p1127 line 15: "IGR" -> IRG

- p1131 line 7: "specificities" -> specifics (also in tables 1 and 2)

- p1131 line 18: It would be useful to say which observations go into the RTG analysis, to help understand the difference from the "AVHRR+AMSRE" analysis.

- Section 3 is longer than is justified by its relevance. I would recommend moving some of the "history" information to the introduction, and to make the rest more concise: there is some information which could be removed with any detriment to the paper.

- p1134 line 20: remove "with"

- p1135 line 27: "control" as a verb has different meanings in English and French. "checked" or "monitored" would do here.

- p1136 line 4: "Most NWP centres publish quality reports on a regular basis". I'm not sure if this is true. -> "Some"?

- p1138 line 11: "It is therefore considered that the QC built from GLORYS2V1 may be applied to other systems". Is this referring to the QC method, or the QC flags for these observations, or something else?

- p1138 section 4.1: "Best analysis" is not a term I recognise. "Best estimate" is a term defined by the GODAE community, though just "analysis" would be fine here.

- p1140 line 5: "We initially checked that all the systems were closer to the observations than the climatology." Slightly ambiguous. Sounds like |model - obs| < |model - clim|. Presumably the intention was |model - obs| < |clim - obs|.

- p1140 line 10-16: These sentences are a little too confusing. I think they are making relevant points, but could be made clearer.

- p1140 section 4.1.2: Although there is some independent information in OSTIA because it assimilates observations not included in RTG or the AVHRR+AMSRE product, it should be pointed out that it is not completely independent because it shares many observations with these products. The "reduced bias" is to some extent a reflection of the fact that OSTIA shares more observations with the AVHRR+AMSRE product than with RTG. In the (cloudy) high latitude coastal regions where the bias is reduced by the move to AVHRR+AMSRE, both this product and OSTIA are likely to be dominated by AMSRE (the only microwave instrument outside of the tropics used in either analysis) and in-situ data and so will be very consistent. I am sure the AVHRR+AMSRE product is a better dataset to assimilate than RTG, but the comparisons to OSTIA don't necessary prove this. I suggest that the authors include some caveats on the conclusions which can be drawn from the OSTIA comparison.

- p1140 line 25: "This SST product has the same quality level as OSTIA and both

display better performance than RTG especially in high latitudes." Include a reference.

- p1141 line 8: "increment ... rejected". This is an interesting metric. It should be defined. (is it = increment - analysis + background?)

- p1141 line 12: "The concurrent effects of bulk fluxes and of IAU correction are not efficient in this region." I don't know what this means. That the increments are not sufficient to control the effect of the fluxes?

- p1144 line 19: "It is smaller than the different internal errors involved in the system." This is unclear. Smaller than the uncertainty in the observations? If so, please include a reference.

- p1145 line 27: "is" -> though

- p1146 line 1: "default" -> defect?

- p1146 line 4: "We checked that the correction of the precipitations lead to a deficit in summer". I don't know what this means.

- p1146 section 4.3: It might be helpful to the reader to point out that the results in this section are based on analyses.

- p1146 line 22: "water masses charecteristics" -> water mass characteristics

- p1147 line 15: "which" -> whose; "teleconnexions" -> teleconnections

- p1147 line 18: "As" -> While?

- p1149 line 9: Please define "cumulative trend".

C289

- p1149 line 23: "The latter confirms the seasonality of the cold bias that is observed," The authors could refer to the earlier section (or figure) where this was discussed.

- p1151 line 25: "The IAU prevent from keeping the correction of the initial condition in the model because of the bulk formulation." This sentence should be re-written.

3.2 Figures

- Figure 3: the words "temporal and geographical" could be removed without changing the meaning.

- Figure 15: A very short caption. Please include something like "skill relative to persistence", and include the "(x100)" in the caption - not just in the plot title.

- Figure 19: The map on each plot showing the location of the section is too small to see on the printed page. Please either include a larger map, or describe the position of the section more precisely in the text.

- Figure 24: The caption says 2010 but the plot title says 2011.

_____

Interactive comment on Ocean Sci. Discuss., 9, 1123, 2012.