**Ocean Science
Discussions**

Interactive comment on "Modelling temperature and salinity in Liverpool Bay and the Irish Sea: sensitivity to model type and surface forcing" *by* C. K. O'Neill et al.

C. K. O'Neill et al.

clne@noc.ac.uk

Received and published: 5 July 2012

Author response to Referee 3 for manuscript os-2012-12

Thank you to Referee 3 for his/her comments.

General comment:

As already mentioned by two other reviewers, the paper is finally difficult to evaluate due to the confusion introduced by the authors in the way they processed some of the data used for model evaluation. Briefly stated, temperature and salinity data used in this study are coming from i) CTD profiles taken on regular cruises, ii) sensors mounted

on board a ferry running from Liverpool to Dublin, iii) a "SmartBuoy" deployed at one location (referred to as site A) in the Liverpool bay and iv) a bed frame co-located with the "SmartBuoy". All these T&S data are influenced by the tide and by the seasonal variability present in the atmospheric forcing and fresh water discharges. However, only the "SmartBuoy" data do properly resolve all the different times scales.

Now, in section 2.2.3 (p.657, l 24-28), just after the description of the "SmartBuoy" data, the authors mention that a running mean (i.e., a mean over 14 M2 tidal cycle) is taken to look at lower frequency signal while tidally-dominated fluctuations are obtained by removing this mean from the original data. In the next paragraph, p. 658 l1-5, it is precised that the tidal signal is removed from the observations and model results, now using a Doodson filter) before performing any statistical comparison.

It is clear that these descriptions were confusing, so we have clarified the text.

If, at some places, the distinction between the different time scales is well mentioned (e.g., Figs. 6, 7, 10 and the parts of text here those figures are discussed), they are other places where this distinction is less clear and the reader finally doesn't know which data he (she) is looking at (e;g., Figs. 3 and 8).

Figs 3 and 8 just show all of the data points over the year. This has been clarified in the text.

It is also rather unclear when and where T&S data coming from the bed frame at site A are used. They are certainly used to produce Figure 7. It seems that they are not used in Figures 3 and 8. But are they used in the computation of some metrics? Captions to Tables 2 and 3 indicate that "all model-observation comparisons" have been used for the computation of r2 and while, apparently, RMS errors are given only for comparisons at the surface in Tables 4 and 5. Is this correct? If yes, could the authors justify their choice?

The metrics are all produced from surface data only, so that they can be compared. The

data from the bed are only used to produce Figure 7. The phrase 'all model-observation comparisons" was intended to mean that the whole year's data were included. This has been clarified.

Another quite disturbing point is that negative values are given, in table 2 and in the text (see, e.g., p 661 l 16), for the squared correlation coefficient, r2. The authors should decide if they "play" with the correlation coefficient, r, or with its squared.

The values used are $r|r|$, so the magnitude is the same as r^2 but keeping the information given by the sign. This has been added to the text.

Accordingly, we consider the paper certainly can't be published as such. It requires a serious revision. It is hard to estimate the time such a revision could take. It's only when all confusing points will have been removed that it will really be possible to review this paper.

Specific comments:

On p 651, l 1-10: the explanation of the so-called strain induced periodic stratification is rather unclear.

This is an interesting if not complicated processes that occurs in regions that are dominated by later salinity induced density gradients. We have modified the text slightly, though really the cited references should be pursued for detailed studies on the process.

On p 652, l 24-25: apparently, there is no explicit horizontal diffusion in the Irish Sea implementation of POLCOMS. What about the AMM implementation?

The AMM model does have explicit horizontal diffusion in deep areas, but it is ramped down to zero in shallow areas so is not relevant in the Irish Sea region.

Section 2.2.2: the frequency at which the ferry is running from Liverpool to Dublin should be indicated.

C671

On p 657: giving the frequency (daily?) at which the ferry is running from Liverpool to Dublin could be of some interest.

The ferry ran daily. This has been added to 2.2.2 the Ferry data description.

On p 655, l 11:is the horizontal resolution of the North East Atlantic NWP model equal to 0.11 in both directions? We would also like to suggest the author present all the horizontal resolution in the same way.

Yes it is 0.11 both directions. This has been clarified.

On p 656, l 4-5: satellite data are independent form model results while model results are not independent form satellite data.

Correct

On p 656: the list of cruises should be presented in a tabular form.

This has been changed

On p 659, l 21: the sentence "RMS errors compared to the ferry data, averaged within 3' by 1.2' bins ..' is unclear. What is averaged? The ferry data? How is this consistent with the sentence (on p 658, l 8: ". . .model results were interpolated in space-time to the locations of the observations." This should be clarified in section 2.2.2.

The averaging is only for the purposes of producing the plot, since there are so many data points. This has been clarified in the text and caption.

On p 659: the header "2.4 Results" should be removed and the following section renumbered accordingly.

Yes this was an error.

On p 660, l 1, Figure 8: we would suggest using the same range for the observations and model results on the different scatter plots event if this could slightly reduce the variability seen on the plots at site A. Showing the regression lines in addition to the

C672

perfect linear regression lines qould be of interest as well.

The scales have been changed.

On p 661, l 16: r2 < 0? As explained above it is actually $r|r|$. The definition has been corrected.

On p 661, l 19-20: if surface salinity at site A is clearly underestimated in both POL-COMS applications, this is not the case for NEMO. Is the same climatological river data used in the 3 models? If yes, the authors should give another explanation.

Yes the same river dataset is used in all models, so the underestimation of salinity is not due to a difference the amount of freshwater entering the system. It is caused by the salinity gradient in the area, particularly in the case of the 12 km model which has only a few grid boxes covering Liverpool Bay. Modelling the freshwater near coastal sources is especially difficult for coarse resolution models given (1) the uncertainty in the freshwater forcing and (2) the insufficient spatial resolution to capture the mixing processes of the freshwater plume.

On p 664, l 16-18: while should r2 be a measure of the model ability to reproduce the seasonal cycle and a measure of model ability to reproduce the tidal variability if this latter has been filtered out by a method or another?

Do you mean how do r2 and Chi2 preferentially filter the tidal and seasonal timescales? The definition of r2 evaluates difference from the mean quantities, and since the annual cycle has a larger amplitude than the tidal SST/SSS signal then this dominates. On the other hand Chi2 has the cumulative sum of differences between instantaneous values and so will score poorly if the tide is, say, always out of phase.

On p 665, Figure 11: we would have expect a negative eastward salinity gradient in the Liverpool Bay.

The scale on Figure 11 is intended to just show the magnitude of the gradient, rather than direction. Naturally the salinity gradient is negative with lower salinity near the

C673

coast. This has been clarified.

On p 665: NEMO better reproduces the horizontal salinity gradient at the latitude of site A (Figure 11) but is unable to reproduces persistent (i.e., staying more than one tidal cycle) stratification that sometimes appears (Figure 7). The more diffusive horizontal mixing scheme used in NEMO is advocated to explain the first point. What about the second? There is nothing in the turbulence closure schemes used in the three models that contribute to the different behaviors?

It could be that the sharper lateral gradient in POLCOMS are sufficient to shut off the vertical mixing when the gradients are vertically shear through SIPS. In NEMO the lateral density gradients are insufficient to shut off the mixing.

Typing corrections:

On p 654, l 4: Lapacian operator(s) should be replaced by either Laplace operator or Laplacian.

Should be Laplacian.