

Dear anonymous referee #2,

thank you for your careful reading and for the detailed comments on our paper. They were very useful to improve our work. In addition to your requests for specific corrections (see point-by-point response below) as well as those from referee #1, we have tried to improve the wording and clarity of the paper.

Yours sincerely,

Karina von Schuckmann and Pierre-Yves Le Traon

Point-by-point response (bold: referee #2 comment):

1) The second line tells that the floats provide a good coverage down to 2000 m depth. Line 15, however, indicate Argo sampling to extend between 10 to 1500 m.

We changed the first sentence of the abstract by: “Argo deployments began in the year 2000 and by November 2007 the array reached its initial goal of 3000 floats operating worldwide.”

We have also modified the sentence in Ln.15/16 by “Long-term trends (15 years) of GOIs based on the complete Argo sampling for the upper 1500m depth can be estimated with an accuracy of ± 0.03 mm/yr for GSSL, ± 0.02 W/m² for GOHC and ± 20 km³/yr for GOFC – under the assumption that no systematic errors remain in the observing system.”

Global GOI time series based on Argo are still short, and as mentioned, Argo achieved its initial target not a long time ago. During 2005, about 45% of all Argo floats were measuring until 2000m depth. This number has increased during the last six years and about 60% of all floats measure down to 2000m depth. Since we start our analysis in the year 2005, we decided to use all measurements until 1500m depth to guarantee maximal available data coverage for our estimation. This is now explicitly mentioned in the paper (see also answer to comment 16).

2) Line 15: long-term trends can be performed: I think it is the calculation of the trends what is really meant here, since the trends are not performed but rather calculated/estimated.

We replaced 'performed' in Ln. 16 by 'estimated'.

3) Line 17: steric rise better change “to ocean level rise (steric component)”.

'Steric rise' is replaced by 'GSSL'.

4) Line 8: “They include a proper estimation ..: better change to “are accompanied by error estimates”

Corrected.

5) Here again (line 19) the depth range down to 2000 m is indicated. Few lines later (page 3, Line 4) the statement is made about Argo floats measuring AT LEAST within the upper 2000 m. Does it mean some floats can measure deeper? Or do you mean possible technology development in the future??

Corrected, see comment 6) just below.

6) Line 6: the fresh-water content has NOT been neglected, and there are papers (Wijffels et al.??) which must be cited here.

The two sentences “Argo provides the capability to assess global ocean heat content (OHC) by measuring subsurface in situ temperature, at least for the upper 2000 m depth. Moreover, the effect of internal global ocean salinity changes can be discussed which had been mostly neglected in previous global analyses due to a lack of large scale direct subsurface salinity observations.” were

changed by

“Argo provides the capability to assess global ocean heat content (GOHC) by measuring subsurface temperature. Salinity data allow an estimate of global ocean freshwater content (GOFC).”

Papers are not cited at this point as a more detailed discussion on salinity/freshwater changes can be found below, page 1002, Ln. 18- page 1003, Ln. 5.

7) Line 10: “steric component of the levels rise ... is a cause of global SL change. It’s better to say that the warming of the ocean leads to the SL increase due to the respective change in the sea water density.

The sentence: “This, in turn, provides the capability to understand global sea level change by evaluating its steric component which is one of the major causes of global mean sea level changes (Cazenave et al., 2009; Bindoff et al., 2007). Estimations of sea level changes are of considerable interest because of its potential impact on human populations living in coastal regions and on islands (>50%).”

is replaced by:

“This, in turn, provides the capability to understand major contributions to global sea level change (Cazenave and Llovel, 2010; Church and White, 2011), i.e. global steric sea level (GSSL).“

Explications of volume effect due to thermal and haline changes are introduced below, page 1001, Ln. 20- page 1002, Ln.3.

8) Lines 12 to 19: the text could be omitted

We believe we should keep these three sentences as they underline the importance of error handling and the aim to improve accuracy of GOI estimations.

9) Line 20: To my mind the SL rise is (partly) driven through the continental ice melt water input into the global ocean, e.g. the water mass increases leading to the level rise. This additional fresh water dilute the salty ocean water: salinity decreases.

Lines 23-24: bad wording. It’s better to say that satellites (which measure the total SL rise signal) help define the SL rise component linked to the input of melt water into the ocean, since this component is NOT measured by argo floats.

We improved the wording:

“Global total sea level derived from satellite altimetry can be partitioned into its steric and mass-related components (e.g. Cazenave et al, 2009, Leuliette and Miller, 2009). Steric sea level is driven by volume changes through ocean salinity (halosteric) and ocean temperature (thermosteric) effects, from which the latter is known to play a dominant role in observed contemporary rise of GSSL.”

10) Page 4, line 5: “the underlying uncertainties... are still unclear” – bad style. The uncertainties arise because smth. is unclear!

We have removed this sentence.

11) Lines 9 -11: The sense of the sentence is not clear. The same with the next sentence, starting with “But nevertheless, ...”. Please ask colleagues profound in English for help!!!

We changed the wording to:

“Differences have been explained by the uncertainty due to the choice of data processing methods (including corrections of instrumental biases) as well as to effects of interannual variability (Domingues et al., 2008; Lyman et al., 2010; Gouretski and Reseghetti, 2010). Recent estimations of GOHC are mostly based on Argo measurements which reduces possible errors due to large data gaps in space and time as well as due to inhomogeneous sampling. Nevertheless, analyses of GOHC during the last decade differ as well among methods (von Schuckmann et al., 2009; Willis et al., 2009; Trenberth and Fasullo, 2010).”

12) **Page 4, lines 18-19. Please, rewrite, otherwise the sentence bears an imprint of tautology: “changes ... induce fluctuations...” (the latter are changes as well!)**

The overall discussions on the issue why the OFC is important (Page 4 Line 18 to Page 5, line 5) should be rewritten. For instance, the thesis “While impacts are local and regional, the causes and patterns are global” needs more discussion, or should be omitted.

We rewrote this paragraph:

“The estimation of GOFC from Argo as a salinity anomaly over a depth layer (e.g. Boyer et al., 2007) is an indirect but potentially sensitive indicator for detecting changes in precipitation, evaporation, river runoff and ice melt (sea ice, continental glaciers and ice sheets). Any change in the hydrological cycle is reflected in the ocean salinity field due to the large-scale balance between the surface freshwater flux (evaporation minus precipitation and continental run-off), salinity variations and the ocean's advective and mixing processes (Durack and Wijffels, 2010). Multi-decadal trends in ocean salinity have been observed on global and regional scales (e.g. Antonov et al., 2002; Boyer et al., 2005; Delcroix et al., 2007). These multi-decadal salinity changes appear to coincide with both, broad-scale surface warming and the amplification of the hydrological cycle (Durack and Wijffels, 2010). Results shown in von Schuckmann et al. (2009) document that ocean salinity and hence freshwater are changing on gyre and basin scales and GOFC is characterized by large interannual changes rather than by a significant trend during the last decade.”

13) **Page 5, line 6: “These discrepancies show...” It is not clear which discrepancies are meant!!**

We changed the wording of the last paragraph of the introduction:

“Thus, while Argo provides data with unprecedented accuracy and coverage, estimating GOIs remains a major challenge. It requires very careful data quality control and analysis. An estimation of errors is needed for a sound interpretation of results.”

14) **Line 9 change “refined error estimates” to “error estimates” Line 11: change “A careful discussion..” to “a discussion”**

Corrected.

Data sets and methods

15) **Line 22: the datasets are processed by processing tool “ISAS-STD” – bad wording. Instead of reporting the name of the software product, it is much more important to know how it works, namely: 1) what is a climatological test? 2) how the observed values are interpolated on standard levels?**

We added a citation to this sentence: “[see von Schuckmann et al., 2009 for more details].” A detailed description of the data processing method is given in this paper, and we have decided not to repeat it here.

16) **Line 27: It was already mentioned in the previous paragraph, that the data for 2005-2010 were used – please, avoid such repetitions! Lines 27ff: what does this compromise exactly mean? Why numerous argo profiles are shallower than 2000 m?**

Due to technological evolutions, first Argo floats were able to measure down to 1000m or 1500m depth. Some of these floats are still working and will be replaced by improved Argo floats measuring down to 2000m depth. Hence, the data coverage below 1500m depth increases with time, and in the future, the calculation of GOIs should be based on measurements down to 2000m depth. As we started our calculation in a period where only 40% (for the year 2005) of all Argo measurements

went up to 2000m depth, we decided to base our estimations on the upper 1500m depth. This is now better explained: “Only 40% of Argo floats provided measurements up to 2000m depth in the year 2005. The data coverage is thus not sufficient for a global estimation of changes between 1500m and 2000m depth [the situation has improved and thanks to technological evolutions 60% of Argo floats provided observations up to 2000m in 2010].“
Moreover, we have removed the repetition ‘2005-2010’.

17) Page 6, Line 10: the gridded in situ product ARIVO – here, again, the name (what is not important) is provided, but the essential details are completely absent. I assume “the gridded in situ product” simply means “gridded climatological fields of T and S – is that true? – the description is totally misleading. It is said, the gridded product is used to extract climatology? Does it mean the “gridded product” is a software tool, which allows to read in a climatology??
The gridded product ARIVO (as described in von Schuckmann et al., 2009) is used to calculate an Argo climatology. To avoid any misleading interpretation, we changed the sentence by: “An Argo climatology (ACLIM hereinafter, 2004-2009, von Schuckmann et al., 2009) is interpolated on every profile position in order to fill gappy profiles at depth...”

18) Further, the argumentation is missing completely WHY the climatology is used to fill the gaps. Using climatology would obviously lead to T or S-anomalies being simply zero? Than what is the motivation here?

As described in the following sentence, OHC, SSL and OFC are calculated from temperature and salinity profiles where gaps are filled with a climatology at depth. Complete profiles are needed to calculate integrated quantities. The OHC, SSL and OFC anomalies are then calculated, which means that ‘filling gaps at depth’ is sensitive to the choice of filling, i.e. a climatology (ACLIM or WOA05). We tested this sensitivity and since the effect was small, we decided to use the ACLIM for this purpose.

We added one sentence: “This procedure is necessary to calculate depth-integrated quantities.”

19) Testing the results against different climatologies is justified. However, the description of the results is unsatisfactory. For instance, I disagree that the differences between the two cases (ACLIM and WOD05 climatologies) are larger at the beginning of the time series: it is simply not seen in the fig. 2!!! Moreover, the statement about the differences being largest for GSSL is simply wrong: how could different variables be compared? Is, say, a 10mm SL difference larger than 0.5 degree C difference???

We changed this paragraph: “Results are presented in Figure 2 for all three parameters. The sensitivity of GOIs with respect to the choice of the climatology is generally small, but is not negligible. Computations of OFC, in particular, are sensitive to the reference climatology [see Boyer et al. (2007) for more details on the freshwater calculation]. Based on this sensitivity test, a climatology uncertainty for each GOI is included in the error estimation as discussed later in section 3.1.”

20) Page 7, line 1: “the variance information to build this criterion ...” needs more explanation. On the previous page the authors mention that a 3-sigma limits have been selected – why? Or is it a subjective decision?

We used 3-sigma as it is a standard method to remove extreme values.

21) Page 8, Line 5: “The GOIs are evaluated from the horizontal data distribution” - I do not understand, what distribution is meant here. Line 6 domain involves the effective coverage – bad wording, please, rewrite.

We rewrote this sentence: “GOIs are calculated within 60°S-60°N, i.e. the effective coverage of the Argo array (Roemmich and Gilson, 2009).”

22) Line 7: “Mean estimations of physical parameters” – what is meant here? Is it the estimation of the mean value???

Page 8 Line 13ff: the usage of the weighting matrix W is not sufficiently justified. Why the anomalies are (much) larger near the coast? Is it an artifact of the averaging within large boxes, or is the variability higher there? What happens (with respect to the time series) if the coastal boxes are omitted??

We changed this paragraph as our description indeed was unclear:

“The global mean indicator GOI(t) is obtained by averaging $\overline{\Theta}_{box}$ estimations weighted their surface area $M_{i,j}$:

$$GOI(t) = \frac{\sum_{i,j} \overline{\Theta}_{box} M_{i,j}}{\sum_{i,j} M_{i,j}} .$$

Error estimation

23) Page 9, Line 1: redundancy again: please, remove two sentences starting with “Using the box-averaged: ...”

Done.

24) Lines 5ff: I was pleased to see the comparison of different argo-based products. However, here again the description of the inter-comparison procedure is unsatisfactory. What is shown in Fig. 3 are three steric height time series. Were these time series obtained by the authors of the submitted manuscript? If yes, how were they calculated?

I disagree that the largest deviations are observed in the beginning of the time series, as the year 2007 exhibits the largest discrepancies.

We modified this paragraph: “Large sensitivities of a GOI like GSSL to different data processing techniques are obvious when comparing different products of gridded Argo fields (Figure 3). We calculate GSSL from three different products which are downloaded from the Argo web-page (www.argo.ucsd.edu), i.e. two products based on Argo and other hydrographic data (ARIVO delivered by Ifremer, and MOAA delivered by JAMSTEC) and one product including Argo only measurements (delivered by Srippls Institution of Oceanography). Detailed information on the gridded fields can be found on the Argo webpage. We chose to evaluate the comparison during the time period 2004 to 2008 for consistency. Amplitudes of interannual fluctuations differ from one product to another (Figure 3). Although the evaluation of GSSL in Figure 3 ...”

25) Page 10, line 8. “... the climatology used to fill vertical gaps and to evaluate the anomaly fields” . To my understanding the climatology provides the reference against which the anomalies are calculated. I do not understand how the climatology is used “to evaluate” the anomalies.

We changed the sentence to: “... the climatology used to fill vertical gaps and to calculate the anomaly fields”.

26) Page 10, lines 11-14. It is not clear to me what residuals are meant which were used to estimate E-clim. Or is it the RMS difference between the two time series which provides a measure for E-clim???

We changed this sentence to: “To estimate the value for Eclim, the standard deviation of the difference of the two time series ...”

27) Page 11, line 1. I think, earlier in the text a different year was mentioned for the time moment since then the coverage was complete.

We removed “where Argo sampling was complete”.

28) Page 11, lines 10ff. I recommend delete from “This approach allows : : :” to the end of Line 17, as all the information here is redundant.

Done.

29) Lines 17-19 “A forecast calculation ... has been established” – how the calculation can be established??? Line 18: here, the same redundancy again: the reader has already been informed many times on when the argo array has achieved its complete shape.

We changed the sentence to: “A ‘forecast calculation’ of the uncertainties of global trend estimations is given in Table 2 assuming GOI error bars during the year 2010 while applying Eq. (A1) of the Appendix.”

30) Page 12, lines 1-4 I do not understand what they want to tell here.

The explanation is now better given by: “Note that our estimations provide an estimation of errors on the trend over a given period. Such trends even if they are statistically significant cannot be interpreted as long term climate trends as they also include the effect of interannual signals. This is clearly the case for the OFC trend.”

Method validation (changed to “Testing the method”)

31) Page 12, lines 6-9. Bad formulation: “Altimeter observations ... correlate with in situ upper ocean observations” Observations do not correlate: the observed parameters may be correlated. “For this purpose, maps of: ...” Which purpose is meant???

We have changed the two sentences to: “Altimeter sea level observations are a useful and nearly global observational record over the ice-free oceans that have been shown to be correlated with in situ SSL and OHC (e.g....). Maps of mean sea level anomalies ...”

32) Line 11: what kind of “in situ estimations” do the authors mean??

The sentence has been changed to: “Using high-resolution altimeter measurements has a proxy for GOI estimations based on in situ data (SSL and OHC) has already been performed in previous studies (...).”

33) Lines 22-24: I suggest to delete the sentence starting with “However, using ...”.

Done.

34) Line 25: “The comparison between the two global averages calculated in DIFFEREWNT ways: : :” Just a few lines earlier the authors note, that for the calculation of the MSLA from the gridded altimeter fields the SAME method as described in section 2, w3as used – this is an obvious inconsistency.

“... calculated in different ways ...” has been removed.

35) The following description of the two MSLA time series is unsatisfactory. The authors note differences between the two curves in “high-frequency variability”. The scale of this variability is not defined, but the highest frequency variability as resolved by the curves in Fig. 4 is rather similar (to my mind) for both curves, and it is for the longer period variability where

the two curves differ. Moreover, the Fig. 4 indicates an offset between the ANOMALY curves. After deletion of this offset the agreement would significantly improve.

Page 13, line 2. Here again the year is indicated when the sampling becomes complete: redundancy!!!

Thanks to your and reviewer's #1 comment we realized that we had submitted an erroneous time series (bold line in Figure 4, Global mean MSLA from AVISO, see resubmitted Figure 4).

We changed the description of Figure 4 in the text as well: "The comparison between the two global averages shows reasonable agreement and their 6 year trends are consistent (Figure 4). There are differences in annual and lower period variability among the curves. This testing shows that our simple box averaging method depicts global mean changes reasonably well and can be used to assess GOIs for monitoring needs of the climate system."

36) Page 13, Lines 8-10. "The calculation ...has been chosen to represent..." WRONG: the three globally averaged characteristics of the ocean state were selected to represent the so-called GOIs!!!

The description which follows is completely unsatisfactory and should be rewritten. For instance, they start with the GSSL (Line 11), but then start to discuss the errors, which obviously decrease with the improvement in the argo array not only for the GSSL, but for all indicators under consideration. Moreover, the decrease in the error magnitude with time is barely visible in Fig. 5 (by the way, the footnotes a), b) and c) are absent in the Figure). The new values of the OHCA should be compared not only with those by Schuckmann et al. 2009, but with several other estimates available in the literature.

Line 21: I do NOT agree that the inter-annual variations of GSSL and OHC are smaller in amplitude compared to the long-term variability. Here, again is not clear, what the long-term scale is. Should it be the whole analyzed period 2004-08, than the typical amplitude of the inter-annual variability is not smaller. Moreover, considering just the size of the error bars, the long-term change in GSSL and OHC is also not significant.

We have rewritten this paragraph for a better explanation: "In this section, GSSL, GOHC and GOFC are derived from Argo data using the box averaging method discussed in section 2 (Figure 5). The 6-year trend estimations are calculated as described in section 3.2. Generally, the error decreases as the number of measurements increases and hence, the GOI errors decrease with time (Table 1).

A significant increase in steric sea level can be observed from 2005 to 2010 with a trend of 0.69 ± 0.14 mm/year (0.48 mm/year for the Earth's entire surface area, Figure 5a). For this decade, different values for global upper ocean thermal expansion have been estimated from in situ data ranging from -0.5 to 0.8 mm/year (e.g. Willis et al., 2008, Cazenave et al., 2009; Leuliette and Miller, 2009, von Schuckmann et al., 2009). The GOHC estimation shows a significant 6-year increase with a rate of 0.55 ± 0.1 W/m² (0.39 W/m² for the Earth's entire surface area, Figure 5b). Our GOHC estimation is slightly lower compared to the composite evaluated by Lyman et al. (2010) and to what was found in our earlier study (von Schuckmann et al., 2009). This can be due to the fact that the later period is confined to a period when the upper layers did not seem to be gaining much heat (e.g. Levitus et al., 2009, Cazenave and Llovel, 2010, Lyman et al., 2010). However, comparisons of GOHC and GSSL values from different studies are difficult to interpret as differences include the effect of interannual variability, instrumental biases or different data processing methods.

One important source of uncertainties in GOI estimations are due to the fact that patterns of interannual variability differ among estimation methods (Domingues et al., 2008; Lyman et al., 2010). Our results show that interannual fluctuations of GSSL and GOHC exist but are small (Figure 5a and b). This is different for GOFC (Figure 5c). Large interannual fluctuations dominate the time series, and

the trend is very much dependent of these large interannual fluctuations. This implies that a longer time series is needed to extract a significant long-term trend of GOFC.”

We have also added labels a)-c) in the Figure.

Conclusions

37) It is not clear for me, why the attention is drawn only to the paper by Lyman et al 2010 (lines 5-7, page 14).

We have changed this: “GSSL, GOHC and GOFC from in situ observations are a useful benchmark for ocean and climate models and an important diagnostic for changes in the Earth's climate system (Hansen et al. 2005; Levitus et al. 2005). Differences among various analyses and inconsistencies with other observations (e.g. altimetry, GRACE, Earth energy budget) requires particular attention (Hansen et al., 2005; Willis et al., 2008; Domingues et al., 2008; Cazenave and Llovel, 2010; Trenberth, 2010, Lyman et al., 2010).”

38) Line 14: there is a significant inter-annual global variability at global scale – oh, what is it!!!

Line 15ff: actually, it is not shown in the paper, that the short-term trends can not be estimated. Again, what does it mean “short-term”???
Line 15ff: actually, it is not shown in the paper, that the short-term trends can not be estimated. Again, what does it mean “short-term”???

Line 17: “uncertainties due to inter-annual fluctuations are not included in our error estimation” –completely unclear, what do they want to tell here. In the next sentence they say, “this (WHAT???) will change with the growing set of Argo measurements”

We changed this:

“Moreover, our trend estimations are estimated over a 6-year time series only and are affected by interannual variability. Hence, an interpretation in terms of long-term climate signals remains questionable. ”

39) Line 22: please, do not pretend to say your method is a proper one

Corrected.

40) Line 26: “uncertainty estimations due to the data handling”: I guess what they really mean here was something like “ estimation of uncertainties arising due to the method which was used”

Corrected.

41) Line 27: of course, it is trivial, that only the period 2005-2010 has been analyzed, and that the trends for this period may differ when calculated for the other periods.

We have removed the text in the parenthesis.

42) Page 15, last sentence before the Appendix should be deleted.

We have deleted this sentence.

Figures

43) Fig.2, page 23: please, insert labels a), b), c) for the respective panels and introduce changes in the main text. The sense of the last sentence in the figure caption is not clear: “... for the choice of reference climatology... two different climatologies are used”. Nowhere in the text is indicated which climatology has been chosen.

The figure caption now reads "..., i.e. ACLIM (green) or WOA05 (blue)". We added "see text (section 2) for more details". We also added labels a)-c) in the Figure.

44) Fig. 3 I guess what is really show here is the globally averaged steric height ANOMALY.
Yes, this is corrected.

45) Fig. 4 Method validation... I guess they wanted to say "Validation method". I do not agree with this terminology. Usually in situ observations are used to validate the results based on the proxy data (in this case it was the satellite altimetry). It is better to speak about the comparison of the two time series, one of which was derived from the satellite altimetric observations.
Obviously, the terminology leads to misunderstanding. Therefore we replaced "Method validation" by "Testing the method". Our aim is not to compare in situ and satellite time series. It is rather to test our box averaging method using satellite altimetry data. Altimeter data are available on a gridded field and hence, they are ideal to apply our method on this data set and compare the subsampled (on Argo position) results to the gridded field results.

46) Fig.5 "Revise estimation of ... of" I suggest to change to "Estimates of..." This comment is also relevant to the main text, as it is not clear which estimates have been revised in the submitted manuscript. Lease add literals a,b,c to the respective panels.
We removed "revised" in the caption and in the text. We also added labels a)-c) in the Figure.