<u>Response anonymous reviewer #1:</u>

We would like to thank the anonymous reviewer #1 for his/her wise comments that helped us to improve the manuscript**.**

General comment

The paper has huge language problems! Some should have been addressed before submission, e.g. obvious spelling errors in headings and legends. But even then a native English speaker should have edited the paper before submission.

We have checked the spelling and grammar, and hope that the manuscript will be easier to read by an English native speaker.

On top of the language issues, the paper is not well structured. The introduction does not give an overview of the different datasets out there and why special datasets are needed. Instead, it sketches already what the CORA dataset is and how data are treated before assimilation.
The next section heading suggests that the CORA dataset evolved from the Coriolis data centre, which makes no sense and indeed, the section discusses the data collection and processing at the Coriolis data centre. The last subsection in here already talks about the data retrieval and organisation of the CORA dataset, which should fit to the next section.
The third section is about data processing and quality control to generate the CORA dataset. The section states that flags done by the Coriolis data centre are kept, but does not explain if they are used to any advantage during the following quality control procedures.
The next section is supposed to deal with diagnostics. However, the biggest part is again about quality control and should have been discussed in the previous section.
The text seems to be a conglomeration of 'copy-and-paste' fragments from different reports and it is hard for the reader to follow the data processing. The reader is confused at the end which quality control procedures were done at the Coriolis level and which were done when collecting data into the CORA dataset. A data flow diagram would help a lot.

We have worked on the structure of the paper. The sections are now organized as follows:

1   Introduction
2   CORA3 dataset content
         2.1   Data sources
         2.2   Organisation of the CORA3 dataset.
         2.3   Data coverage
3   CORA3 data processing

We rewrote the introduction, shifting specific points to the following sections. We also added text that provides an overview of the different existing datasets (mainly World Ocean Database and EN3 dataset), how the data are quality controlled in these datasets and which corrections are applied.

We took care of separating the description of the CORA3 data processing from the diagnostics. We have moved the description of real time quality controls which was previously hared between section 2 and section 4 in an expanded section 3 (3.2.1).

To improve the flow of the paper we have moved the description of the CORA3 data coverage (previously in section 4) to section 2. Section 2 now focuses on the CORA3 data content, with description of the sources, the organisation of files and the plots showing the data coverage.

We have added a data flow diagram (beginning of the section 3) to clarify the different steps of data validation and correction, those performed in real time and those performed in delayed time.

Specific points:

1- I will not go into type setting here and correcting the spelling and language problems as this would go beyond the duties of a reviewer. Instead I will highlight some specific points to support the general comments.

We have checked the spelling and grammar.

2- There is no consistency in using the acronym 'CORA' and 'CORA3' in the text. The authors explain that CORA3 is the recent version of the CORA dataset, but they then do not stick with one acronym. I therefore assume that when 'CORA' is used, all versions of CORA are meant, while 'CORA3' means that some method or correction is specific to this latest version only? This needs further explanation within the text.

First two versions of CORA had been released in 2007 and 2008 (CORA1 and CORA2 respectively). However, the process used to produce CORA has changed and the procedure described in the paper is specific to CORA3 (and further versions). The use of the two acronyms 'CORA' and 'CORA3' was confusing and we decided to stick with 'CORA3'.

3- A lot of acronyms were used, but not explained (e.g. T, S, ATLAS, GODAE, PMEL, GT-SPP, WOA09, SST, SLA, XB, BA and TE, i.e. BATHY and TESAC); others are defined and then never used (e.g. GOSUD) and others defined twice (e.g GSSL).

We have checked that the acronyms are defined upon their first use.

4- Not sure about the OS policy about web links within the text. But my feeling is that most of the internet links occurring in the text could have been incorporated as references.

We have replaced web links with references each time it was possible.

5- p1274 l23: Which decade?

The 2000s decade.

6- p1277 l7: What is MyOceanII?

The aim of the MyOcean project is to pool the resources of the various European ocean forecasting centres in order to create a Global Monitoring for Environment and Security (GMES) Marine Service by 2014. MyOcean2 is the ultimate step before ECOMF (European Centre for Ocean Monitoring and Forecasting)

7- P1277 l17: What determines if the Coriolis centre collects in real-time or in near-real time mode?

Coriolis centre collects in real time or near real time mode depending on data sources and availability.

8- P1278 l19: What do quality flags of 0 and 5-9 mean then?

A flag 0 means that no quality control was performed, a flag 5 means that the value was changed, flags 6 and 7 are not used, flag 8 means that values are interpolated and flag 9 means that the value is missing. We add a table with the complete definition of flags.

9- P1278 l25: How large do the residuals have to be to produce an alert?

For each profile, a residual is calculated at each standard level and a mean is computed in each layer (0-100m, 100-300m, 300-700m, 700-1000m, 1000-1600m, 1600-2000m). The residual is the difference between the observed value and the analyzed field normalized by $\sqrt{\sigma_{UR}^2 + \sigma_{ME}^2}$, where $\sigma_{UR}^{\acute{e}}$ represents small scales unresolved by the analysis and considered

as noise and σ^2_{ME} corresponds to the instrumental errors. An alert is produced each time the normalized residual is larger than 8 at a standard level or the mean normalized residual is larger than 4 in a layer.

10- P1283 l12: I am confused here. The figure 4 shows the percentage of suspicious temperature and salinity profiles in GLORYS2V1 and not in the CORA dataset.

It is true that figure 4 shows the percentage of suspicious temperature and salinity profiles in GLORYS2V1. This reanalysis assimilated profiles from a preliminary version of CORA3 in situ profiles data base (this preliminary version contains the same profiles as the final version of CORA3 but covered 1990-2009 only). The suspicious profiles identified in GLORYS2V1 have been transmitted to the Coriolis data centre in order to check them and flag them, if necessary, in the final CORA3 data base.

11- P1283 l16: The authors explain the peak in the number of suspicious salinity profiles around 2000. But what are the consequences? Are the profiles removed from the dataset as they are not conform with the climatology (bad idea) or are they kept within the data collection and flagged as 1 (false alarm)?

In GLORYS2V1, the quality control (QC) performed has been done off line, when the whole reanalysis was produced. The QC then provides a grey list of suspicious profiles that was sent to the Coriolis data centre. Based on the visual check of these profiles it was decided or not to flag as bad the corresponding profiles. About half of the suspicious profiles identified in GLORYS2V1 reanalysis were really bad profiles and are flagged as bad in CORA3 data base. The other half are false alarms or profiles for which it is difficult to say whether they are bad profiles or not. Those profiles have a flag value equal to 1 in CORA3.

12- P1283 l23: Why do the authors expect the erroneous profiles to be randomly distributed in space? As you say, instruments with defective sensors would generate a trace. Very interesting is the increased number of suspicious salinity profiles at the northern end of the Antarctic Circumpolar Current. Could this be the result of a mismatch between the in-situ profile from e.g. within an eddy or meander to the climatology?

The reasons for erroneous measurements in an ARGO float profile are numerous. It can be due to a sensor (pressure, temperature, salinity) defect or ageing. There may also result from transmission errors. These reasons are *a priori* independent from the Argo float location and that is why we a priori expect erroneous profiles to be randomly distributed in space. Of course, when an Argo float has a permanent sensor failure, one expects some accumulation points on the map related to its trajectory (see the comment in the article).

It is apparent in Figure 5 (bottom) that there are more salinity profiles than temperature profiles below 30°S. In 2009 there are 56 suspicious salinity profiles below 30°S and after a visual check, about 25 of them are due to a bias in salinity probably due to a sensor failure

(only 7 profiles have an apparent bias in temperature). This kind of failure is common for Argo float salinity sensors.

The possibility that there would be a real large difference between the in situ profiles from within an eddy and the climatology can not be fully excluded. A similar behaviour is found in the Kuroshio region. However, this does not seem to be the case in the Gulf Stream or Aghulas regions, thus it is difficult to conclude.

13- P1284 l9: Why doesn't the check at the Coriolis centre find all duplicates?

The duplicate check at the Coriolis centre is run when the data is collected. Because it is embedded in a real time process, the duplicate check procedure experiment can present temporary failures and not run for a short while. Moreover the duplicate check program run at the Coriolis centre has evolved since 2001 but the newest versions were not systematically re-run on the data already loaded onto the Coriolis database.

14- P1284 l23: Does it mean that 1.5

1.5% of the profiles were found to have at least one duplicate (most of them have only one duplicate). We removed the duplicates. A large part of the duplicates were found in the BA and TE files (data from the GTS).

15- P1284 l2: Why is the name CORIOLIS in capital letters here?

We have corrected the error.

16- P1288 l22: Figure 5 does not show the percentage of profiles with bad positions!

We have corrected the error.

17- Figure 2: The text introduces the reader to quality flags ranging from 0-9 and explaines flags 1 to 4. However, this figures shows quality flags of 24. Please explain.

These quality flags of 24 are only for the purpose of the plot and indicate that this test (acceptable range) fails at some observed levels of the profile. This profile was visually checked before any flag were modified in the CORA3 dataset. After visualization, all the temperature values measured at depth below 360 m were flagged as bad data (flag 4) in the CORA3 dataset. This was clarified in the legend, so as to avoid confusion.

18- Figure 3: the red dots are hard to see and the font size is too small.

We have improved the figure 3 quality (and figure 1 as well)

19- Figure 4: This figure is based on the GLORYS2V1 dataset, while the paper describes the CORA dataset. The legend should also explain why the profiles are classified as 'suspicious' or at least refer to the text.

It is true that the legend was a bit confusing. We changed it according to your suggestion.

Fig. 4. Percentage of suspicious temperature (black) and salinity profiles (red) as a function of year in GLORYS2V1 reanalysis quality control. The profiles grey list was provided to the Coriolis data center in order to improve CORA3 quality. See the text for more details.