Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2017-361-RC2, 2017

Interactive comment on "Mechanisms of the Sea–Air CO_2 Flux Seasonal Cycle biases in CMIP5 Earth Systems Models in the Southern Ocean" *by* N. Precious Mongwe

Anonymous Referee #2

Received and published: 31 October 2017

I think there is the nucleus of a solid paper here, but major revision of the present manuscript is required. The English is generally good although there are some quirks of usage. The Discussion/Conclusion is somewhat repetitive; I would consider merging Discussion and Conclusions.

Major points

(1) I think the methodology could be better explained. Perhaps the logic of equations 1-3 is explained more thoroughly in the previous paper by Mongwe et al 2016, but this critical reference is missing from the reference list. Assuming they mean the paper in Ocean Modelling 106: 90, I will agree that equations 1 and 2 can be derived from

equations 3.2-3.4 in that paper. But the LHS of equation 2, which does not appear in the previous work, is physically speaking, a fairly nebulous quantity. The Takahashi et al 1993 estimate of 0.0423 Kˆ-1 mainly expresses the change in partial pressure due to changing temperature for a given concentration of CO2 ([CO2*]), with a small contribution from the partitioning of DIC among CO2/HCO3-/CO3– due to the temperature dependence of the equilibrium constants. DIC does not change as a result of changes in temperature, except indirectly through gas exchange.

So what we have here is an observed change of pCO2 with changing temperature, convoluted into a change of DIC by application of several highly empirical conversion factors (more about this below), as an estimate of the changes in DIC not attributable to biological uptake/remineralization (and therefore primarily attributable to gas exchange). This in itself might be inoffensive, but I would prefer if its relationship to actual physical processes were better explained. The equations that are taken as a starting point are highly empirical, and we should not invest rearrangements of these with an outsized significance.

0.0423 Kˆ-1 is intended to be an average value for a broad range of ocean conditions, but it is stated to be valid for salinities from 34-36 and temperatures from 2-28 C (Takahashi et al 1993). In the Southern Ocean one will encounter conditions outside, or on the far edges of, these ranges. What are the implications of this for the analysis shown here? This seems like something that could be evaluated. Similarly, the calculations assume a constant Revelle factor, but it should be quite straightforward to calculate Revelle factors from the model outputs, giving a range for the range of environmental conditions characteristic of the study area. The conclusions are probably robust to these assumptions, but I see no reason why this can not be tested.

Finally, isn't the total DIC variability by definition the sum of the various components? So I'm not clear why the temperature driven component should ever be larger than the total. I find equation 3 and the discussion on 304-311 to be the most confusing part. We have an observed rate of change of DIC (which is never actually defined), which one would think would be the sum of the contributions from gas exchange, biological uptake/remineralization and entrainment. But in this case, the index that is considered is that the total is either greater or less than one of these three components (whose physical meaning is nebulous). To confuse matters worse, we have a reference to "the total DIC seasonal cycle (dDIC/dt)" (306-307). Doesn't dX/dt imply an instantaneous rate-of-change that will itself vary over the annual cycle? I really do not understand what is being asserted here. (Also, the text should say something about exactly what sort of discretization was used in calculation of trends, e.g., does delta-X/delta-t for November represent a value for Nov. 1 based on a difference of October and November means, or is it something else? If this is the case, figure axes should indicate that calculated values are for the first of the month and not the mid-month.)

The discussion of entrainment is also confusing and poorly connected to actual physical processes. Equation 4 does not have the units of a flux, but rather of a rate of change within the surface layer. The proper quantity here is not DIC concentration at MLD(T+1) but rather the difference between DIC at MLD(T+1) and at MLD(T). A simple example: we have an 80 m mixed layer that deepens to 100 m over a 24 h period, with an initial linear gradient of DIC concentration from 2.0 mol mˆ-3 at 80 m to 2.1 at 100 m. The total amount of DIC entrained is 0.5 * (20 m) * (100 mmol mˆ-3), for a net vertical flux of 1000 mmol mˆ-2 dˆ-1 or a surface rate of change of 10 mmol mˆ-3 dˆ-1 based on the 'new' MLD as per equation 4. According to equations 4 and 6 in this paper (5 is redundant and unnecessary), this rate would be 210 mmol mˆ-3 dˆ-1.

(2) There is no discussion of the effects of SST and and SSS on CO2 solubility, other than changes in temperature arising from air-sea heat flux. When the mixed layer deepens in autumn, changes in surface T+S caused by entrainment of subsurface water could affect solubility. This may be a second order effect, but it wouldn't hurt to discuss it, and at times the text is ambiguous as to what exactly is being considered. For example, on 395-396 it is stated that "While surface cooling strengthens CO2 solubility in autumn, the concurrent MLD deepening has an opposing effect". So at first

it appears that what is being referred to is the effect of entrainment on solubility via T+S. But when you read a bit further it seems that they are actually talking about DIC entrainment, which of course has no effect on solubility. I think entrainment effects on solubility via SST+SSS should be discussed. If the authors do not wish to address this topic, at least they should modify this passage to make clear that this effect was not considered.

(3) Similarly, the authors could discuss bias in alkalinity estimates as a source of potential error. Upwelled deep waters contain excess alkalinity as a result of accumulated dissolution of biogenic carbonates. Whenever deeper waters are entrained, the Lee formula will generally underestimate the alkalinity. This bias may be small but it is systematic and should be evaluated. (The authors should list in the text the regression coefficients from Figure S1 as well as the correlation coefficient. It is possible to have a very strong correlation but a large systematic bias. In this case the bias is small, and this should be stated in the main text.)

(4) The use of chlorophyll as a proxy is not really explained, when primary production and export production are generally available as model output fields. One might justify this by saying that observations are available only for chlorophyll, but this should be stated explicitly. There are also observation-based estimates of primary production available (see below Terminology).

(5) I would like to see some discussion of the possibility that the apparently greater temperature control in the Pacific sector (259-263) is a real effect that arises from iron limitation. Because terrestrial sources of iron are much greater in Atlantic sector and the western half of the Indian sector (see e.g., Graham et al 2015 DSR I 104: 9; Tagliabue et al., 2012 Biogeosciences 9: 2333), it seems logical that the effect of seasonal biological drawdown on pCO2 would be greater than in the Pacific and in the eastern half of the Indian sector. These regions also overlap the regions where the wind speed and the amplitude of its annual cycle are greatest (e.g., Trenberth et al., 1990, JPO 20: 1742), which will also tend to reduce the influence of biological uptake

relative to temperature. Note also that the wind speed peaks in the spring and fall transition periods, particularly over the Pacific sector (Trenberth op cit their Fig. 4).

Some more questions about methodology

Why use only ten years of model output (124)? The results could be biased by internal variability; the more usual averaging period would be 20 or even 30 years. With a reference year of 2000, this would require using emissions scenarios, which is perhaps a reason not to do it, but the differences among scenarios are very small in 2005-2015 (because the scenarios are constructed precisely around the assumption that there is some inertia in human societies and abrupt changes are unlikely). I think the authors should (a) pick one model, recalculate the results for 1990-2010 and 1985-2015, and estimate the potential error associated with aliasing of internal variability. And (b) if this error turns out to be large, repeat the calculation for the full suite of models.

Why use GLODAP1 data for DIC? GLODAP2 has been available for almost two years. There are also more up to date data sets available for MLD. ARGO has made the Southern Ocean a much less undersampled region than it was in 2004. (The abbreviation MLD is not defined at first use (117)).

Some stylistic advice

Sir Peter Medawar is said to have advised authors of scientific papers to never begin with "a resounding banality". We all do this. "CO2 is one of the most important greenhouse gases." "The ocean is a critical component of the climate system." But it's actually good advice. The abstract to this paper starts weak, but ends even weaker. The final sentence is very poorly worded and bordering on incomprehensible. What is in between is generally good. But if we want people to read the paper, we have to end the abstract in a way that generates interest.

Similarly, the main text (Conclusion) ends by saying that "the inability of the CMIP5 ESMs to resolve CO2 biological uptake during spring might be crucially related to the

sensitivity of the pCO2 to temperature". This again leaves the reader at a loss as to what exactly the authors are trying to say, at a critical point. Is the problem here model errors in SST, or the way that solubility is calculated from T+S? Almost certainly the former, but it's hard to tell from the way this is phrased.

Terminology

I think the authors should acknowledge that the FCO2 data product is not really 'observed' in the sense that pCO2 is. I think they should compare modelled and observed pCO2, and then discuss what this means for modelled estimates of CO2 flux, without referring to the Landschuter FCO2 estimates as observations. CO2 fluxes in models and data products like this are actually quite different conceptually. When you estimate CO2 flux from observed pCO2, the errors in the flux are a linear function of errors in wind speed (or uˆ2, assuming a quadratic parameterization) and the piston velocity. In numerical models, pCO2 and DIC self-regulate to dampen these errors when monthly averaged fluxes are considered, e.g., if both the wind speed and the DIC are too large, the enhanced outgassing flux will reduce the pCO2 and DIC error. Higher wind speed or piston velocity will tend to drive pCO2 towards atmospheric, and not necessarily towards the 'correct' value if over- or undersaturation exists, so there is no straightforward way to correct for this difference. But I think that the authors should acknowledge that it exists, and that in comparing modelled and 'observed' fluxes they are to some degree comparing apples and oranges. (With regard to point (4) above, if there is no observed primary production, there is no observed CO2 flux either: both of these are extrapolated from the primary observed field using models of unknown accuracy.)

I think it is stretch to call the Takahashi et al (2009) flux estimates a 'dataset' (103). I think it would be more accurate to refer to CO2 flux estimates of Takahashi et al, or CO2 flux estimated by the methods of Takahashi et al (as do Takahashi et al themselves). Note that the description of this data product (104-106) refers only to pCO2 measurements and says nothing about how flux was estimated.

Don't refer to climate model projections as predictions (15, 85, 496). Similarly, I would remove the term "mode" (60, 488) as the intended meaning deviates from the usual meaning of this term in climate research. (Rule of thumb for young authors: don't coin neologisms without some compelling reason, and don't appropriate existing terms and give them new meanings.)

Similarly, the intended meaning of the word "phase" is unclear, and the usage of this term is not usual. (The same is true of "coherent": see below 252, 339.) This begins in the Introduction, where it is stated that the models "disagree on the seasonal cycle of CO2 flux and they are out of phase with observations" (58). Does this mean the phasing of the seasonal cycle in the models differs from that in the observations? This would seem to be the most plausible explanation, but the present wording does not communicate this (or anything else) effectively. The later invocations (e.g., 231-232, 260-266) are better but could still benefit from the being a bit more explicit about exactly what is being discussed. Instead of referring to the models and observations being out of phase with each other, one could refer to actual physical processes: e.g., that the modelled seasonal cycle of SST is out of phase with that in the WOA climatology. Assuming, of course, that this is actually an accurate characterization. Or is what is meant here simply that the models do not reproduce the phase of the annual cycle accurately? When we say that two cycles are out of phase, we usually mean offset by 180 degrees.

Similarly, "the onset of primary production" could be better phrased. Primary production is not zero in the winter, although clearly seasonality is large in high-latitude environments. I would probably say something like "the increase in biological uptake in spring". More importantly, primary production per se has little or no effect on DIC: in principle it could be entirely respired within the surface layer. It is true that net community production is usually positively created with primary production, especially in highly seasonal high latitude environments. But since the focus of this paper is DIC and pCO2 I don't see what is gained by referring to primary production rather than

simply biological uptake (e.g., 482).

Details

"ingassing" is variously spelled as ingassing, in gassing, in-gassing (e.g., 62, 64, 205)

Model names appear to be misspelled in places (e.g., 203, 233, 283, 319, 321, 401, 431, 466, Figures 3, 4, and 6)

16 remove the phrase "used a specialized diagnostic analysis"

27 change "contradicting" to "counteracting"

43 change "oceanic" to "ocean"

43-44 "The century scale evolution of the Southern Ocean CO2 sink is expected to change as a result of anthropogenic warming, however the anticipated change is still disputed." The century scale evolution of the Southern Ocean CO2 sink is expected to change as a result of anthropogenic warming, however, the sign and magnitude of the change is still disputed.

52 "especially during the winter season" especially during winter

63 missing "Ref"

62-63 "with a weaker in-gassing or even outgassing state in winter" with weaker in-gassing or even outgassing in winter

64 add a comma after "weakens"

68 "The increase of sea surface temperature (SST) with summer weakens the surface CO2 solubility" The increase of sea surface temperature (SST) in summer reduces surface CO2 solubility

86-89 "Efforts to .... coupled simulations". Break this into two sentences at "however".

92 "exploring the mechanisms of the observed model biases" exploring the mecha-

nisms underlying the model biases

114 delete "the" before CO2SYS

137 delete "in the Southern Ocean"

114 delete "the" before F_DIC

175 Fig. 6 appears to be cited out of order here

236 not clear what "ensemble" means in this context

237 not clear what "decadal" means in this context

238 "In the Sub-Antarctic zone for all three basins, observed FCO2 show a weakening of CO2 uptake during winter (less negative values in JJA) with values close to the zero flux at the onset of spring" In the Subantarctic zone, in all three basins, there is a weakening of CO2 uptake during winter (less negative values in JJA) with values close to zero at the onset of spring

242 "In the Antarctic zone, the observed FCO2 seasonal cycle is similar in all three basins (Fig. 3d-f), possibly resulting from the limited number of observations." In the Antarctic zone, the observed FCO2 seasonal cycle is similar in all three basins (Fig. 3d-f). This apparent uniformity may result from the limited number of observations, and additional interbasin differences may arise as new observations are collected. (I am guessing at the intended meaning here.)

252 change "coherent" to "consistent"

255 "though they agree in the phasing, they have some differences in magnitudes" although they agree on the phase, the magnitude varies considerably

267 "Group A shows an overestimation of the CO2 uptake, while group-B shows an underestimation of CO2 uptake with respect to observations. This disagreement is accompanied by a large standard deviation, showing some inter-model differences in

magnitudes (Fig. 2d-f)." Group A shows an overestimation of the CO2 uptake with respect to observations, while group-B shows an underestimation. The large standard deviation indicates considerable differences among the models (Fig. 2dâĂŘf).

275 "CMIP5 models have a general positive bias against observations during summer and/or autumn with the exception of group-A models in the Sub-Antarctic zone." CMIP5 models have a generally positive bias relative to observations during summer and/or autumn, with the exception of group-A models in the Subantarctic zone.

283 change "an" to "the", delete "the" before "MPI-ESM"

284 change "harmony" to "consistency"

284 How can the observations be biased? Relative to what?

286-287 I would just delete everything after "representative".

293-295 This seems to imply that the sign change (-ve to +ve) is the same in spring and fall.

299 not clear what "amplitude" means in this context (see also 466)

302 delete "the" before "pCO2"

319 "It shows group A models (HadGEM-ES, NorESM2 and MPI-ESM) at the bottom of Fig. 6, indicating that these models are mainly DIC driven." It shows that group A models (HadGEM-ES, NorESM2 and MPI-ESM) are mainly DIC driven (Fig. 6). (I would double check all the model names here)

332 change "rich-CO2" to "CO2-rich"

339 "In the Antarctic zone CMIP5 models are largely coherent at the onset of MLD deepening (February), however significantly variable at the winter maximum depth." In the Antarctic zone, modelled MLD is quite consistent across models at the onset of MLD deepening (February), but more variable at the winter maximum.

371 "models cluster together in all the basins of the Sub-Antarctic zone with the only exception of CESM1-BGC in the Atlantic" models cluster together in all basins of the Subantarctic zone, with the exception of CESM1-BGC in the Atlantic

374 change "scattered" to "divided"

386 delete "preceding"

391 remove ()'s around "biased"

404 add "convective" before "mixing"

408 change "is contrary to" to "differs from"

410 I would delete "add to the entrainment rate and"

416-424 this passage seems repetitive and could be shortened.

426 delete "flux"

439 "This is important because though annual means are useful" This is important because, although annual means are useful

455 "In the Antarctic zone, all selected CMIP5 models are in general agreement, as well as with observations on the dominant role of DIC in regulating FCO2 seasonal cycle" In the Antarctic zone, all of the CMIP5 models analyzed are in general agreement, and consistent with observations, on the dominant role of DIC in regulating the seasonal cycle of air-sea CO2 flux

458 "models and observations show little inter-basin differences in the seasonal cycle of FCO2, suggesting that mechanisms driving FCO2 are less localized" models and observations show weak interbasin differences in the seasonal cycle of FCO2, suggesting that mechanisms driving FCO2 are less spatially variable

461 "emergence of basin specific spatial characteristics of FCO2 might be inhibited by lack of observational coverage" basin-specific features may be obscured by incomplete

data coverage in the observational data products

468 delete "analyzed"

469 change "proposed" to "hypothesized"

469-470 "As evident in Fig. 8, bottom or subsurface DIC could not be the main driver of these amplified DIC amplitudes because the Antarctic zone shows comparable lower entrainment fluxes (Fig. 8)." This assertion does not make a lot of sense, and the figure legend and caption are ambiguous: I can't really tell which panels are for which regions. Anyway, what exactly are they trying to say here? That entrainment is shown to be a second order effect because rates are comparable in the Antarctic and Subantartic? The text as written makes it sound like the former is being contrasted with itself. The authors need to be clear about what exactly is being asserted here, and make sure that the assertion is actually supported by the data cited.

477 "this is consistent with Rosso et al., 2017 findings" this is consistent with Rosso et al. (2017)

473 "the seasonal cycle of chlorophyll in Fig. 9 show coherence (symmetric $\sim$ negative correlation) with the analyzed rate of change of DIC" the seasonal cycle of chlorophyll is anticorrelated with the rate of change of DIC (Fig. 9)

484 "may have an important role as suggested by Rysgaard et al., 2011 and Rosso et al 2017 and they should be investigated as part of a future study" may have an important role as suggested by Rysgaard et al. (2011) and Rosso et al. (2017) and should be investigated further

491 add "flux" after "heat"

493 "influence the surface heat regulation" I can't tell what this means

495 "Therefore these analyzed temperature bias pose an important predicament with respect to our ability to predict future earth system changes, particularly the carbon

cycle. We propose this bias as an important consideration to the model developing community as it relates to future biogeochemical and CMIP ESM development." These temperature biases pose an important problem with respect to our ability to model future changes in the carbon cycle. We suggest that diagnosing and reducing this bias should be a priority for future model development.

506 not clear what "exaggerated" means in this context

510 change "on" to "of"

522 "We find that though some models exhibit comparable chlorophyll magnitudes with observations" We find that, although some models exhibit chlorophyll concentration comparable to observations (see also 526-527)

535 delete "the presentation of"

602, 687 references are incomplete

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-361, 2017.