

## **Second review of “Evaluation of real time ... Mercator Ocean”, by Lellouche**

### **By Referee #1**

I'm satisfied with the additional detail the authors have included in the revised paper about the assimilation system and how it's configured and implemented. I'm also satisfied with the explanations about the results. However, I'm less satisfied with the author's actions responding to my comments regarding the localisation and sub-sampling. Details follow.

Re: point 1 – subsampling and localisation

The authors acknowledge a flaw in their configuration in their response to my review, but they don't address it in the paper. They later (responding to point 3) state that addressing the issue identified by doing less subsampling (taking every 2<sup>nd</sup>, rather than every 4<sup>th</sup> point) “does not bring a significant difference”. I remain sceptical about this. Perhaps the analyses they produce look superficially the equivalent – but the dynamical balance of the analysed fields must be different. The authors are effectively reverting to data insertion – a very bad approach.

The authors point out that their approach is suitable for their 1/12 degree model. I agree with this. But their approach is certainly unsuitable for their 1/14 degree model. Yet they want it published without acknowledging the problem in the final paper.

### **Review of Lellouche et al. by Mike Bell**

This review suggests a lot of minor changes to improve the clarity of the presentation. The two main points are marked by \*\*\*. Most of the points are easy to respond to. I expect you to respond to all of them.

#### Abstract (page 1)

Lines 13 & 14. “zoom” would be better described as “nested model”. “open boundaries” change to “open boundary data”.

Lines 22 – 23. Suggest you replace “, starting with an introduction of” with “and”. This is quite a major change to the abstract !

Page 1 lines 24-27. This sentence seems to conflate data quality control (section 3.3) with the scientific assessment (section 3.1).

\*\*\* The abstract should summarise the main results presented in the paper. At present it presents very few results.

#### Introduction

##### Page 2

Line 5: First sentence needs re-writing; it doesn't make sense. “implemented” is the wrong word and “which” seems to refer to GMES.

Line 9: “prevention” would be better as “response”

Line 10: ice sheet surveys ?? Do you mean sea-ice monitoring ? Please append “management” to “water quality and pollution”.

Line 16: define MFCs with capitals. Don’t define again in line 30.

Line 20: “synthetic” is an odd word to use. “comprehensive” might be better

### Page 3

Line 5 replace “as a difference” with “by the differences”

Line 9: “is responsible” change to “has primary responsibility”

Line 17: should say it is a Mercator grid with 1/12o grid spacing.

Line 30: “features” should be “implementations” ?

Lines 31-33 Please use “Section” rather than “Sect.”

### Page 4

Line 17: “homothetic” is a little used word.

Line 19: “stacked in” replace by “within”

### Page 5

Lines 20 – 23. Does daily fluxes mean daily average fluxes ? This would contradict line 13 about 3h sampling. It sounds odd to generate idealised diurnal cycles. Please clarify.

You should mention somewhere in Section 2.1.1 that the system does not include tides so does not attempt to simulate continental shelf areas with large tides.

### Page 6

Is it possible to say how are  $N$  and  $v_{\max}$  chosen ? For example is the aim to represent the forecast errors in the ocean mesoscale field or El Nino ?

### Page 7

Line 24: How are these scales calculated ? For example do you use the Argo and GLORYS2V1 data with the Hollingsworth et al technique ?

Lines 28 and 29: This statement is too vague. Scientific results should (in theory) be reproducible from the information within papers.

Figure 3: Too much of the globe is covered by the dark blue which covers 0 – 100 km. You need a colour scale in which 0 – 50 is distinguished from 50 -100 km.

#### Page 8

\*\*\* First para: I agree with one of the reviewers that using a  $1^\circ$  analysis grid over most of the domain for the  $1/4^\circ$  model with localisation radii of 200 km is not satisfactory. Could the authors please note in this section or (probably better) the conclusions section that re-assessment of this aspect of the assimilation is a high priority for future work.

Lines 10-12. Why is 75 km mentioned here. Figure 3 does not show results from this test.

Lines 19 – line 12 on next page. This approach means that the analysis is 3.5 days out of date by the time it is calculated in an operational system. The approach is not unusual for re-analyses. Also the assimilation increments have not all been added to the model 3.5 days into the IAU stage so it is doubtful that it is the best analysis achievable.

#### Page 9

Line 15: replace “later” by “last”

Line 17: Why are the large-scale biases in the model confined “under the thermocline” ? The surface fluxes are likely to have large scale errors and result in large scale biases near the surface too.

Lines 26-27: These tendencies are applied up to three months after the time at which they are calculated. This could give quite a large phase error in corrections to the seasonal cycle. Does the bias correction have a significant impact on the results ?

#### Page 10

Line 3: “kind of parametrization” does not the right description. It might be changed to “approach”.

Lines 14-16: does including velocities in the control vector change the results when no velocity data are assimilated ?

Line 23: the Celtic Seas and North Sea are tidal areas (see comment earlier)

Line 31: “covariance” is a “variance” in this case ?

#### Page 11

Line 1: how were the spatial correlation radii modified ?

Lines 2-3 and 11: Starting from October 2009 rather than October 2006 would by itself have a large impact on unrealistic salinities in waters where there are few observations available to the system.

Line 26: Garric et al. 2011 is not a very satisfactory reference. (EGU abstract)

Lines 30-32 (point 4): This is informative but may confuse the reader, the surface height budget (rather than surface mass budget) having other contributions as well as the freshwater flux.

#### Page 12

Lines 15-17 (point 9): It seems odd that a quality control of T/S profile is only now being implemented. It should perhaps be mentioned that qc flags from Coriolis have previously been available.

Line 32: Suggest you change to “instantaneous or time mean”

#### Page 13

Line 8: Suggest you change to “calculating various statistics of the differences between all available ... and their model equivalent ...”

#### Page 14:

Line 16: move the comma: “observations, such as currents”

Line 32: background errors are often represented by SOAR functions rather than Gaussians.

#### Page 15

Lines 7 – 13. Is this the only q/c test that has been implemented ? If so you probably need to refer to the Coriolis data centre q/c tests again if you are also using them.

#### Page 16

Line 2: For clarity say “all our systems”

Line 10: Delete “while” and put “demonstrating” in its place and tidy up the sentence.

#### Page 17

Figure 7 and 8 Is it known which upgrades from V1V2 to DEV were responsible for most of the improvements in the N Atl and Med ?

Lines 27-29: The biases in the Med HRZ\_DEV are much smaller than HRZ\_V1V2.

#### Page 18

Line 20: replace “off” by “of”

Line 26 delete “experience”

Figure 9: caption (e) “system after 7 days” should read “system every 7 days”. Just a comment (no action required) the mean analysis errors in (a) and increments (e) are rather large.

#### Page 19

Figures 10 and 11: It is strange to show IRG\_Dev ice concentrations in Oct 10 but not Mar 10 and to omit IRG\_Dev sea-ice drifts.

Line 10: “again” implies that IRG\_Dev previously discussed

#### Page 20

Line 4: Strange not to call it the Grodsky correction ?

Line 12: Fig 14b presents a “speed” bias rather than a “velocity” bias

Figure 14: It would be helpful to state the value of the half width of the distribution function in Fig 14c.

#### Page 21

Section 4.2.1: Could you explain (or refer to a reference describing) how the sea level anomalies are calculated, e.g. are they simply 1 second along track averages or has some filtering been applied to them ?

Lines 18-19: replace “strongest” by “largest” and “as well as” by “and”

Line 31: The mean innovations look to be 1 cm or even 2 cm at times. This is not close to zero in my view.

#### Page 22

Lines 9-21. Does this skill score penalise high resolution models ? In general the HRZ scores seem to be lower than the IRG scores.

Section 4.2.2:

I agree with one of the reviewers that it is disappointing that this section only presents results for the North Pacific region.

#### Page 23

Lines 3-19: This is really an investigation of deficiencies in the surface fluxes and vertical mixing processes so belongs in section 4.3 rather than 4.2.

Fig 20: The “blue-green” colours for the salinity plots are difficult to make out – some are positive values and others are negative. The scale probably needs to be reduced by a factor of 5 to -0.1 to +0.1 psu.

Line 18: wave mixing – surface waves or internal waves ? The seasonal cycle of surface waves is represented through the seasonal cycle of surface winds.

#### Page 26

Figure 26 should really present the sum of the temperature increments made to the model over a 1 year period. But this would not be a flattering plot. Does Figure 27 show the 1<sup>st</sup> EOF for the temperature increments made every 7 days ?

Lines 27-29. The analyses are not suitable for climate studies. This does not mean they should not be released. They may be useful for other studies.

#### Page 27

Line 12: Why does this conclusion only apply down to 500 metres (Argo floats go much deeper)

Lines 22-23: Replace “detail” by “detailed”.

Line 30-31: the reproduction of sea-ice edge is likely to be related to the assimilation of SST data as the next sentence confirms.

#### Page 28

Lines 13-14 “The HRZ\_DEV” should say “The drift in the HRZ\_DEV”