The authors thank anonymous Referee1 for the positive and constructive feedback on the manuscript. Below are point-by-point responses to all of Referee1's comments. We believe that after addressing all of the Referees' comments, the current revised manuscript is substantially improved compared to its initial version.

**Referee#1: P3037 line 23: the ocean model in NorESM is indeed unique among other CMIP5 models in using isopycnic coordinate system. The authors explain the advantages and issues of this algorithm later in the manuscript. However, it would be very useful for the users of CMIP5 data to learn about them earlier.**

We agree that it is useful and add the following statements in the introduction section:
"The ocean carbon cycle in NorESM is unique to most other Earth system models due to its coupling to an isopycnic coordinate ocean general circulation model. One of the advantages of such coupling is more accurate representation of the transport and mixing of biogeochemical tracers along isopycnals in the interior ocean. The isopycnic model also avoids physically inappropriate splitting of transport and diffusion processes in horizontal and vertical components as done in a z-coordinate model (Bleck, 1998; Haidvogel and Beckmann, 1999). Through the vertically adaptive grid, areas of high horizontal and vertical density gradients can be simulated well by the model. Earlier study by Assmann et al. (2010) also shows that higher spatial gradients in tracer distributions can be achieved as well. On the other hand, depending on the number of density surfaces, an isopycnic coordinate model may or may not represent well the buoyancy driven circulation. In areas of low density gradients, the model cannot simulate velocity shear and surface processes are also more difficult to simulate in outcropping layers, which is avoided through introduction of a non-isopycnic surface mixed layer."

**Ref#1: P3038 lines 4-5 it is stated that 'biogeochemical states . . . strongly depend on the quality of the physical fields in the model'. While providing a rather detailed analysis of the ocean's physical state (including ocean temperature, salinity and MLD), the paper omits analogous evaluation of the physical fields which influence the terrestrial carbon cycle (for instance air temperature, precipitation, surface radiation). This omission leaves the land carbon cycle evaluation incomplete.**

Within the same special issue in GMD "The Norwegian Earth System Model: NorESM; basic development, validation, scientific analyses, and climate scenarios", Bentsen et al. also submitted an accompanying manuscript, which evaluates several physical components related to the terrestrial carbon cycle in more details. For example, they assess the model simulated annual mean surface (sensible and latent) heat flux from the land model, surface air temperature over land, cloud fraction, and mean precipitation. We have included this information and reference to the revised manuscript (see Introduction section). We have also included figures of mean surface temperature (over land) and precipitation in the revised manuscript.

**Ref#1: P3039, line 5: Here and in several other locations in the paper, an improvement of one or another model component is mentioned. However, the implications of this (and further) improvement(s) are not discussed. Hence, it sounds rather unjustified whether these improvements were really necessary.**

We now included a new reference (Kirkevåg et al., 2012 GMDD, same special issue), which discusses further the improvement of the aerosols module and aerosol-cloud-radiation interaction in the atmospheric component of NorESM. We have elaborated in the revised manuscript that the new aerosol module especially reduces the bias in near surface mass concentrations and aerosol optical depth. Other improvements are now also discussed further (see also below responses).

**Ref#1: P3040, lines 5-10: Have these recent model developments been documented and evaluated elsewhere? More details on these model improvements would be helpful.**

In the revised manuscript we have extended the ocean general circulation model description section to detail the recent modifications relative to the previous model generation. In addition, a dedicated separate manuscript, focusing on improvement on the ocean physical component of the NorESM, is planned.

**Ref#1: P3040, ocean carbon cycle model description: It would be helpful to see all major features of the model summarized in a table. Such a table could include a list of model tracers treated prognostically in the ocean.**

We have now included a new table (now Table 1) summarizing the main biogeochemical features of HAMOCC, as well as listing the prognostically simulated biogeochemical tracers as suggested. The new table is now also referenced within the text.

**Ref#1: P3041, lines 1-2: Please provide a citation for the sediment model. Also mention here if sediments were included in your CMIP5 experiments.**

The reference to the sediment model and further information on its inclusions in the CMIP5 experiments are now included in the revised manuscript.

**Ref#1: P3041, line 3: Mention that the NPZD model in HAMOCC is extended by DOC.**

Done.

**Ref#1: P3041, line 16: add 'constant' before Redfield.**

Done.

**Ref#1: P3041, line 24: add 'the upper' before 100 m.**

Done.

**Ref#1: P3041-3042: Description of ocean carbon cycle model misses information on which light scheme was used in the experiments.**

We have included the following statements to clarify the light scheme used:
"The available light is formulated based on the prognostic incoming solar radiation from the atmospheric model reaching the ocean surface. Light penetration decreases with depth according to an exponential function with a gradual extinction factor formulated as a function of water depth and chlorophyll (phytoplankton-to-chlorophyll constant ratio is used) concentration (Maier-Reimer et al., 2005)."

**Ref#1: P3041-3042: Please add a description of how weathering fluxes were treated in your simulations.**

In the revised manuscript, we stated that the HAMOCC model does not include any weathering fluxes. For our integration timescales, this usually should not affect the tracer inventories significantly. We have now a version with prescribed matter inputs (from NEWS2, Mayorga et al., 2010), but this is not the version submitted to GMD.

**Ref#1: P3043-3044: It would be useful to see a summary of all land carbon cycle compartments/pools (dead and live) in the model listed in a table.**

A summary of ALL land carbon/nitrogen pool is documented in CLM4 technical report (Oleson et al., 2010), and now stated in the revised paper (section 2.3). We have added a summary of major carbon pools in a new table (now Table 3) as suggested. The new table also lists their values as simulated within the NorESM framework as well as from observational estimates with references.

**Ref#1: P3044: Specify how autotrophic and heterotrophic respirations are calculated in the model as these processes are critical for NPP variations on land.**

We have extended the terrestrial model description section to include the following clarification: "The autotrophic respiration is simulated as the sum of maintenance and growth respiration processes. In living biomass, maintenance respiration is a function of temperature and tissue N concentration (Thornton and Rosenbloom 2005). Growth respiration is calculated as a constant factor of the carbon allocated to growth of new tissues. For computation of heterotrophic respiration, CLM-CN uses a converging cascade representation of soil organic matter dynamics (Thornton et al., 2002, Thornton & Rosenbloom (2005). The model has three litter pools (labile, cellulose, and lignin) and a coarse woody debris pool together with four soil organic matter pools. The three litter pools differ in base decomposition rate, with turnover time ranging from 20 h to 71 days. The four soil organic matter pools differ in base decomposition rate (turnover time is 14 days to 27 years) and C/N ratio (10–12). There is no distinction between surface and belowground pools. The soil organic matter dynamics is conditioned by the soil-nitrogen cycle. In case of nitrogen mineralization, the soil organic matter base decomposition rates are functions of soil temperature (Lloyd and Taylor, 1994) and soil water potential (Orchard and Cook, 1983; Andrén and Paustian, 1987). However in the case of nitrogen immobilization the decomposition is limited by the nitrogen availability and by the plant demand for mineral nitrogen."

**Ref#1: P3045, line 15: It takes several tens of thousands of years for deep-sea sediments to reach an equilibrium state if you initialize them from zero values. Give more details with regard to sediments initialization and spin-up.**

To clarify this comment, we have included the following statements in the revised manuscript (see Section 3):
"Due to the limited computational power, it is currently impossible to spin-up the sediment compartment to reach a steady state. In the future, we plan to spin up the model sediment with an acceleration technique. Nevertheless, for the current CMIP5 experiments set up (i.e., integration times of a few hundred years), the sediment water column interaction contributes little to the ocean tracer inventories (a reason for which most modeling groups do not consider the sediment at all)."

**Ref#1P3046, line 17: Also discuss if the phasing of the simulated variability is in line with observations.**

The simulated global mean temperature variability and attribution of different controlling factors is discussed in accompanying papers and therefore, to avoid redundancy, we have added the following statements in the revised manuscript:

"Comparison between the observed and simulated global mean surface temperature trend over the historical periods and contributions of different elements in the simulated variability are discussed in accompanying manuscripts of Bentsen et al. (2012) and Iversen et al. (2012). Bentsen et al. (2012) show that the mean temporal trends of three historical NorESM ensemble members follow the observed trend closely. For example, both the observation and model simulations yield 0.14 K $decade^{-1}$ warming trend for the 1961-2010 periods. In their study, Iversen et al. shows that the increasing warming trend since the 1970s is predominantly attributed to the combination of opposing radiative forcing of greenhouse gas and the aerosols."

**Ref#1: P3046-3047: Section 'Ocean biogeochemistry' starts with an extended comparison of the physical fields (T, S, MLD) with observations. Consider to separate this very useful evaluation of physical parameters in a new subsection.**

Done. In the revised manuscript, the section 4.2 is replaced from "Ocean biogeochemistry" to "Ocean physical and biogeochemical properties" and we have added two new subsections " 4.2.1 Physical fields" and "4.2.2 Biogeochemical tracers".

**Ref#1: P3048-3053: The globally integrated values (for instance, primary production, export production, etc) and budgets for the ocean are now discussed and spread over several locations in the text. It would be very helpful if they were summarized in a separate table, perhaps similar to your Table 1 for terrestrial parameters.**

We agree that such table would be useful. In the revised manuscript, we have added a table (now Table 2) summarizing the globally integrated annual mean net primary production, organic and inorganic carbon, and silicate exports, as well as sea-air $CO_2$ fluxes computed over early historical (1850-1859) and present-day (1996-2005) periods. We have also referred to this table throughout the text where it is relevant.

**Ref#1: P3048, lines 18-19: specify what this is improvement is due to.**

In the revised manuscript, we have added the following sentences:
"The improvement was achieved through implementation of new turbulent kinetic energy balance equation following Oberhuber (1993) and updated parameterization of mixed layer restratification by eddies following Fox-Kemper et al. (2008)."

**Ref#1: P3049, line 9: again, how this improvement was achieved?**

It is difficult to pinpoint the main reason for the improvement as many parts of the models are modified continuously from the previous generation model (BCM-C). Nevertheless, in the revised manuscript, we have included the following statements that state the likely reason behind this particular improvement:
"While the updated model was modified considerably from the previous model version, this improvement in the surface concentration is likely due to the doubling of the phytoplankton nutrient uptake half-saturation constant from 0.1 to 0.2 umol P $l^{-1}$. Higher half-saturation constant reduces the nutrient uptake when the surface nutrient concentration is low, and hence increases the mean nutrient concentration near surface."

**Ref#1: P3049-3050: In your discussion on ocean biogeochemical tracers, you attribute their distributions to water masses proper ties and other physical properties in the model. This is alright, but what is the role of biogeochemical parameterizations of for instance,**

**the sinking fluxes, remineralization (particularly specific to the model HAMOCC) in the distributions of phosphate and oxygen?**

Following the discussions of model-data discrepancies of phosphate concentration in the deep ocean, we have added the following sentences in the revised manuscript:
"The parameterization of the biogeochemical processes, such as particle sinking speed and remineralization rates of dissolved and particulate organic matters can also influence the nutrients and oxygen distribution at depth. For example, high vertical sinking speed would translate to higher nutrient at depth and high remineralization rate would increase nutrient and decrease oxygen concentration of younger water masses. However, these controlling parameters were not modified considerable relative to the previous version. In addition, in the low-resolution version of the model (i.e. NorESM-L), where the simulated overturning circulation is much more reasonable at ~18 Sv (Zhang et al., 2012), the phosphate concentration at deep ocean is much more realistic. The NorESM-L also simulates older "ideal age tracer" in the deep ocean than the medium resolution version."

**Ref#1: P3049, line 25: How well does the model perform with regard to iron?**

We have included a new figure (now Fig. 9) illustrating the surface distribution of dissolved iron simulated in the model. In addition, the following statements are now included in the revised manuscript (section 4.2.2):
"Dissolved iron is a limiting macronutrients for marine biological production as simulated in HAMOCC5. The main source of iron concentration in the surface is through aerial dust deposition, which is transported out of deserts over land (e.g., the Sahara). Since the model used the same climatology iron (dust) deposition as the previous model version, the distribution of surface iron concentration is very similar to the one shown in Fig. 13 of Assmann et al. (2010) (see also Fig. 9). Maximum surface iron concentration is simulated in the Mediterranean Sea with slightly higher than 2 nmol Fe $L^{-1}$. Several regions such as the North Atlantic, northern part of Indian Ocean, and parts of the Southern Ocean also have relatively high surface iron concentration, ranging between 0.4-0.6 nmol Fe $L^{-1}$. The Pacific Ocean is mostly depleted with regards to the iron concentration. This feature is consistent with the limited observational-based estimates, as shown in Parekh et al. (2005)."

**Ref#1: P3050, lines 24-25: An explanation why is it not easy to simulate the correct TA would be helpful for the readers without background in running global ocean biogeochemical models.**

Simulating the correct alkalinity remains to be a major challenge for the ocean biogeochemical modelers. Fig. R1 (below) shows surface alkalinity concentration (in mol eq/$m^3$ units) simulated by five CMIP5 models. Most of them overestimate the observed GLODAP values throughout most of the ocean regions. Identifying the reasons for this caveat is beyond the purpose of this manuscript. However, in the revised manuscript (last paragraph of section 4.2.2), we offer some insights of what potentially could contribute to the bias in the simulated alkalinity: bias in the GLODAP data used for model initialization (interpolation bias as well as that the GLODAP represents anthropogenic period), bias related to the simulated salinity, and too much CaCO3 production in the model, among others.

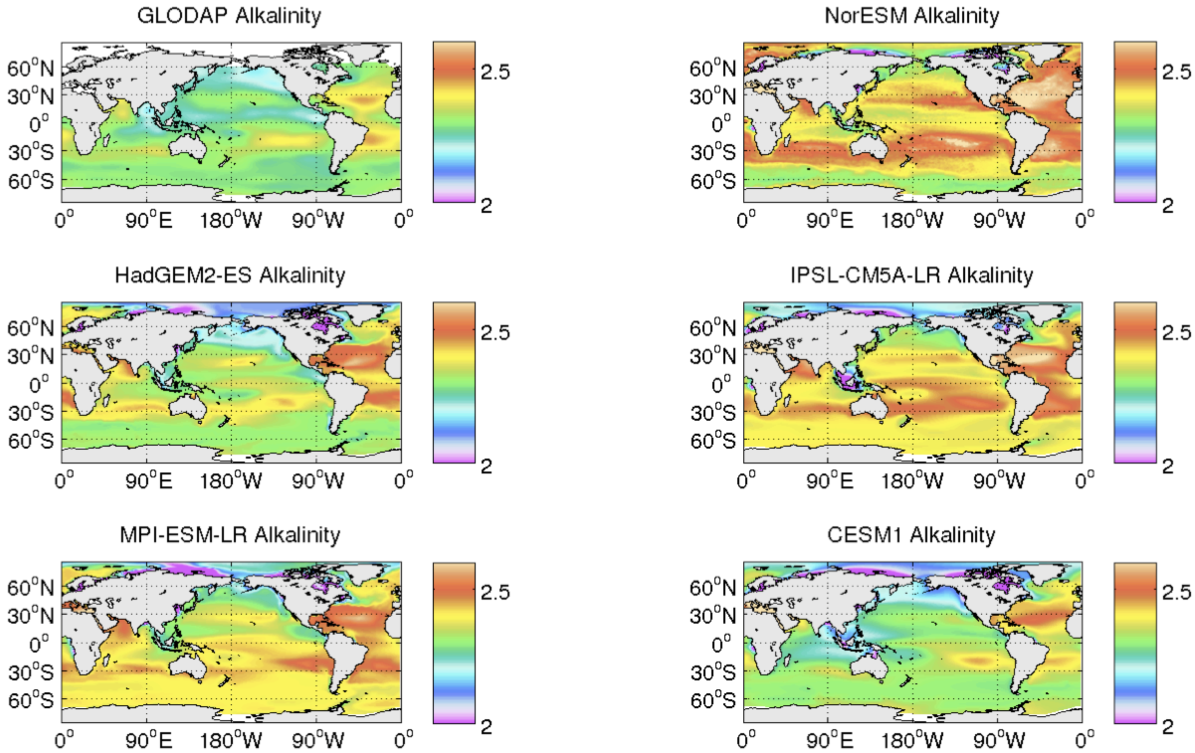

Figure R1. Annual mean surface alkalinity from GLODAP and five CMIP5 models.

**Ref#1: P3050, line 28: add a 'd' in 'compare'**

Done.

**Ref#1: P3051, lines 3-6: Discuss what implications does an overestimated TA have on the ocean buffer capacity.**

The following discussions are now included in the revised manuscript:
"Overestimation of alkalinity alone (i.e., without overestimation of DIC) would give higher carbonate ion concentration, and consequently would increase the buffer capacity. Nevertheless, as DIC also contains carbonate ion, a better approximation for $[CO_3^{2-}]$ concentration can be determined by alkalinity ($Alk \approx [HCO_3^-] + 2[CO_3^{2-}]$) minus DIC ($DIC \approx [HCO_3^-]+[CO_3^{2-}]$), as defined in Sarmiento and Gruber (2006). Since both alkalinity and DIC in the NorESM are overestimated by a similar factor, the simulated buffer capacity is not altered considerably."

**Ref#1: P3051, lines 7-8: It is unclear why do you need to show Alk minus DIC.**

Please see the previous response.

**Ref#1: P3052, lines 21-22: Model performance with regard to silicate distributions not shown/discussed. Hence, while maybe true, it sounds rather speculative to attribute the discrepancy in PIC export to surface silicate concentrations. You either need to include silicate in your evaluation or cite previous studies.**

We have modified the texts in the revised manuscript, added a new figure (now Fig. 12) showing a comparison in the surface silicate simulated by the NorESM and previous generation model (BCM-C). We have also added the statements pointing out that the figure shows that in regions of high biological productivity such as the North Atlantic, North Pacific, equatorial Pacific, and vast area of the Southern Ocean, the surface silicate concentration in the NorESM model is considerably larger than the previous generation model. This would explain the higher simulated PIC simulated in NorESM relative to the BCM-C model.

**Ref#1: P3053, line 9: Explain why the Southern Ocean is a region of increasing interest.**

In the revised manuscript, we clarify the importance of Southern Ocean by adding the following statements:
"While observational-based study (e.g., Le Quere et al., 2007) indicates a weakening $CO_2$ sink in the Southern Ocean, a model study by Tjiputra et al. (2010b) shows that, due to its efficient northward subduction of intermediate deep water, the Southern Ocean could continue as the dominant anthropogenic carbon sink in the future."

**Ref#1: P3053, lines 19-20: Why is this relevant here (the three regions playing a key role in the future)?**

In the revised manuscript, we replaced the sentence
"Therefore, we assume that these three regions will play key role in controlling the oceanic carbon fluxes as the climate evolves in the future."

with

"Over a long term period, a study with the Bergen Earth system model (Tjiputra et al., 2010b) reveals that, due to their water mass transport characteristics, the equatorial Pacific and the polar Southern Ocean could take up more $CO_2$ under a business-as-usual future scenario. On the other hand, the $CO_2$ uptake rate in the North Atlantic would stabilize toward the end of the 21$^{st}$ century, predominantly associated with the slowdown in the overturning circulation."

**Ref#1: P3054, lines 5-8: You are aware of course that the DIC anomaly is not the same as anthropogenic carbon. It may be a suitable approximation though. This has to be explicitly mentioned.**

We agree, and as suggested, have explicitly clarified the approximation in the revised manuscript. We have also used a better approximation for determining the anthropogenic $CO_2$ storage in the reised manuscript (i.e. HIST(yr1994) minus CTRL(yr1994), instead of HIST(yr1994)-HIST(yr1850)).

**Ref#1: P3054, line 17: Explain why the model estimates of CO2 uptake are lower than observed.**

In the revised manuscript, we have included the followings statements:
"The strong AMOC strength could contribute to the higher anthropogenic carbon storage in the North Atlantic, as absorbed anthropogenic carbon in this region is transported faster to the deep ocean. It is still unclear, however, why the model underestimate the anthropogenic carbon in the equatorial oceans. A study by Matsumoto and Gruber (2005) has indicated that the delta C* method adopted in Sabine et al. (2004) study has many limitations as well (e.g., they show that the method overestimates anthropogenic carbon in the equatorial regions)."

**Ref#1: Section 4.3: Is another paper focused on the assessment of the terrestrial biogeochemistry in NorESM planned? If not, this section needs a somewhat more detailed**

**analysis, i.e. starting with the evaluation of physical model parameters that are critical for terrestrial biogeochemistry. Evaluation of the terrestrial carbon cycle component could be expanded by discussing model performance with respect to reproducing radiation / surface albedo.**

Presently, there is no plan for a separate terrestrial biogeochemistry assessment. As now mentioned in the revised manuscript (e.g., see introduction section), the main development and maintenance of the terrestrial model was done at the National Center for Atmospheric Research (NCAR) in the United States. We are mainly user (not developer) of the model. In addition, a number of recent studies by the NCAR scientists focusing on assessing different components of the CLM-CN model have recently been published. Here, we have further clarified and briefly summarized what have been assessed in these studies (see also section 4.3). We also add recent multi-model studies (Arora et al., in revision; Jones et al., accepted), which indicate little difference between the terrestrial carbon cycle in NorESM and CESM1 (both use CLM4). The reason for this is also due the large similarity in the atmospheric model (CAM).

As mentioned in above responses, discussions of latent and sensible heat flux are discussed in accompanying paper (Bentsen et al., 2012). A thorough assessment of the biogeophysical impacts of the CLM4 simulated surface albedo on climate within the CCSM4 model framework is already documented by Lawrence et al. (2012, Simulating the biogeochemical and biogeophysical impacts of transient land cover change and wood harvest in the Community Climate System Model (CCSM4) from 1850 to 2100, J. Climate).

**Ref#1: P3055, lines 1-3: Explain why the terrestrial carbon uptake is lower.**

We have clarified in the revised manuscript that this is due to the strong nitrogen limitation formulated in the land model, which damps out the "$CO_2$ fertilization" effect. This is also shown in other studies using same land model, such as Lindsay et al. (submitted to J. Climate, now mentioned in the revised manuscript).

**Ref#1: P3055, line 7: A reference to precipitation and temperature distributions is unfounded without showing them.**

In the revised manuscript, we have added new annual mean surface temperature and precipitation plots, and referred to these plots in the discussions. We also note that a more detailed model-data evaluation of the simulated temperature and precipitation is discussed in a separate accompanying paper (Bentsen et al., 2012).

**Ref#1: P3060, lines 1-5: I am not sure this outlook is really relevant for the paper.**

We think that the outlook with regards to the river flux is relevant in the paper because in the manuscript, we discuss the drift in the simulated nutrient budgets in the model, which partly can be attributed to the missing riverine flux.

**Ref#1: Figure 13: The colorbar is extremely confusing: some colors (e.g. blue, red) are repeated several times. Does blue in continents shading stand for no data or values close to zero, or close to 100?**

In the revised manuscript, the figure is revised with improved colorbar.

**Ref#1: Figures 14-16: Likewise, the colorbar does not provide necessary details. For**

**instance, it is unclear if the dark blue in Antarctica for GPP values in Fig. 15 is zero or not.**

As suggested, we have reproduced the figures with improved colorbars (for Figs. 14-16) in the revised manuscript.

**Ref#1: Figure 16, legend: Clarify which respiration is shown here.**

We have clarified in the revised manuscript within the respected figure caption that it is the sum of autotrophic and heterotrophic respirations.