Response to Reviewer 1

We thank both reviewers for the careful reading and suggestions on how to improve the manuscript, and in particular Reviewer 1 for pointing out ways to add more discussion in the analysis of results, further to many technical corrections. Below, we answer in details to all the points arisen from Reviewer 1. Sentences in italic are the reviewer's comments, while our reply is in bold, and includes the reply and the proposed modifications to the revised version of the manuscript.

It is observed that there exist similar global reanalysis datasets, e.g., in in marine.copernicus.eu, ECCO or US HYCOM etc. Many components of the method description have also been published by the authors in other papers. However, these products haven't been inter-compared with the V5 data, either qualitatively or quantitatively. This makes it difficult to evaluate the "state-of-the-art" and "uniqueness" of the data.

We thank the reviewer for this comment. In the framework of the MyOcean project, our V4 reanalysis was extensively compared with the other three reanalyses produced in the framework of the project (Mercator-Ocean/GLORYS2, ECMWF/ORAP5, UniReading/UR4) and the main outcomes are included in Masina et al., (2015, Climate Dynamics). Within that intercomparison, it turned out the C-GLORSv4 is a state-of-the-art reanalysis for all parameters cross-compared (SST, temperature in the 0-800, and 0-2000 layers, sea-ice concentration, AMOC, volume transports in selected WOCE cross-sections) and also salinity content that however was not included in that publication. This cross-comparison allowed us to use C-GLORSv4 as the starting point to compare C-GLORSv5 with. Unfortunately, it is not straightforward and beyond the scope of this work an in-depth cross-comparison of our reanalysis with other global ocean reanalysis.

To account for this comment, we have modified the Introduction to explicitly reflect that C-GLORSv4 was already included in the MyOcean cross-comparison and proves a state-of-the-art reanalysis. We however have removed the reference to "state-of-the-art reanalysis" in the Abstract (replaced by "latest") to avoid any possible subjective definition.

The new data (V5) have been validated and compared with the old data (V4), which shows improvements in simulating variability of the sea ice and AMOC. However some features of the products have been degraded. Overall T/S validation in Fig.5 shows that only water temperature in upper 80m in V5 has smaller RSME than V4, T/S in other layers for V5 are worse than V4. The global SST comparison also show significant signals in several areas of degradation (Fig. 6). These results have not been analyzed in the paper for details.

We thank the reviewer for pointing this out and we modified part of the text in the Abstract, Results and Summary sections to explicitly state the metrics that get worse in V5 and justify them whenever is possible. Our

experience suggests that when a consolidated product is upgraded, it is difficult to have all diagnostics that show improvements; nevertheless we believe that V5 contains important improvements and it is at the moment our recommended reanalysis product, which is also being widely used for a large number of applications and from many users worldwide.

Detailed comments

Acronyms: There are many acronyms not given their full name when first time shown, eg EN3, EN4, MDT, SLA, SSH, DMSP, RMSE, PIOMAS, OAFlux, ISCCP etc.

Thanks for pointing this out; we have explicitly added the meanings of all acronyms.

P1, L10: "a state-of-the-art ocean reanalysis". Reviewer: as the article does not perform any cross-validation with other reanalysis products from different systems, it is very hard to say it is a "a state-of-the-art ocean reanalysis".

Please see the answer above to the general comment. The cross-comparison of v4 with other reanalyses (Masina et al., 2015, Climate Dynamics) is now explicitly mentioned, while we replaced "state-of-the-art" with "latest".

P1, L18-19: "the new reanalysis outperforms the previous version, especially in representing the variability of global heat content and associated steric sea level, the upper ocean temperature and the thermohaline circulation." Reviewer: the presentation can be more precise: "the thermohaline circulation" is AMOC, the "upper ocean temperature" is "slightly improvements in upper 80m but worse below for water temperature, and salinity is consistently worse in the upper 500m".

Please see the answer above to the general comment. Strengths and weaknesses of the new reanalysis have been precisely stated in the revised version of the manuscript, as the Reviewer suggests.

P1, L23: "a ocean" -> an ocean

Corrected

P3, L29: "OcenVar" -> OceanVar

Corrected

P4, L4: describe what is x and xb

We have added their description.

P4, L10-15: VV and Vv shall make the same.

Thanks for noting this; we have corrected it (all operators have lowercase index)

P6, L28: in the equation (5), ð˙IZ˙ij has been used in equation (4), but are different parameters.

Thanks for noting this; to avoid confusion the "threshold" of the background quality check is now indicated with the "gamma" Greek letter.

P6, L32: "cost" -> coast

Corrected

P8, L20-21: the analysis on Fig.5 results should more precise and detailed, notifying that temperature is only slightly better in upper 80m but not 100m, maybe it's good to give a quantitative value of how much the temperature and salinity are better or worse in different levels.

We thank the Reviewer for pointing this out: as mentioned in the answer to the general comment, we have precisely indicated the vertical range of improvement/worsening. According to Reviewer 2, we have also added panels on the skill score in the Tropical Pacific and North Atlantic Oceans.

P8, L28: while there is a 3.2% decrease of RMSE for global SST, there exist significant degrade in Gulf Stream, Kuroshio extension and circum-Antarctic ocean. It should also be mentioned is the accuracy of the NOAA SST, ie. about 0.6C.

We have added a discussion on the significant degradation in the regions mentioned by the Reviewer, which may be explained with the increased background-error variance in those areas (see also the Response to Reviewer 2). This would lead to heavier weights given to assimilated observations in the reanalysis, resulting in a worsening of the skill score in the validation against NOAA SST in areas with large variability. In the revised version, we also mention the accuracy of NOAA SST as suggested by Reviewer 1.

P9, L9-10: "C-GLORV5 data start in 1980 unlike C-GLORSv4 (1982)" should be changed to "C-GLORV5 data start in 1980 unlike C-GLORSv4 in 1982"

Corrected

P9, L10-15: it is suggested to use longer period for comparison the trend, e.g. 1982- 2013 where NODC, V4 and V5 all have data. Use 2003-2011 for trend inter-comparison may be affected by the statistical significance due to very small number of samples.

We have added also estimates for the trend within the period 1982-2013, for comparison with NODC estimates.

P9, L18-25: Figure 8 is not convincing to show that V5 has a better heat content trend han V4. Fig. 8c has 4 areas with large differences from V4, i.e., N. Atlantic including Gulf Stream, Kuroshio extension, South China Sea and circum-Antarctic ocean. Considering results from Fig.5 and Fig. 6, I would not conclude that these areas with significant differences are improved. The evaluation should reflect

both significant sig- nals of strength and weakness.

We have now better discussed the strengths and weaknesses: while skill score results seem worse in the Gulf Stream area, here the increase of heat content trend in the middle of North Atlantic gyre drives the sustained global signal in the last decade (see also the Response to Reviewer 2).

P9, L28: "RAPID-MOC" -> should this be RAPID-MOCHA?

We prefer to keep distinguished the sources for the volume transports and overturning (RAPID-MOC) and that for the meridional heat transport (RAPID-MOCHA), according to the fact that observational data belong to the two projects, respectively.

P10, L5: 26N -> 26oN

We have inserted the "degree symbol" in all occurrences

P10, L21: "by a realistic" -> by a more realistic

Corrected

P11, L3-4: "Based on this assessment, C-GLORSv5 proves a reliable tool for investigating the ocean and sea-ice interannual variability in polar regions." Reviewer: I am not sure if this statement holds. In terms of ice volume, V5 has more reasonable results than V4. But it is hard to justify that V5 is a reliable tool for investigating the ocean and sea ice interannual variability in polar regions. Even PIOMAS is just another model. Regarding to real interannual variability in ice volume, one cannot say much due to lack of data.

This has been shown by a recent study by Mayer et al. (2016, GRL) that makes use of C-GLORS data to investigate energy budget in the Arctic Ocean. We have softened the sentence and added the reference and a corresponding sentence. The new paragraph reads:

Based on this assessment, C-GLORSv5 may be used as a tool for investigating the ocean and sea-ice interannual variability in polar regions. For instance, Mayer et al. (2016) make extensive use of C-GLORSv5 data (ocean heat content, sea-ice concentration and thickness, sea-ice velocities) to investigate the Arctic region energy imbalance.

P11, L7-9: it would be good to have a reference here.

We have added the previous reanalysis paper (Storto et al., 2016a), and the intercomparison work (Valdivieso et al., 2015) that uses the same forcing configuration used in C-GLORSv4.

P12, L10-12: the summary here on V5's quality on T/S should be more precise, and reflect results from Fig. 5 and Fig.6.

Please see the answer above to the general comment. Strengths and

weaknesses of the new reanalysis have been precisely stated in the revised version of the manuscript.

P22, Fig. 3: title of vertical axis is missing

Added in the revised version.

P23, Fig. 4: the legend "180d": is this wrong? In the text it says 3 months, i.e., 90days.

Thanks for pointing this out, it was mistakenly reported as 180d, although it referred to 90d, i.e. 3 months. We have corrected the figure legend.

P25, Fig. 6: the Caption should be rewritten

We have rephrased the Caption: Map of differences of SST Root Mean Square error between C-GLORSv4 and C-GLORSv5. The RMSE is computed against NOAA SST ¼ daily analyses (Reynolds et al., 2007).

P26, Fig. 7: the correlation coefficients can be removed, as they are not explained and used in the text. Showing correlation generates a couple of issues: i) have the trends been removed before calculating the correlation? ii) the number of samples used and significant level of correlation

We thank the Reviewer for pointing this out: we have now mentioned the correlation in the text, reporting that the correlation is significant at 99% (bootstrapping) when greater than 0.25 for the 108 samples (monthly values from 2003 to 2011) considered here. We prefer to use this metrics rather than the correlation w.r.t. to NODC, because the latter provides only pentadal (5-year) means and it is strongly affected by the climatology during the first 10 years of the reanalysis period.

P28, Fig. 9 lower panel: the legends need to be corrected, "Sv"->PW .

Corrected