Response to Dr. F. Perez (Referee #1)

We thank Dr. F. Perez (Referee #1) for his thoughtful comments. We have tried to take his comments into account in the revised manuscript as follows:

Minor Comments

Page 12166 Line 17-25. *The description of the Ocean Interior* CO_2 *inversion method is rather simplistic. The manuscript linked the use of the Green Function(GF). However, the cited articles (Gloor et al. 2003; Gruber et al. 2009) did not mention GF although they described the use of 'dyes' to calibrate the age of the water mass parcel. The GF was introduced earlier (Haine and Hall 2002; Primeau 2005). Khatiwala et al. (2009) used this method to evaluate the anthropogenic* CO_2 *assuming a steady-state circulation. Other important fact about this method is that it uses the anthropogenic* CO_2 *estimations given in GLODAP (Key et al. 2004) as an additional constrain. This should be addressed.*

A. We thank the referee for pointing out some potentially confusing text in our discussion of the ocean inversion. The ocean inversion described in Gloor et al. (2003) and Gruber et al. (2009) uses Green's Functions that are derived from ocean model simulations in which dye tracers are released at the surface of the ocean. The Green Function (GF) approach pioneered by Haine and Hall (2002) and used to estimate anthropogenic carbon uptake by Khatiwala et al. (2009) is quite different from the ocean inversion, although both are constrained by the GLODAP data. A detailed comparison of these techniques is presented in the RECCAP paper on ocean storage of anthropogenic carbon (Khatiwala et al. 2012).

To the best of our knowledge, the GF method has only been used to estimate the anthropogenic component of carbon flux, while this manuscript is concerned with the net air-sea fluxes (e.g. the sum of the natural and anthropogenic components.). Therefore, we have not included it here.

In order to avoid confusion and highlight the ocean data used in these estimates, we will re-phrase this paragraph as follows:

"The ocean inversion estimates regional air-sea CO_2 fluxes from ocean interior observations of DIC and other species from the Global Ocean Data Analysis Project (GLODAP) (Key et al., 2004) and ocean model simulations that describe how fluxes at the surface influence tracer distributions in the interior ocean (e. g. Gloor et al., 2003; Gruber et al., 2009). The inversion initially estimated fluxes from 30 ocean regions, which were subsequently aggregated to 23 regions (10 regions in the Pacific). The anthropogenic and natural air-sea CO_2 fluxes are estimated separately (Mikaloff Fletcher et al., 2006; Mikaloff Fletcher et al., 2007), and the net air-sea fluxes shown here represent a sum of these components and a riverine flux estimate following Jacobson et al. (2007), i.e., +0.08 PgC yr^{-1} in the North Pacific extra-tropics and +0.04 PgC yr^{-1} in the tropical Pacific. These contemporary air-sea CO_2 flux estimates were initially reported in Gruber et al., 2009, but the anthropogenic component has been scaled for this manuscript to match the RECCAP period of 1990-2009. Since this method does not resolve seasonal and interannual variability, the results are only used to compare the 1990–2009 mean air-sea CO_2 fluxes."

Page 12167 Line 20. *In addition to the restoring SSS, some OBGCMs are also constrained by using the Carbon data, both the one observed and the one computed as anthropogenic, given in GLODAP. If it is the case, this fact should be addressed.*

A. There are some exploratory efforts underway that assimilate biogeochemical data, such as the work of Valsala et al. (2010). In the long term such efforts certainly offer promise, but to date the critical activity of skill evaluation has only just begun. More importantly, the product of these efforts need to be distinguished from OBGCMs, as they are assimilation efforts instead of forward models. Please note that in the interests of being inclusive and in recognition of the value to the scientific community of assimilation efforts, we have included the output of Valsala (2010) in this manuscript.

Page 12172-12173. *In terms of inter-annual variability in the Tropical Pacific, the wind-products can introduce an additional source of variability. Several papers are cited between lines 5 to 15 in page 12173 but they could use different wind-products making the discussion a bit of a mess.*

A. As pointed out by the referee, the use of different wind-products coupled with different parameterizations of gas exchange coefficients can introduce variability in the air-sea CO_2 flux estimates. The other source of variability in the air-sea CO_2 flux is the variability in the field of $p\mathrm{CO_2sw}$. We think there is merit in discussing these two sources separately, and discussed the former in Section 6.1 and mentioned the latter in this section by referring previous works. In order to avoid confusion, we will re-phrase this paragraph (Page 12172 Line 23 - Page 12173 Line 15) as follows:

"In terms of the phase of inter-annual variability, the results from most OBGCMs are consistent with those from diagnostic models demonstrating larger CO_2 efflux during the ENSO cold events and smaller efflux during the warm events. However, OBGCMs appear more sensitive to the ENSO warm and cold events (Table 4 and Fig. 7), particularly during the 1995–1996 cold event and during the 1997–1998 warm event. The reason for the larger ENSO sensitivity in OBGCMs than diagnostic models is yet to be determined but is likely to be attributable to the larger response of the $p\mathrm{CO_2sw}$ field to the change in the wind field (see Section 6.1) associated with the ENSO events. It is also likely that diagnostic models more or less smooth out the variability through the regression analyses of $p\mathrm{CO_2sw}$ as a function of SST and other parameters that are used to correct the implicated under-sampling in observations. To date, two modeling studies have evaluated the skill of the diagnostic method originally developed by Lee et al. (1998) and Loukos et al. (2000) in simulating air-sea CO_2 flux variability over the tropical Pacific (Christian et al., 2008; Park et al., 2010). For both studies, modeled outputs of three-dimensional (x, y, t) fields of SST and other parameters in the tropical ocean are sampled from an OBGCM, and three-dimensional field of $p\mathrm{CO_2sw}$ is reconstructed with a diagnostic model. The air-sea CO_2 flux is then calculated from the reconstructed $p\mathrm{CO_2sw}$ field using the same gas exchange coefficient as in the OBGCM, and this estimate is compared with the fully resolved explicit flux in an OBGCM. The study of Christian et al. (2008) argued that diagnostic models can underestimate the amplitude of interannual variability in the flux from tropical Pacific by up to a factor of 50% depending on the model and grid resolution used. The study of Park et al. (2010) also found that their diagnostic model underestimates 15–20% of the variability, while it overestimates 25–30% for the full fluxes, in the tropical Pacific."

Page 12174 and line 2. *0.20 $\mathrm{PgCyr^{-1}}$ should be 0.17 $\mathrm{PgCyr^{-1}}$*

A. We thank the referee for pointing out the confusing text here. The mean annual air-sea CO_2 flux in the North Pacific extra-tropics north of 18ºN estimated from the diagnostic models of Park et al. (2010) and Sugimoto et al. (2012) is -0.47 and -0.57 PgC yr^{-1} for the period 1990–2009, as mentioned in the text. However, those for the period 2002-2008, i.e., the same period that Nakaoka et al. (2013) estimated the mean flux with their neural network technique, is -0.47 and -0.60 PgC yr^{-1}. These values had not been mentioned in the text. We will describe these values of fluxes for 2002-2008 from Park et al. (2010) and Sugimoto et al. (2012).

Page 12186 lines 1-13. *Wind-speed products are pointed to be one of the elements responsible of the discrepancies between OBGCMs and the diagnostic model. However, these discrepancies are quite low in NP and the Tropical regions in comparison with the South Pacific. Any possible cause for this behaviour should be addressed.*

A. Mean of regionally-integrated air-sea CO_2 flux in the sub-basins of the Pacific Ocean is tabulated in Table 4. The difference in the mean flux between OBGCMs and diagnostic models (diagnostic models – OBGCMs) are +0.05 PgC yr^{-1} in the North Pacific, +0.13 PgC yr^{-1} in the tropical Pacific, and +0.11 PgC yr^{-1} in the South Pacific. When riverine flux are added to the results from OBGCMs, the difference becomes -0.03 PgC yr^{-1} in the North Pacific, +0.11 PgC yr^{-1} in the tropical Pacific, and +0.11 PgC yr^{-1} in the South Pacific. The discrepancies in the tropical Pacific and in the South Pacific are comparable in both cases.

As the referee has pointed out, differences in surface forcing fields including wind-speed is one of the underlying causes of the discrepancies in the flux estimate in a single OBGCM as discussed in section 6.1 as well as among different OBGCMs. Other potential causes of the discrepancies among OBGCMs include ocean model resolution, ocean physical parameterizations, and representation of ocean biogeochemical processes. In the case of diagnostic models, the difference in the wind field coupled with the different gas exchange coefficients as discussed in section 6.1 as well as the difference in pCO_2sw field that is reconstructed from SST and other parameters is the cause of the discrepancies in the flux estimate. However, we will not make a comprehensive analysis to identify the underlying causes of the discrepancies in detail within the context of this synthesis.