We thank “xiao boot” for the comments on our manuscript submitted to *ocean science*. Here we provide our responses in blue.

Overview: In this study, observed and simulated full-depth ocean heat content changes for 1970-2005 has been compared in detail. It is found that the rate of the observed full-depth OHC increasing for 1992-2005 is nearly twice than that for 1970-2005. In addition, the ensemble mean of the CMIP5 models is consistent with the observation.

Reply: Thanks for your evaluation. Yes, we carefully quantify the OHC changes of the improved observational estimates of OHC changes and the drift-corrected CMIP5 model simulations. The model ensemble mean shows a better comparison with the updated observations, but models show large uncertainties.

The paper can be published in Ocean Science. Specific comments are listed as below:

1 In the Abstract, the authors noted that “We suggest that OHC be a fundamental metric for climate model validation and evaluation.” I think this conclusion can be revised like “We suggest that the ensembles of the CMIP5 models be appropriate for mechanisms for the OHC changes”.

Reply: In this study, we systematically and carefully compare the global OHC changes between observation-based estimates and CMIP5 model simulations for different depth layers and different time periods. We found that the model ensemble mean is consistent with observations but there is large uncertainty in models, which indicates OHC a fundamental metric for model validation. The importance of OHC as a vital climate metric is well-discussed in a recent study of von Schuckmann et al. 2016.

We agree that the CMIP5 models could be very useful for understanding the mechanisms of the OHC changes, but it goes beyond the scope of this study (we do not focus on the mechanisms for OHC changes). We will mention this point in the discussion section in the revised manuscript.

2 In the Introduction, the authors stated that "We note that the work presented here is broadly similar to the recent study of Gleckler et al (2016) and provides an important independent verification of some of their key findings."This statement is confused because if the results of this study is broadly similar to the recent study of Gleckler et al (2016), why your studies are different ?

Reply: Following this statement, we mentioned the differences: "*the present study also makes use of a larger number of CMIP5 models (24 compared to 15) and observation-based estimates of the 0-700m ocean heat content changes (8 compared to 3), including improved XBT bias corrections and new mapping approaches. We are therefore able to more fully characterise the uncertainties associated with CMIP5 models and place our new observation-based estimates of OHC in the context of several previous estimates (including those used by Gleckler et al).*"

3 In the Data and Methods, the authors used two approaches to map the OHC data. I think these approaches may be not perfect because the ocean state in data-rich areas is not the same to that in data-sparse regions. These approaches should be supplemented with the different ocean dynamic conditions.

Reply: We agree that the ocean state in data-rich regions (mainly in Northern Hemisphere) and data-sparse regions (mainly in Southern Hemisphere) are not the same. That is why we argue that some traditional methods those assuming equal OHC change in data-gaps with that in sampled area are probably biased and should be more carefully applied. That is also why we proposed two simple ways to solve this problem: one shown in CZ14 and the other in this study. Also we agree that the two approaches are not perfect, and they are very simple transparent methods as we discussed in the manuscript. There is no "perfect" method since one can never re-observe the ocean in the past. It is very important to develop a more comprehensive method to solve this problem, but this should be provided in another study and the method should be fully evaluated.

4 In the Data and Methods, the authors noted that "Because the upper 700-2000m oceans show an approximate tripling of the heating rate from 1992-2005 compared to 1970-1991 (as shown in Fig. 2, green curve), we assume a proportionate increase in heat uptake in the deep ocean (2000m-bottom)." I think the upper 700-2000m oceans

are controlled by the wind while the deep ocean (2000m-bottom) is controlled by the thermohaline circulation. The dynamic conditions are very different. So the assumption is not so appropriate. You can give some observation evidence for this assumption.

Reply: This is a good point, but the thermohaline circulation also depends on mixing from multiple sources. In fact it is now more commonly referred to as the MOC (meridional overturning circulation) because it is much more than thermal and haline effects, the winds are also very much involved. We added in the manuscript that "the upper 700-2000m oceans are controlled by the wind while the deep ocean (2000m-bottom) is controlled by the thermohaline circulation". The scale we used is a simple/empirical choice. The problem is we don't have sufficient observations below 2000m before 1990, we have to make some assumptions. We discussed the uncertainties induced by this assumption in the manuscript:" *We show that the difference of this lower and upper bound of the 700m-bottom OHC change is equal to ~13% (~10%) of the full-depth OHC change during 1970-1991 (1970-2005), which indicates the maximum error induced by this assumption.*" This size of uncertainty will not impact the key conclusion of this study.

5 In the Observation-based full-depth OHC estimates, the authors stated that "The OHC change after the two volcano eruptions is approximately assessed by subtracting the OHC one year before the eruption from the OHC in the second year after eruption" I think the ocean response to the volcano eruption is not so quick, you can remove the effects of volcano eruptions by longer time delay.

Reply: This is the best choice for assessing the impact of volcanoes. According to Church et al. 2005, there is a strong radiative forcing and heat budget decrease soon after the volcano eruptions (within one year). After two or more years, the natural variabilities (i.e. ENSO) and greenhouse gas forcing comes in and dominate the global EEI, so it experienced a faster recovery.

6 In the Observation-based full-depth OHC estimates, the authors noted that "There is also indication of substantial heat discharge from the upper 700m ocean following the extreme 1997-1998 El Niño event". What is the reason of substantial heat discharge from the upper 700m ocean following the extreme 1997-1998 El Niño event?

Reply: The heat loss by the global ocean after El Nino has been well-documented by studies such as Trenberth et al. 2002 and Balmaseda et al. 2013, I quote a discussion in Balmaseda et al. 2013 here." *there is an additional cooling episode following the huge 1997– 1998 El Niño event after 1998, which mainly affects the upper 700 m. The event led to a global warming of the atmosphere and made 1998 the warmest year on record to that point as heat came out of the ocean, largely through evaporative cooling [Trenberth et al., 2002].*"

7 In the Observation-based full-depth OHC estimates, the authors wrote that "Although the comparison between the observational and CMIP5 full-depth OHC results in an insignificant difference, CMIP5 models show a large spread (Figure 3, 4, 5)". What is the dynamics of large spread in these models? I suggest you can explain it in detail.

Reply: This is a good question, and this is definitely a hot topic in climate community recently, but it is not solved yet. Unfortunately, some of this is simply that some models are not very good, they may not conserve properly, their physical parameterizations are not state of the art, they do not replicate the past very well, and so forth. We just provided several possible reasons according to the recent literatures:" *The reasons why the models have large divergence are still an actively discussed issue. Frolicher et al., (2015) discussed the large range of model results and attributed a contribution of this to the differences in indirect aerosols. Additionally, CMIP5 has been missing post 2000 volcanic eruptions in these simulations as discussed in Glecker et al (2016), but this effect is shown to small and less than 0.1 W m^{-2} as indicated in Trenberth et al., (2014).".*