The authors thank anonymous Referee2 for the constructive feedback on the manuscript. The manuscript has now been revised considerable to address all of Referee2 suggestions. Below are list of point-by-point responses to all of Referee2's comments. We think that the revised manuscript is now considerably improved compare to its initial version.

Ref2: Land/Ocean balance. I acknowledge that fact that the authors have tried to combine the land and the ocean carbon cycle analysis in the same manuscript. But, as a reader, you still very clearly see that the ocean and land parts have been written by different people. The abstract is a good example of that, listing a few global numbers for the land, and focusing on deficiencies in the spatial representation of some marine tracers. Again, in the result section, the ocean analysis is based on 13 figures, whereas the land one only uses 4 of them. The ocean section is structured with sub-sections it is not the case for the land one. Of course, a perfect balance is not possible (and not wished), but more integration / balance is needed.

With regards to the lacking terrestrial component in the manuscript, we now explicitly states in the manuscript (e.g., in the introduction section) that we are not the main developer of the land component, and that inclusion in the NorESM model framework does not introduce any significant difference with the native ESM framework of Community Climate System Model (CCSM4). In addition, many publications on the CLM model (both description and evaluation) have been recently published by the main CLM developers. In the revised manuscript, we tried to refer to these previously published papers as much as possible to reduce redundancies and better elaborate what has been previously discussed. Nevertheless, in the revised manuscript, we have added more figures (now 6 figures for land as compared to 13 figures for the ocean) and table (2 tables for each the ocean and land) related to the terrestrial component. We also note in the paper that several terrestrial components such as simulated sensible and latent heat flux from the land, and evaluation of the simulated temperature and precipitation are included and discussed in the accompanying NorESM physical component evaluation study, which is submitted to the same special issue in Geoscientific Model Development journal (Bentsen et al., 2012).

As suggested, we have also refined the abstract, added subsections into the terrestrial analysis section to improve the readability, and also moved the discussions of temporal evolution of sea-air and land-air CO_2 fluxes into a new subsection to increase the integration between the land and ocean discussions.

Ref2: Improvements from a previous version. A general goal in model evaluation is, to my view, to show how a newer version of a model is better / worse than the previous one and for what reasons. This is only touched upon in the present manuscript and should be better discussed here. I would integrate in the main article some of the figures that are now only in the supplementary material. Is this possible to draw conclusions for the land part as well?

In the revised manuscript, we have included additional discussions on specific improvements or worsening parts (including several related comments raised by Referee1) of the model performances. Among them, we elaborate better the improvement of surface nutrient concentration in the high latitude regions, how the

improvements in the mixed layer depth was implemented, as well as how we think the AMOC is responsible for the worsening deep ocean biogeochemical travers distribution.

With regards to moving the figures from supplementary to the main manuscript, many of the figures in the supplementary materials, in some way, has been shown in the previous studies (Assmann et al., 2010; Tjiputra et al., 2010, both in GMD) as they referred to the previous generation ocean-biogeochemistry model (MICOM-HAMOCC). They were mostly reproduced in a different plotting formats and styles to improve the visualization when comparing the version in the NorESM and one in the Bergen ESM. Thus, they are kept in the supplemental section to reduce redundancy. On other hand, in the revised manuscript, we have added surface silicate concentration comparing the previous model version with NorESM (now Figure 12) because the surface silicate concentration has never been shown in previous publications.

When discussing the improvement from the previous terrestrial model generations, contrast to the ocean carbon cycle, the NorESM adopted a new terrestrial component in the Community Land Model-Carbon Nitrogen (CLM4-CN) model as compared to Lund-Potsdam-Jena (LPJ) dynamical vegetation model used in the previous generation Bergen ESM. Thus, a direct comparison is beyond the scope of this manuscript, which main purpose is to describe provide basic description and evaluation of the carbon cycle components.

Ref2: Links between physical biases and biogeochemistry/carbon cycle. This point is illustrated for the ocean part - and could be even better discussed with the inclusion of figures showing previous modelling results. But it is clearly lacking on the land side. I would strongly recommend some additional analysis to strengthen the manuscript.

As also suggested by Referee1, we have refined the discussions in the manuscript by clarifying how certain improvements in the ocean components was achieved from both the physical and biogeochemical perspective. For example, we have elaborated how the improvement in the mixed layer depth was achieved, and how we argue that the stronger AMOC (rather than the biogeochemistry parameters) is responsible for the worsening deep water tracer distributions (see Section 4.2.1 and 4.2.2). As mentioned above, figures from previous models are kept in the supplementary to reduce repetitions and the bulk number of figures. From the land side, a direct comparison with the previous model is not feasible as they are two different models. And it is not the main goal of this paper to evaluate the difference between these two terrestrial carbon cycle models (CLM4-CN and LPJ).

Ref2: Ocean analysis. I feel that the evaluation of 3D variables first (nutrient, dic, alkalinity) and then NPP and air-sea carbon fluxes are a bit disconnected. They should be better integrated. For example, could the authors show the different nutrients and the most important limiting nutrient spatially, and how it relates to NPP?

In the revised manuscript's section on biological production, we have added several references back to the 3D nutrient figure (i.e., phosphate). A map showing the

limiting nutrient spatially has been shown in previous publication (Assmann et al., 2010). Here, we show the different productivity-limiting annual mean surface nutrients (phosphate, nitrate, and dissolved iron) in one figure (now figure 9) panel with similar unit (equivalent to micromole P/L). The figure illustrates regionally which nutrients essentially act as the main nutrient limiting factor. We have also included additional discussions in the beginning of section "4.2.3 Biological production" to better integrate the relationships between the nutrients and biological production simulated in the model (how the nutrient usage for biological production is formulated in the model). In the same section, we also included a new silicate plot, which is relevant in determining the export of inorganic carbon simulated in the model.