This manuscript describes the latest version (v5) of the ocean re-analysis system C-GLORYS. The manuscript is well written, and while there are some things I would have presented somewhat differently, there is nothing I could suggest that would substantially alter the manuscript, and the authors should feel free to treat my minor comments below as mere suggestions.

My largest concerns were to do with the presentation and discussion with regards to the AMOC in Section 3.3. Firstly, it is somewhat difficult to make out the C-CGLORYS overturning underneath the RAPID observations in my paper version of Figure 9 – although I suppose that is one of the virtues of electronic media, as it is much easier to see in a zoomed in electronic version. It would also be illuminating, but not necessarily convenient, for the authors to show the non-Ekman component of the overturning streamfunction. All models, pretty much by definition, would replicate the Ekman component of the overturning derived by the RAPID observations, since that is solely determined by the wind stress forcing which is typically identical, or near identical to that used in the RAPID calculation. Replicating the density driven circulation, on the other hand, is more difficult — and ideally it would be the correlation between the non-Ekman portions of the overturning in the RAPID observations and C-GLORYS re-analysis that would be most interesting [Roberts et al., 2013]. In the absence of that calculation, however, it is noteworthy that neither C-GLORYS analysis appears to pick up the early period peak in the RAPID observations — although they do appear to pick up this 2005-2010 decrease in the RAPID overturning, and subsequent increase
after 2010 from about 2007 onward — which coincidentally would be when the ARGO float array is fully deployed, and their analysis potentially morphs from being largely driven by the SST nudging to one where the sub-surface profiles are playing a substantial role. Perhaps the authors may wish to substantiate on that further.

**Minor Comments**

1. p. 7, l. 2: 9% of the observations affected by the bug seems large. Are you inferring that 9% of the profiles were actually only surface measurements? Furthermore, is this 9% of the profiles, or 9% of all the profile observations at all levels?

2. Figure 3: There is a large spike in the monthly inflation value in 1993, very close to the coming on line of the altimeter data. Coincidence? Spurious?

3. p. 8, l 18: “namely the floats used represent a fairly independent dataset.” Firstly, tacking this onto the end of the previous sentence does not make grammatical sense, but more importantly, you are being unduly brief with what in my mind is a fairly complex statement. What I believe you are saying is that because you are comparing observation minus background for floats (as opposed to say moored buoys), measurements at any given point can be consider independent, since (in principle) no one float makes repeat measurements at the same location. Perhaps it would be better to expand your statement somewhat – making it a complete sentence while you are at it.

4. Figure 4 and discussion p. 8, ll. 20-23. Only the global average observations minus background stats are shown. It would be worth at least showing the tropical (possibly Tropical Pacific) statistics that you note as significantly improved. Can this be attributed to the decrease in background-error standard deviation in the tropical Pacific. Conversely, observations minus background for the North Atlantic where the background-error has been increased and the skill decreased could be illuminating. Is the Gulf Stream more misplaced in the non-assimilated version of v5 compared to v4? Finally, you attribute the decreased skill at high latitudes to differing sea ice cover – but the sea ice cover should be largely constrained by the sea ice concentration observations. Is is not simply that you have increased the background error in this region as well? Note that the SST rmse is reduced at high latitudes as well.

5. Figure 8 and Section 3.2: The more accurate (compared to the NOAA/NODS estimates) trend in 2000m heat content in the Gulf Stream region seems at face value at odds with earlier statements regarding a loss of observation minus background skill in this region. However, there is also a decrease in SST rmse here as well, presumably due to the increased background error standard deviation – although the lack of flux correction could also play a role. Is the trend in 2000m heat content largely surface driven here?
6. Figure 9 and Section 3.3: As mentioned above, a comparison of the non-Ekman component of the Atlantic overturning streamfunction would be useful, but not essential.

7. Sections 2.2.5 and Section 3.4: Note on using PIOMAS as data. While PIOMAS does validate well with the sea ice thickness over the period it was validated – mostly ICESat data. However, it does not validate as well over data from more recent periods, possibly overestimating March ice thickness. However, I have no citable literature to back my claim, so this amounts to hearsay. Nevertheless, the Arctic ice volumes in C-GLORYSv5 are undoubtedly more realistic then those of v4. Undoubtedly, the spatial thickness patterns are close to those that are begin imposed by PIOMAS, nevertheless, a spatial map could be useful, especially if it can be compared with a satellite observations for a particular period. Laxon et al. [2003] could be used for an early altimeter based thickness estimate.

8. Figure 10: A yearly timeseries (with collapsed vertical axis) along with a seasonal cycle climatology might be more easily decipherable than the monthly timeseries shown. It might even be possible, with dual vertical axis, to plot area and volume on the same plot so that the number of sub-figures remains the same.

9. Section 3.4 and Figure 10. There is (presumably?) no ice thickness restoring performed in the Antarctic, yet the volume field in v5 is also more stable here than in v4. No mention was made of any sea ice improvements in the model, so can this be attributed to either the removal of the the flux corrections, or the changes (increase?) in background error covariance for the profiles. I note the error RMS error in SST is also reduced in both the Antarctic and Arctic. Would there have been similar improvement in the Arctic volume without the PIOMAS restoring?

**Typos and Grammatical Errors**

1. Title: I was always taught that a colon(:) should be followed by capitalization – and a complete sentence.

2. Section 2.1.1 Heading: Ocen → Ocean

3. p. 5, ll. 6-7: “constant value of 10m along the latitudes ···” I would have simply said “a globally uniform value of 10m ···”

4. p. 6, l. 15: “Differently” Use “Conversely”

5. p. 6, l. 30: Not really a typo, but insert $\alpha$ into the sentence to make this more explicit: “the threshold, $\alpha$, being increased from 6 to 9”

6. pp. 6-7, last and first line: You can’t be shallower than the first level, but you can be shallower than the **middle** of the first level.
7. p. 7, l. 18: “passing” → “changing”
8. p. 7, l. 30: “which we found caused by” → “which we found was caused by”

References
