

## ***Interactive comment on “Verification of the mixed layer depth in the OceanMAPS operational forecast model” by Daniel Boettger et al.***

### **Anonymous Referee #1**

Received and published: 1 June 2018

#### General comments:

This is a good paper. It is a regional comparison between two different models utilizing a local contribution to a major community project (GODAE OceanView Class-4 Inter-comparison) as part of the validation of a model upgrade. I found the explanations and figures to be clear and concise, with sufficient background on the methods of MLD estimation as to carry the discussion without necessitating an immediate literature search. It also used subregional groupings (latitudinal zones) that helped highlight the differences between the two MLD algorithms in the two models and how they compared to in-situ observations.

#### Scientific comments

C1

I am a little unsure of the purpose of section 5.2, the sensitivity study based on modifying the accepted MLD threshold definition. In particular, this sentence stands out: "Clearly an optimal MLD definition may include spatially varying thresholds; however, unless carefully implemented this may introduce other errors." To my understanding, the MLD definition used in the paper (de Boyer Montégut et al., 2004) is an optimal definition, calibrated on a global scale on a 2-degree grid. The validation study is dependent on using consistent community-accepted metrics, as is the GODAE OceanView data set it uses. Suggesting that there may be a regional issue with the metric definition may be outside of the scope of the paper, or possibly the subject of another, more in-depth study. At the very least the statement needs support from the literature.

A general comment on the MLD statistics: the de Boyer Montégut 2004 paper suggested that "Comparisons with ocean models should be made with MLDs that are computed at each timestep and then averaged, yielding more consistent comparisons." Considering that the validation study uses the GOV data set which is comprised of daily averaged profiles from which the MLD is then computed, this would appear to present something of a problem. Possibly some discussion about the particulars of the GOV data set construction, perhaps a comparison of daily averages of instantaneous MLD versus MLD from daily averaged profiles would strengthen the discussion. To my understanding, when calculating statistics from non-linear processes, it's best to compute the statistic from the data and then average the results, rather than average the data and then compute the statistic.

#### Technical comments

Section 4.3 line 16-17: "Both models are more skilful than climatology within that data limits (+144 h), and in each region, there is a distinct improvement in version 3.0 compared to 2.2.1." Not sure if "skilful" ought to have two ells (skillful), or if its a regional usage. Also, "that" appears to be a typo. In any case the sentence is not clear. Suggest saying something like this: "Both models are more skillful than climatology within their respective data limits out to 144 hours, and within each region there is a distinct

C2

improvement in version 3.0 as compared to version 2.2.1."

Figure 4(a): The observed mean line (shown in black) is not evident in the image, due to the close overlap of the blue MAE line. The figure is relatively small, and blue/black differences are slight and difficult to perceive. Suggest changing the MAE color to something brighter, possibly using a dashed line for the overlapping MAE line to allow the mean line to show through.

Figure 5: A label on the secondary y-axis for the MAE would greatly help the graphic. It's mentioned in the caption, but a label would clearly identify the axis values.

Figure 6: The contrast between the light and dark red/blue stripes is not very strong (though that may be a product of my printer calibration). Perhaps more contrasting colors for the respective climatologies? This is more of a minor stylistic opinion on my part.

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-69>, 2018.