Answer to reviewers.

Interactive comment on "Forecasting the mixed layer depth in the north east Atlantic: an ensemble approach, with uncertainties based on data from operational oceanic systems" by Y. Drillet et al.

Dear editor,

We sincerely thank all the reviewers for their review of our paper. The aim of this letter is to detail how we addressed their comments on our manuscript No 2014-31 untitled Forecasting the mixed layer depth in the north east Atlantic: an ensemble approach, with uncertainties based on data from operational oceanic systems.

Anonymous Referee #2

Received and published: 1 August 2014

GENERAL COMMENTS This study is focused on the mixed layer depth in the North East Atlantic near the Porcupine Abyssal Plain for May 2013. This phenomenon is studied using in-situ observations and the hindcast/forecasts products from 4 different forecasting systems. An introduction and contextualization of the studied process is missing as well as a short general evaluation of the four forecasting systems. The structure of this paper is not very well organized and it is quite difficult to follow the discussion (in particular subsection 4.2 and 4.3). The methodology applied is not always clearly explained and justified. There is a very detailed description of the results but their interpretation is very poor. Both the abstract and the conclusion paragraphs are quite general without a clear focus on the aim/results of this study. The achievements of this work are not very clear. Why have you decided to use 5-day forecast even if the 4 forecasting systems considered in this work produced 7-day forecast?

We thank the reviewer for his comments. The introduction has been rewritten to give a brief overview of the evaluation of the four forecasting systems, and to present the studied process and the interest of studying it. The aim of this study has been more precisely explained, and a larger part has been given to discussion of the results. In this study we use 5-day forecast because the Ibi36 system provides only 5-day forecast; this information was missing and it has been added in table 1.

SPECIFIC COMMENTS

ABSTRACT: Page 1436 line 14-15: reformulate this sentence. It is not clear the meaning of "forecast time delay" and the general statement of this sentence.

The abstract has been modified, the last sentence is now: "The impact of atmospheric forecast error, and for the wind field in particular (miss or time delay of a wind burst forecast), is also quantified in terms of occurrence and intensity of mixing or stratification events."

We hope it gives a clearer description of the objectives of the paper.

1. INTRODUCTION:

Page 1437 line 5-6-7: please specify which is/are the physical process/es involved

A description of the physical processes at play has been added in the introduction.

Page 1437 line 17: “forecast length” -> please specify what “forecast length” means.

We mean by forecast length the time elapsed from ocean initialisation This expression has been clarified in the text.

Page 1437 line 17-18: “ several . . .. Atmospheric forcing were used” -> It seems from the description of the 4 forecasting systems that are all 4 forced by the same ECMWF analysis/forecast fields. Please explain.

The four systems are theoretically forced by the same ECMWF analysis/forecast fields, but for technical reasons it is not always the case. The systems are launched sequentially, one after another, so that during the interval of time between the first and the last launch, the atmospheric forcing may have been updated (the atmospheric forcing is updated every 6 hours). Moreover bulk formulae are used to compute atmospheric fluxes and this computation is another source of differences between atmospheric forcing used by each systems. These explanations have been added in the text.

2. FORECAST PRODUCTS AND OBSERVATIONS

Page 1438 line 2: Is the Atl12 a stand-alone system or it is nested into the GLo4 or GLo12 system?

The Atl12 system is nested in the Glo4 system: the Glo4 system provides boundary conditions to the Atl12 system. A line has been added to Table 1 for boundary conditions.

Page 1438 line 16

– Table 1 – Could you please specify if you assimilate L3 or L4 SST products?

Table 1 now specifies that the systems assimilate L4 SST product.

Page 1438 line 26-27: Please explain why the ATL12 free experiment has been necessary. This is an ad hoc experiment carried out for this study. The reader doesn’t have to wait till section 4 to understand it.

The ATL12 experiment has been necessary to better evaluate the role of the initial state in the uncertainties of the MLD estimates. This explanation has been added in section 2.

Page 1439 line 10: Please explain what do you mean as “system” in this contest.

System refers to Mercator Ocean forecasting systems, it has been modified in the text.

Page 1439 line 12-13: which criteria is adopted to subsampled to one observation profile for each instrument per day? The description given is very short and not detailed. Please clarify and give a detailed description and explanation.

The description of the subsampling method has been clarified in the text.

Page 1439 line 17: why don't you (re-)write the equation used for the mixed layer computation ?

This part of the text has been rewritten for clarification. The reason why we did not use an equation is that it is rather an algorithm than an equation, and we hope it is now better described.

Page 1440 line 1-2-3: there is some literature about this phenomenon?

The spring stratification is a crucial phenomenon for the onset of phytoplankton blooms in this area (Mahadevan et al, 2012). This reference has been added in the introduction.

Page 1440 line 7-8: it continues to be not very clear which kind of pre-processing, interpolation, mean . . .. has been done to sub-sampled, filter, . . . the available pool of observations (see also comment on Page 1439 Line 12-13). Please explain.

Changes in the text have been made, also in response to the previous comment.

STATICS

Page 1441 formula (1): please explain why you have decided to use H instead of Obs. Give a more detailed explanation on how you have computed the Persistence of initial state and of the observations.

The formula written here was false and has been replaced by the right one (which uses Obs instead of H). Note that the calculations had been preformed with the right equation.

The persistence is computed by replacing the forecast fields by the hindcast fields (initial condition of the forecast), so that the hindcast field of the initial day is repeated each day of forecast.

Page 1441 line 18-19: Explain why the Glo4 performs like this.

This is due to the fact that Glo4 misses the first stratification event S1 (added in the text).

Page 1441 line 21-22 "The results are very similar for H up until the 1 day forecast": don't you expect higher accuracy for H respect the forecast? Page 1441 line 22-23-24 "The dispersion for all the systems is small.." do you have an explanation ?

One would expect a slightly better accuracy for H respect to the 1-day forecast. The difference is not significant mainly due to the fact that the atmospheric forcing which plays the dominant role in this process does not differ very much from hindcast to 1-day forecast (good predictability of short range numerical weather predictions, see for instance ECMWF technical reports on forecast verification). Additionally, the update of ocean initial conditions with data assimilation only takes place once a week in Mercator Ocean systems. In consequence most of the time H and F0 differ only by their atmospheric forcing. And as the atmospheric forcing is the same for all the systems (almost the same, as explained before), the dispersion is small.

Page 1442 line 2 -> to fully understand this statement, I need to know exactly how you compute the persistence (see first question on this section)

Question answered previously.

Page 1442 line 4 " . . .except F0 in GLO4": explain why

Done in the text.

Page 1442 line 6: ok but it is important also to state that from F0 to F1 the relative increment of the forecast respect the persistence is the highest .

This is now explained in the text.

4 MIXED LAYER DEPTH FORECAST DURING MAY 2013 4.1 Description of the mixing and stratification events

Pag. 1443 line 23: "..while for M3 the response is faster (only one day)": could it be because the wind relative increase is the biggest (around 10 m/s) (for M1 and M2 is around 5-7 m/s) and because there are 3 consecutive days of constantly increasing wind ? In M1 there is 1 day of increasing wind and in M2 there is a less homogeneous wind increasing phenomenon.

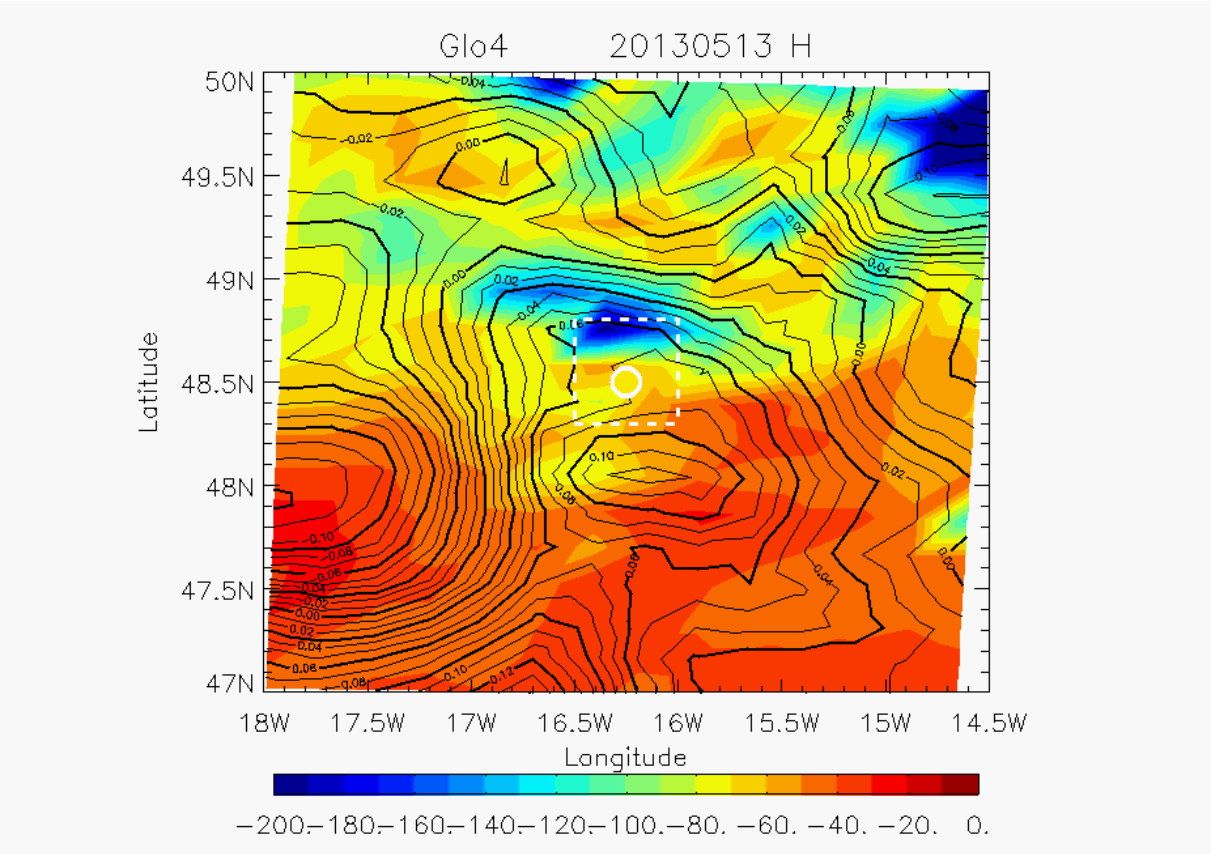
We agree with the reviewer assumption and added it in the text.

4.2 Evaluation of the hindcasts

First part (comments of figure 7): the description is quite detailed but there are very few explanation of the characteristics/differences described. Second part: it is mainly based on figures 9-10. These figures are not very clear. The fields in the small dotted box are difficult to inter-compare among the different systems. A zoom in is needed to better understand the comments done in this section. The observations represented in figure 2 should be overlapped to this figure in order to have a full view of the situation.

Fisrt part:some explanations have been added in the text, especially on the differences between the ocean circulation features of the four ocean forecasts.

Second part: we agree with the reviewer comment. The area displayed by the figure is now smaller so that the central box appears bigger. The observations of the day is now overlayed inside a white circle.



Page 1445 line 15-16-17: this consideration should have been written in the previous section where the method and the experiments strategy are explained.

An explanation is now given also in section 2

4.3 Discussion for the forecasts

4.3.1 Forecast of the 1st mixing event (M1)

Page 1446 line 12-13-14-15: such a sentence needs an adequate explanation. In principle the hindcast and the first days of forecast have a better atmospheric forcing and should perform better than the 3/4-day forecast.

An explanation is now available. One would have expected F0 to be more accurate than F4. As already mentioned in section 3, the Glo4 is better at long forecast lengths than at short forecast lengths for wrong reasons. The 4-day wind forecast is less than the analysis wind (4m.s-1 rather than 13m.s-1; purple dots for 9 May in Figure 8, top panels). The Glo4 system seems to produce too much mixing in response to realistic wind (H, F0 to F3), and thus produces a more realistic mixed layer depth when the wind is underestimated (F4). The other systems, Glo12, Atl12 and especially Ibi36 have a more realistic answer to wind forcing and for these systems the error increases with forecast length.

Page 1446/1447 line 24-3: are there differences in the wind field used for the forecast only between the GLO4 and all the others systems (IBI36, ATL12, GLO12)? In figure 8, top panel there are shown

only the wind from GLO4 and Atl12. The ECWMF forecast fields used by one system could have been released 24hr before the ECMWF fields used by another system?

The time difference between two ECMWF fields releases is 6 hours, so the ECWMF forecast fields used by one system could have been released 6hr before the ECMWF fields used by another system. There can differences in the wind field used for the forecast between all the systems, not only GLO4 and ATL12. But as GLO4 is launched first, and ATL12 is launched third, it is more likely to have differences between GLO4 and ATL12 then between GLO4 and GLO12 (which is launched after GLO4).

Page 1447 line 3: "(using for example the previous analysis cycle)": please could you explain? It's evident from figure 8 that there are several differences for F4.

ECMWF fields are released every 6 hours.

There are not only differences for F4 but for all the forecasts days. It appears more clearly with F4 because there is more dispersion for the 4-day forecast then for the other.

4.3.2 Forecast of the 1st re-stratification (S1) and 2nd mixing (M2) events

Same comment done for figures 9 -10 in the previous section.

As already mentioned, new figures 9-10 are available

4.3.3 Forecast of the 2nd and 3rd stratification (S2, S3) and 3rd mixing (M3) events

4.3.4 Atmospheric forcing vs. initial state in the uncertainties

Is the statics of the atmospheric forcing computed on the original ECMWF analysis/ forecast fields or the fields computed via bulk formula from the model ? It seems that is it the latter. In this case as explained in the previous section (section 4.3.1, page 1446, lines 24-27), there is no homogeneity in the atmospheric forcing used by each system (see comment on Page 1446/1447). Can you quantify how this inhomogeneity is taken into account by your statics?

The reviewer is right, we use bulk formula, and do not take into account the inhomogeneity that it produces in the forcings. This fact is now mentioned in the discussion of the results even if the quantification of this uncertainty is not perfomed.

Page 1449 line 14-15: "This small decrease in correlation indicates that the initial state has a small effect" -> justify this sentence, the explanation given in the text is quite poor and not clear.

Done in the text.

Page 1449 line 19-20-21: as above, provide a stronger justification to this sentence.

Done in the text.

Page 1449 line 24: ". . . error id due to fresh water flux" . . . can you explain why, as shown in fig. 11 top panel, F2 water flux correlation is less than F1?

This is not the correlation but the normalized standard deviation. First, current numerical weather prediction systems have difficulties to produce realistic water fluxes over the ocean, in analysis mode

as well as in forecast mode. It is known that ECMWF analyses and forecast both overestimate precipitations in the tropical regions. Second, water fluxes may vary a lot inside a given day, or between two instances of a weather forecast. In consequence errors in water fluxes averaged over one day may behave this way due to pure random effects, and a bigger sample may be necessary in order to derive robust statistics for this variable. This comment has been added in the text

Page 1450 line 10-11: explain better this sentence or remove it.

The sentence is now removed

Page 1450 line 22-23: is the ATL12 FREE experiment initialized with the same Initial Condition of Atl12? Please specify.

This simulation was initialised in March 2013 with the analysis provided by Atl12 system, and forced using the atmospheric forecast analysis to the end of May 2013. This is now better explained in section 2.

In consequence the ocean forecasting system has been run for one entire month without data assimilation before the beginning of May, which is enough for the system to "forget" all corrections from the data assimilation.

Page 1452 Line 4-5-6: this sentence is quite ambitious. The description and explanations you give are quite preliminary and not very detailed.

We have improved the explanations in the text.

5. Conclusions

Page 1453 line 2: ".. less than 20m. . .": this number is strongly dependent from the model vertical discretization, it should be written and explained.

It is now explained in the conclusion.

Page 1454 line 19-20: "The effects of horizontal circulation, particularly around eddies or along strong fronts, have been illustrated for the model mixed layer": in this work there isn't a detailed study of these phenomena.

The impact of the mesoscale is illustrated by the figures 9 &10 and this paper shows that Ibi36 is the best among Mercator Ocean system in reproducing the mixed layer depth and variability. It can be inferred that this result is linked with the ability of Ibi36 to reproduce mesoscale due to horizontal resolution and specific model tunings. This is now better discussed in the conclusion.

Page 1454 line 26-27: this sentence is very general. Please specify better which developments can have a strong impact on the improvement in the mixed layer depth forecast.

Improvements can be obtained by using the full resolution of the atmospheric forcing (except for Ibi36, the native 1/8° atmospheric fields are put on a ¼° grid), or by modifying the TKE mixing scheme or implementing the k-ε mixing scheme (already used in Ibi36). A better vertical resolution could also

improve the MLD forecast, as well as introducing the mixing due to waves. All these potential improvements are now mentioned in the conclusion.

TECHNICAL CORRECTIONS

Page 1437 line 18-23: the numbers of the sections does not correspond. Please correct.

Corrected

Page 1444 line 2 . “Fig. 8”: fig. 8 is mentioned here but is not commented in the following lines.

Corrected

Page 1454 line 16: typing error . . .”atmospheric focing” ! atmospheric forcing

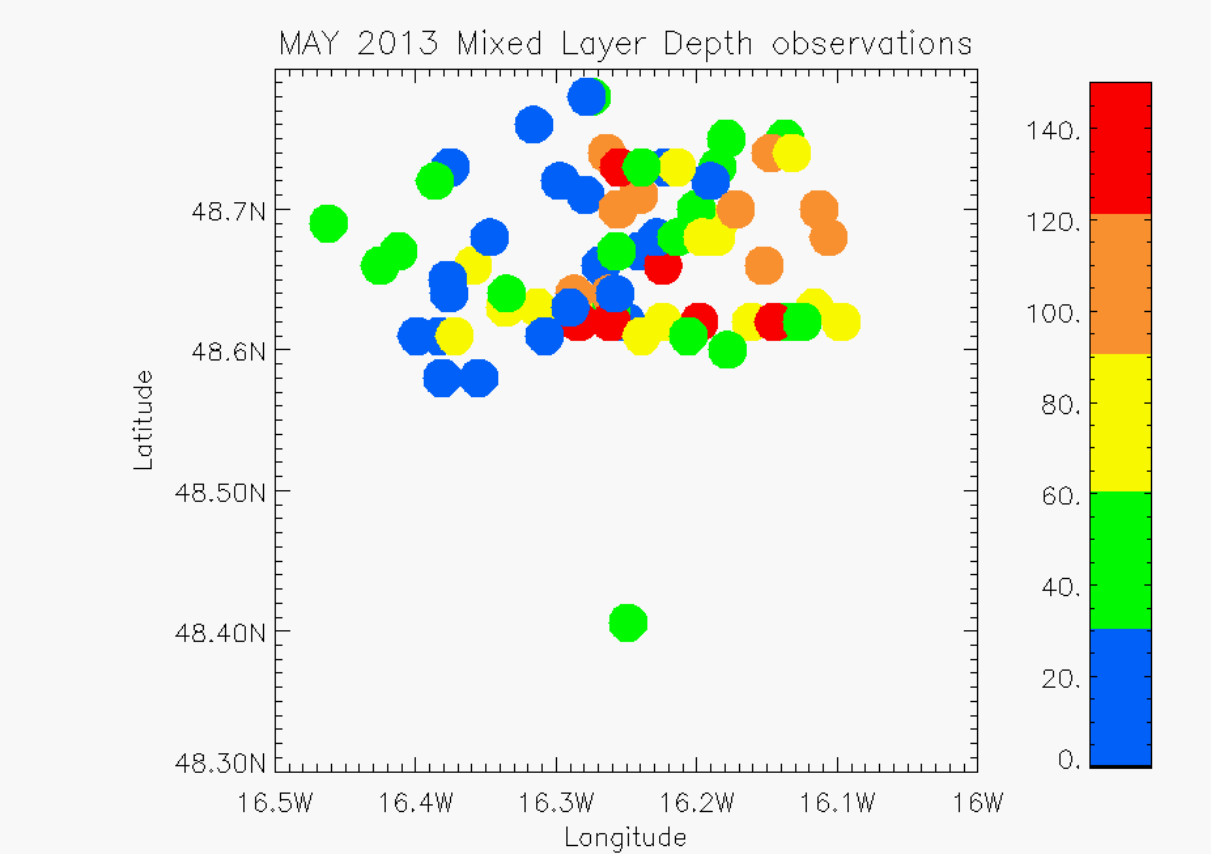
Corrected

Page 1454 line 22: typing error “..validate properly, The coverage. . .” -> “. . .validate properly. The coverage..”

Corrected

Figure2: most of the numbers of the circles are not readable.

Figures have been improved showing the date only on the zoom and with less color



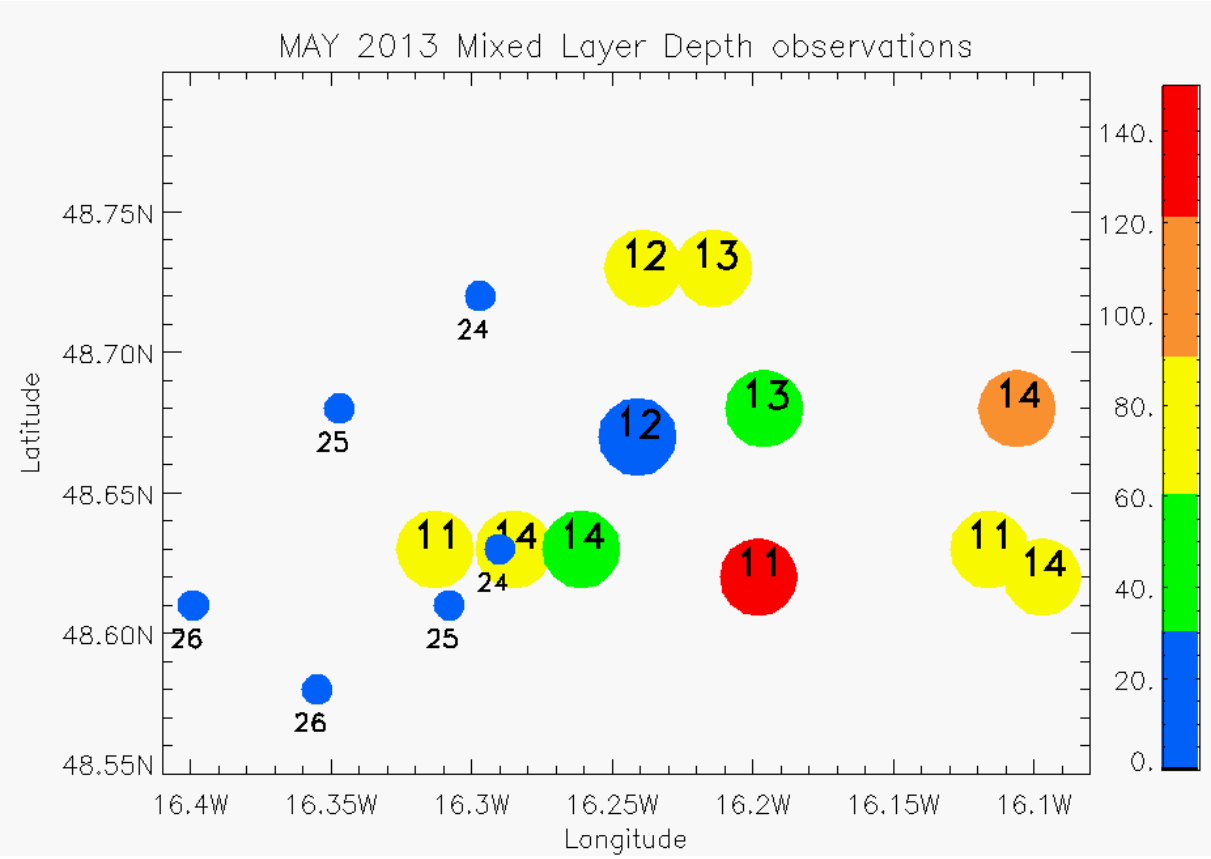


Figure 8: bottom-right panel! typing error in the title “Frash”! “Fresh”

corrected

Figure 12: the title of each panel is difficult to read. The character font is too small.

corrected