This document is the author’s response to review os-2016-105-RC2.

We will address the comments of Anonymous Referee #2 in the following fashion:

- Referee’s comment (bolded).
- Our response to the comment(s).
- Our change(s) to the manuscript.

1.

The authors do not look at uncertainty in the secular (non-cyclical) change from the model, rather looking at variability around a secular linear trend. The authors note (on page 2, top) ’the variability about the warming trend has yet to be reliably quantified due to the underlying uncertainties’. I am not sure the authors have done quite that here. In their breakdown of OHC, they represent secular change as a trend (deltaQ/deltat). Not much is said about how this is done, but secular change represented by a trend will unavoidably influence the OHC anomaly [if the trend is overestimated, anomaly will be biased negative, understimated, biased positive, etc]. The obvious solution would be to run the model without secular change - meaning without anthropogenic change. Then only cyclical natural (non-human induced) variability would contribute to the OHC anomaly. But the authors note that the model they use is ’run without variability in anthropogenic forcing’. Unless that constant forcing is zero, there will still be a secular change and a contamination of the OHC anomaly.

The referee’s comments are correct, but are unfortunately misdirected by our awkward writing to have formed a misconception on the referee’s part. We took ‘anthropogenic forcing’ to be related to *changes* in the atmospheric greenhouse gas concentration. Since the model was run without changes in greenhouse gas concentration, we took this to mean that there was no anthropogenic forcing in the model. This statement should be re-written as ‘run with constant forcings’, as there were no anthropogenic changes as well as no volcanoes, no changes in solar insolation over time, and other terms that could be considered ‘variable’ forcings. I.e., the greenhouse gas concentration is held fixed near the 1990 value. Note that even though the forcing is held constant, there may be some secular trend due to remaining model spin-up in some variables, thus a detrending is still required.

We will change the text on page 9, line 14, from (Orig.) to (Edit):
(Orig.) …, is run without variability in anthropogenic forcing, …
(Edit) …, is run without variability in the external forcings, …

2.
I also think the authors could use some better terminology in the paper. ... The authors use the word ’truth’ throughout to mean results from the geographically complete model run as opposed to the ’observed’ results from the geographically subsampled and objectively analyzed fields. They also occasionally refer to the ’real’ ocean. Not being a modeler, it is somewhat incongruous to hear the model results referred to as ’truth’. Maybe ’model truth’ or ’complete model’ juxtaposed against ’subsampled’, whereas the ’real’ ocean would be the ’truth’.

We understand the referee’s concern here as we have struggled with the terminology ourselves. We agree with the referee that the ‘model truth’, ‘observed model’, and ‘observed ocean’ are more sensible than the terms we used; ‘truth’, ‘observed’ and ‘real’.

We will change these terms throughout the manuscript.

3.
The authors define ’natural variability’ as ’the variability of the ocean not forced by multi-decadal or slower linear climate change and seasonality’. But there is no term for multi-decadal climate change in their OHC equation, except the anomaly term. It would make sense to equate more closely the term ’natural variability’ to the OHC anomaly, as the terms are used interchangeably in the paper.

The referee’s comment is correct. We originally defined OHC anomaly to be equivalent to the natural variability and understand the confusion in wording the natural variability as ‘the variability in the ocean not forced by multi-decadal or slower linear climate change in seasonality”. Since the record length in our case is 23 years, multi-decadal variability would be difficult to distinguish, but in general this isn’t necessarily true.

We will change the text on page 8, line 11, from (Orig.) to (Edit).
(Orig.) … to be the variability of the ocean not forced by multi-decadal or slower linear climate change and seasonality. …
(Edit.) … to be the variability of the ocean about the record-length seasonal cycle and linear trend. …

4.
The final term which could use amendment is ’ISAS13 observing strategy’. ISAS13 is a set of gridded fields of temperature and salinity. It does not have a set plan for sampling the ocean, an observing strategy. What ISAS13 has is an ’observed data distribution’ from which it calculates t and s fields.

We understand the referee's concern for the term 'ISAS13 observing strategy'. We wish to express the combination of the observed data distribution and objective analysis methodology used in the creation of the ISAS13 data set in as few words as possible, and we originally used the term 'ISAS13 observing strategy' without a thorough discussion of its meaning. Therefore, we will add a descriptive definition of what we mean by the 'ISAS13 observing strategy' and by the term 'observing strategy' in a general sense in Section 4.2.

We will change Section 4.2 to read the following:

ISAS13 is a set of gridded global fields of surface and subsurface ocean temperature and salinity on a 0.5x0.5 degree grid spanning the years 1990 through 2012 (Gaillard et al. 2016). ISAS13 is constructed by gathering the observations from a variety of sources, including Argo, (etc.) and applying an objective analysis (OA) procedure to generate the gridded global fields from the point observations. The OA uses the optimal interpolation technique presented by Bretherton et al. (1976) and depends on a mean reference (or first guess) field, a field of the expected variance about the mean field, and de-correlation length and time scales. In the construction of the ISAS13 data set, this OA uses the monthly means provided by the World Ocean Atlas 2005 (WOA05) monthly climatologies (Locarnini et al. 2006, Antonov et al. 2006), monthly variances estimated from the observations relative to these climatologies, and decorrelation time scales chosen to represent 1) the resolution of the Argo network and 2) the first-mode Rossby radius of deformation on each point of their chosen analysis grid.

While a strategy assumes an a-priori plan, we define an 'observing strategy' as the methodology employed to produce a global subsurface ocean data set, whether the observational meta-data was planned before-hand or not. In our case, the ISAS13 observing strategy is defined as the set of observation locations and times plus the OA used to generate the ISAS13 data set, even though the observations and OA parameters were available beforehand.

We will also change the last paragraph of Section 1 (starting on line 21) to the following:

We demonstrate the method by quantifying the uncertainty associated with the spatio-temporal variability of ocean observations and specific choice of the mapping method used to construct global gridded fields from these observations associated with the ISAS13 data set (Gaillard et al. 2016). Specifically, the uncertainties are quantified for the estimate of global ocean heat content (OHC) variability down to 700m between 1990 and 2013. This 'observing strategy' is applied to 37 independent model segments from the Community Climate System Model version 3.5 (CCSM3.5), from which the statistics of OHC variability will be compared between the 'observed model' segments and their corresponding 'true model' segments across a range of time scales. The application of the methodology is described in Section 4, and the results are discussed in Section 5.

5.

… There is little in the way of description of the ISAS13 field calculation. The first time ISAS13 is mentioned there needs to be a reference. This reference should include details of the objective analysis procedure and full details of the OHC calculations - including climatology used and XBT bias correction applied.

These comments are addressed in the response to the previous comment.

6.
The authors further note twice in the paper, including the concluding sentence, that varying mapping method (a term which should be defined) will not affect the results of the present work. But previous work has shown that the method used to extrapolate and smooth irregular data does have a large contribution to uncertainty. So, even if mapping method would not affect the results of this work (a speculation in the authors conclusion), it does have an impact on the uncertainty of the trend in OHC change. Likewise XBT bias correction. So saying that OHC anomaly uncertainty alone is low enough not to contaminate the calculated trend is not sufficient. Other sources of uncertainty should also be factored in, or if the OHC anomaly uncertainty is an uncertainty which incorporates these other uncertainties in some way, this should be explained in detail.

Here is the concluding sentence of the submitted manuscript:

"Different mapping methods do differ in errors and estimates of global ocean heat content, but the broad similarity of the various methods leads us to speculate that our finding that subannual variability is significantly contaminated by the observing strategy is robust, and that only interannual and longer timescales of global ocean heat content variability and trends can be skillfully measured."

In this statement, we do not mean to state that varying the mapping method will not affect the results of our work. What we mean to say is that the spatio-temporal sampling dominates the error between the 'observed model' and 'true model' global OHC estimates in our study, and changing the mapping method won't change this conclusion. This conclusion comes from the large errors seen in the pre-Argo era versus the Argo era. Since the mapping method is the same for both eras, but the spatio-temporal sampling is very different in the two eras, we conclude that the effect of sampling is larger than the effect of the mapping method on the global OHC estimate. However, if the Argo era is examined alone, it is impossible to determine from this work alone how large the choice of mapping method will affect the estimates of global OHC variability. The manuscript will be re-worded to make this distinction clear. Additionally, since the observations are taken in a 'perfect-model' scenario, choice of XBT bias correction (and other related bias/error corrections) is not something that can be assessed in this methodology.

We will change the concluding paragraph of the manuscript quoted above to the following:

From the large differences seen in the pre-Argo era relative to the Argo era, since the mapping method is the same for both eras, we find that the spatio-temporal variability in sampling is the largest source of error in the ISAS13 observing strategy. It is impossible to determine from this work alone how large an effect the choice of a mapping method will have on these errors, but the similarity amongst the objective analyses used in the cited literature leads us to speculate that this result will remain unaffected by the choice of mapping method. We therefore conclude that subannual variability is significantly contaminated by this observing strategy, and that only interannual and longer timescales of global ocean heat content variability and trends can be skillfully measured. While other works have come up with this same conclusion, this work is perhaps the first to quantify this measurement skill, as given previously.

7.

I would like to see more explanation of the uncertainty inflation factor. Is this a standard practice? Why is it statistically valid here? The authors note that ISAS13 variability is about 2.5 that of the model variability - but the model was run with steady anthropogenic forces which I would expect to dampen variability (at least secular variability). So the low variability may be due to the unrealistic forcing.

We do expect the model to have unrealistic variability, and as stated in the second paragraph of Section 4.4, page 9, lines 12-19, “the estimates found using this model are based on a necessarily less variable climate system”. In order to increase (‘inflate’) the model variance to that of the real-world ocean, we multiplied our uncertainties by the ratio of the observed ‘real world’ variance to the ‘observed model’ variance. This is not standard modeling practice, but without significant change to our experimental design (e.g., examining a higher resolution model and doing another comparison study), we wanted to suggest an estimate for the real world uncertainty. We will edit the manuscript to make it more clear to the reader that this is done only as a rough estimate of this estimate.

We will create a paragraph break after the sentence ending with “...shown in Figure 7.” in section 5.5 on line 3 of page 15. The following sentences will be added to the end of this paragraph:

It should be noted that this variance inflation method is not standard practice, but this method gives a rough estimate of real world uncertainty.

8.

… Figure 7 raises a number of questions, starting with the units. It is impossible (for me) to see the power of 10 for the y axis in the figure. This, and the caption and accompanying text noting ’(low frequency) observed global ocean heat content’ make it hard to know what exactly is being shown. Is this Qbar which includes mean OHC, a huge number in comparison with deltaQ/deltat and QsubL, or is this just the latter two terms? The uncertainty of the anomaly is necessarily very small compared to the entire

global OHC (as the anomaly is just a fraction of the global OHC), shouldnt just the deltaQ/deltat and QsubL be shown with the QsubL uncertainty?

The referee is correct, deltaQ/deltat and QsubL are much smaller than Qbar, and therefore it makes more sense to plot only QsubL + deltaQ/deltat * t to better visualize the magnitude of the temporal variability and uncertainties. This change will be made in the edited manuscript. The figure caption will also be changed to clarify what is being shown, per the referee's request.

We will change the caption of Figure 7 to the following:
(Orig.) Annual running mean of the ISAS13 observed global ocean heat content and one-sigma uncertainty estimated by the EOSSE.
(Edit) QsubL(tau=12) + deltaQ/deltat * t as calculated from the ISAS13 data set. The shaded region indicates the one-sigma uncertainty due to the ISAS13 observing strategy as estimated by the EOSSE methodology introduced in this work.

9.
I am also not sure how the jumps in year 1997 has anything to do with the change in observing system. There were no Argo floats in 1997, and only a handful of profiling floats. I would also like to see more explanation for the peak around 2004 and subsequent drop in relation to the change in observing system, since the observing system was already dominated by Argo in 2004.

We have emailed Dr. Tanguy Szekeley regarding this very concern. Here is the response we received addressing these jumps, copied here with some minor grammatical corrections:

The 1996 and 2003 jumps are associated to changes in the ocean sampling (see the PCTVAR field). The Indian ocean measurements almost disappear in 1996 leading to a decrease of the solution accuracy. In 2003, the begin of the worldwide ARGO deployment change the solution in numerous badly sampled zones (pacific ocean, South Atlantic Ocean, Antarctic Ocean, etc....)
This problem will be solved in the next version of CORA (to be released on April 2016, a paper with details on the dataset description and validation should been published then).

Due to recent release of the new data set, we were unable to re-run our analyses using the updated CORA data set (the computations were originally run in Fall 2015 through Winter 2016). We will update the manuscript to state this, as it is currently unknown whether the new CORA dataset does or does not contain these jumps at ~1996 and ~2003.

Together with the comments given in (**6.**), the last paragraph of Section 5.5, page 15, beginning on line 7, will be changed to the following:

One sees that after smoothing with an annual running mean, the uncertainties introduced by the observing system are small compared to the low-frequency variability and trend. We note that there may be uncorrected biases between pre-Argo and Argo instruments which produce the jumps seen at 1997 and 2003 which are an additional source of uncertainty. According to personal communications with Dr. Tanguy Szekeley, the curator of the CORA data set, these jumps and other problems have been resolved in the newest version of CORA referenced by Gaillard et al. (2016).

10.

Can the authors please add more explanation to the text and figures for Section 5.5?

This has been addressed in (**7. - 9.**), but we have edited the text slightly in multiple places in Section 5.5, and have changed the caption of Figure 8 to the following:
(Orig.) Annual running mean of the ISAS13 observed global ocean heat content and one-sigma uncertainty estimated by the EOSSE.
(Edit) Annual running mean of the ISAS13 observed global ocean heat content (solid line) and its estimated one-sigma uncertainty bounds (shaded region). The uncertainties are estimated separately for the pre-Argo and Argo era's by taking the uncertainties derived from the EOSSE (see Figure 7) and multiplying them by the ratio of the variance in the ISAS13 Argo-era data to the variance of the full 'model truth'; 2.6.

11.

There is an extra 'non' in 'non non-uniform' at line 6 on the first page. There is a missing 'a' in 'anomlies on line 5 of page 15.

We thank the reviewer for pointing out these errors. They have been corrected in the updated manuscript.

12.

Figure 1 has incorrect labeling. Blue line should be profiling floats. In addition the counts for other instrument types are too low. A quick look shows there were more than 80,000 XBTs dropped in 1990 for instance, an average of more than 6,000/month.

This error was addressed in Author Comment #1 and will be implemented in the revised manuscript.