Interactive comment on “Observed and simulated full-depth ocean heat content changes for 1970–2005” by Lijing Cheng et al.

M. Balmaseda (Referee)
Magdalena.Balmaseda@ecmwf.int
Received and published: 9 May 2016

The manuscript presents a new estimates of global ocean heat content temporal evolution for the period 1970-2005, and compare them with previous observational estimates. The authors then use an ensemble of observational estimates to evaluate the OHC trends in an ensemble of CMIP5 model integrations. They find that the median of the CMIP5 ensemble agrees well with the observational estimates of OHC global trends, both of them showing an acceleration of ocean warming during the period 1992-2005. They and propose to use OHC as a metric to evaluate climate models. The paper is clear and well written: the problem in question is well introduced, the results clearly presented, and there is a levelled correspondence between the numerical findings and the interpretation given.

I have some questions and comments that the authors may want to take into account.

Abstract: “We suggest that OHC be a fundamental metric for climate model validation and evaluation”. The current study only deals with trends of global OHC for a given period of time. Maybe this should be the specific metric proposed. Otherwise the current statement in the last sentence of the abstract is far too generic, and open to mis-interpretation.

Why the validation period does not extend beyond 2005? The period post-2005, when the so-call hiatus started, is of large interest. Can the authors comment on their choice of period? Would the choice of period change their conclusions?

In the observational estimates, the corrections by Durack etal (2014) seem to be included in some of the ensembles. Those use CMIP5 model information to fill the gaps. Then, these corrected estimates are used to validate CMIP5 models. It seems to me like a circular argument. How would the results be influenced by removing the Durack etal (2014) corrections?

If the median is chosen against the mean in recognition of the non-gaussianity of the distribution, the use of Gaussian estimations for the confidence levels (twice the standard deviations) to evaluate the significance of the median seems inconsistent. Are there any other ways of estimating confidence levels for the median using non-parametric distribution?

It is said in the text that the estimate of OHC by Smith and Murphy is discounted because the values are smaller than the others. This is quite an adhoc reason. Can the authors provide a more solid motivation for excluding the estimation?. The estimate is not removed from figure 3, which is misleading.