We are indebted to all the reviewers for all the comments. Some have helped us avoid serious mistakes (like the nasty units prefix error not noticed in the first round of reviews). So helped us focus the message even if we disagree with them. **Generally the discussion showed the sub-communities (air-sea interaction, geochemistry etc) do not always speak the same language**. We believe the discussion shows that the question of how the choice of gas transfer velocity formula influences calculated flux values, strangely a gap in the body of published knowledge, does deserve its own paper. At least this is how we see it.

We thank reviewer #2 for his second round of comments.

1)SOCAT is not a climatology and should not be referred to as such (see the Abstract). Whatever was done with the FluxEngine is fine, but when talking about SOCAT it should not be called a climatology.

This has been changed everywhere from "climatology" to "database" with the exception of lines 116-117 where we have added a reference (Goddijn-Murphy, et al., 2015) explaining how a climatology is calculated from the SOCAT data.

L28: 'We also compare the available *p*CO_2 **climatologies** (Takahashi and SOCAT), *p*CO_2 discrepancies in annual flux values of 8% in the North Atlantic and 19% in the European Arctic. The seasonal flux changes in the Arctic have inverse seasonal change in both **climatologies**, caused most probably by insufficient data coverage, especially in winter.'

Change to: 'We also compare the available *p*CO_2 datasets (Takahashi and SOCAT), *p*CO_2 discrepancies in the annual fluxes values of 8% in the North Atlantic and 19% in the European Arctic. Seasonality of the flux changes in the Arctic are opposite to one other in both datasets, most likely caused by insufficient data coverage, especially in winter.'

L114: The SOCAT databases have been converted to climatologies using methodology described in Goddijn-Murphy, et al. (2015).

L250: ….(using both SOCAT **datasets** results in a larger sink in summer and smaller in winter compare to Takahashi),…

L253: … with Takahashi and SOCAT dataset derived climatologies resulting…

L315: Although, using both Takahashi climatology and SOCAT *p*CO_2 **dataset** (Fig. 8)….

L317: This may have been caused by slightly different time periods of the **datasets** (SOCAT is more recent).

L337: We compare Takahashi and SOCAT *p*CO_2 **datasets** finding that…

L340: The seasonal flux changes in the Arctic have inverse seasonal change in both **datasets…**

L599 and 814: Figure 8. Comparison of monthly values fluxes of air-sea CO_2 fluxes calculated with different *p*CO_2 **datasets**…

L388: **Goddijn-Murphy L., Woolf D. K., Callaghan A. H.: Parameterizations and Algorithms for Oceanic Whitecap Coverage, J. Phys. Oceanogr., 41, 742-756, 2011.**

2) First review and Abstract - The authors misunderstood my criticism of the word uncertainty. I do not mind the concept of comparing different k parameterizations. I mind use of the word uncertainty. The fact that parameterizations result in fluxes that are 10% or 50% different does not mean that is uncertainty in the k value or the resulting flux. The uncertainty in the k param. has to do with the methods used to make the measurement. I would like it better if the difference in fluxes related to choice of k was called something else, rather than uncertainty (for example, range or difference).

We believe there is a misunderstanding between the views of the reviewer and our use of the word uncertainty. We use it in the metrological meaning such as in this Wikipedia derived definition of Measurement_uncertainty "*In metrology, measurement uncertainty is a non-negative parameter characterizing the dispersion of the values attributed to a measured quantity*." Please take note that in statistics dispersion may be several things, for example standard deviations (either one or several) but also a range of values, which is exactly as we use the term.

3) In my initial review, I made a comment about Figure 7 as well as Figure 8. The authors misunderstood. Point about fig 7 - This figure does not add anything over Figure 6. Why do you need both figures? Point about fig 8 - This is an interesting result and should be described in more detail. It is fine that the authors think this is outside of the scope of the current manuscript. I just want to make it clear that there were 2 seaparate comments. Here I would like the authors to address only the comment about Figure 7.

Figure 7 is used to visualize the uncertainty (in the meaning of range) of values while Figure 6 shows the actual annual flux values. Even if the former clearly may be derived from the latter they convey different messages. As concerns Figure 8, we agree that removing it and the discussion of the difference between Takahashi and SOCAT datasets is possible. However, if were are not allowed to comment on the differences between the results resulting from the use of the two datasets, we should also remove all mentions of one of them (otherwise what would be the point of using two if they cannot be compared?). However in our opinion the readers of the paper would lose out not obtaining the information of *additional* uncertainty coming from the choice of the datasets.

4) Again, I think it is strange to cite workshops here. It doesn't mean the point of the authors is not valid. I leave it to the editor to decide, but I think the paper reads less professionally when workshop proceedings are cited. There should be substantial support in the literature for the points made by the authors other than 'these folks said it is important at a workshop'.

This is not a random, unimportant source. The SOLAS Open Science Conferences are possibly the most important meetings of the air-sea interaction community. They involve special Discussion Sessions convened to give answers to the most important questions facing the community. The sessions have their official Rapporteurs and the results are published on the SOLAS International website and in it bulletin. Here we had a session convened by the very authors of most of the formulas used in the community with a large representation of the community (and authors of other formulas) present. And the subject of the discussion session was exactly which parameterizations should the community use with a conclusion listing the ones which are consistent with the data we have and indistinguishable within experimental uncertainty.

In any case, the OS journal rules allow “grey literature” where no peer-reviewed sources are available. Even if we concede that this source is “grey literature”, it is still at present the best source for the statement that the three parametrization are within present measurement uncertainty.

Anyway, we removed the text about the discussion session while leaving the citation to this. (L264)

=====

We thank Reviewer#3 for his comments that will be accounted for to improve the manuscript. **The manuscript was re-checked for correctness of the language, so we do not write point-by-point of each amendment, corrected typos, because it have not changed it the context of a sentence.** All postings and comments by the reviewer the bulleted were introduced and used to the proposal. Below the point-by-point are only those changes and responses to comments, which are required to develop, supplement and responses to comments and resulted in a change in the context of the sentence.

Major comments:

Although it is useful to quantify uncertainty attributed to the gas transfer formulation, there is very limited new information for the community. Given that the authors have done the computation for the whole global ocean, a more detailed investigation on regional features would be very useful, i.e., in addition to the North Atlantic. The authors state that the uncertainty could be substantial in regions of high wind speed. Is this only valid for the North Atlantic or also for other strong wind regions such as the Equatorial Pacific and the Southern Ocean?

We chose the North Atlantic and the European Arctic in scale like it is due to the fact that we are limited, to the results of the measurements, from the satellite and also because of the most intense CO_2 sink area on the per unit area basis (Arctic), cold temperature, strong photosynthesis, high wind speeds and high alkalinity. The importance of North Atlantic for ocean carbon fluxes comes from the fact about 80% of global deep waters are produced there. We also study the North Atlantic and the Arctic sea because this is the part of a PhD thesis one of the author (Iwona Wrobel).

To compute uncertainty highlighted in this study, one would only need to provide surface pCO_2, SST, SSS, wind speed, and atmospheric CO_2 fields. In addition to the above point, I think the authors should also consider studying how this uncertainty would evolve in the future, regionally and temporally. To do this, one can use the CMIP5 model outputs, which

are publicly available. Questions such as will uncertainty due to atmospheric CO_2 concentration dominates the uncertainty due to wind speed formulation, etc. has not been studied previously and can be addressed here.
In which regions, and when in the future, the uncertainty become significant thus provides some recommendation for improving the monitoring network to the observational community.

The analyses the reviewer wants is preparing the North Atlantic wind climatology from the CMIP5 model ensemble (the modelling work for IPCC AR5). Such analysis would be overly burdensome and actually pointless.

a) <u>Why I believe this recommendation is overly burdensome?</u> We would need to prepare several climatologies from the output data of almost 30 models used in the CMIP5 ensemble. Several, because we do not know the future course of our carbon emissions and therefore the modelling is done for different future forcing assumptions. At least three would be needed: RCP2.6, RCP4.5 RCP8.5 (the lowest, middle and highest emission scenarios). That would be needed to be done for each month on a 1x1 deg net that FluxEngine uses (which is a finer net than most of the CMIP5 models use so it would involve a lot of interpolation work, different for almost every model (see yourself: https://verc.enes.org/data/enes-model-data/cmip5/resolution). We never did this kind of work and we did not plan to do that in foreseeable future. First of all we would need to purchase additional hardware just to store all of the modelling output. I estimate the amount of work in terms of months and the cost in thousands of Euros (Should I add that I have no guarantee that the operators of FluxEngine could not agree to such addition to it making it necessary to make a local copy of the engine and all the databases it uses and run our version locally). All this would be an absolute waste of time and money because…

b) <u>Why I believe the effort would be absolutely pointless</u>? Because it has been done and the results are published. Actually in the the very IPCC AR5 WG1 report the modelling was made for. The results are presented in Fig. 12.19:

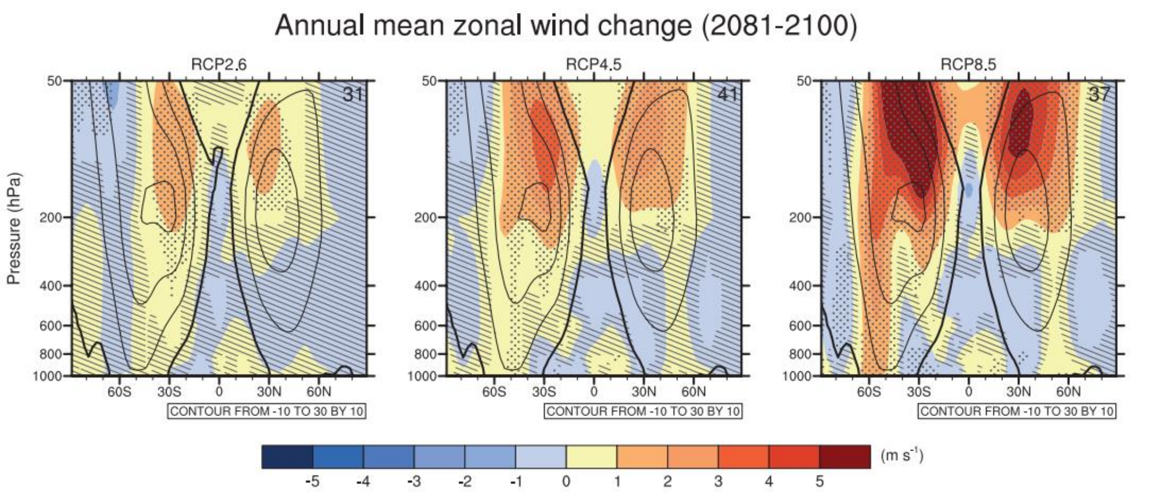


Figure 12.19 | Coupled Model Intercomparison Project Phase 5 (CMIP5) multi-model ensemble average of zonal and annual mean wind change (2081–2100 minus 1986–2005) for, from left to right, Representative Concentration Pathway 2.6 (RCP2.6), 4.5 and 8.5. Black contours represent the multi-model average for the 1986–2005 base period. Hatching indicates regions where the multi-model mean change is less than one standard deviation of internal variability. Stippling indicates regions where the multi-model mean change is greater than two standard deviations of internal variability and where at least 90% of models agree on the sign of change (see Box 12.1).

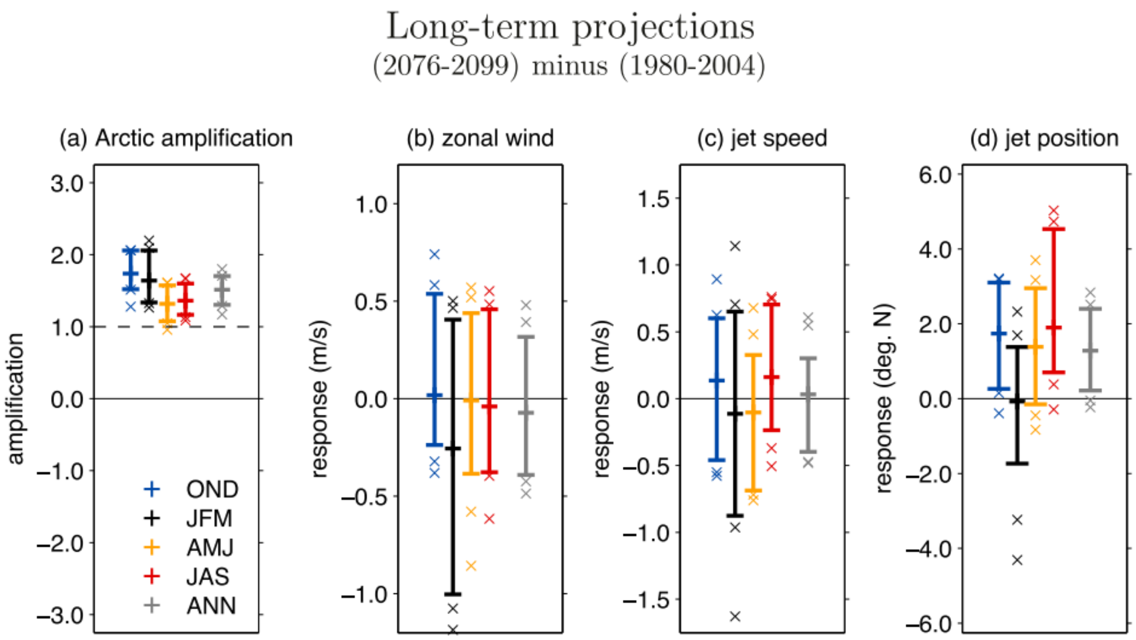
The problem is there are no robust predictions for the zonal winds (most of the wind energy and variability is in the zonal direction) at the sea-surface level for the North Hemisphere (the bottom of the graphs to the right of 30 N). For RCP2.6 and RCP4.5 all the surface wind results for the extra-tropical North Hemisphere are hatched, which means the multi-model mean is change is less than one standard deviation (sigma) of internal variability. Only for RCP8.5 there are some results over one sigma but still under two sigmas. This means the confidence level of over 66% but under 95%. This is not anyone would call robust predictions (neither statistically significant). Please take note that the ensemble average wind speed change in everywhere smaller than 1 m/s (with either negative or positive sign).

Therefore, it is no wonder that **everything** the IPCC report has to say about North Hemisphere extra-tropical surface winds is:

> 12.4.4 Changes in Atmospheric Circulation
>
> [...] In the NH [North Hemisphere], the response of the tropospheric jet is weaker and complicated by the additional thermal forcing of polar amplification (Woollings, 2008). Barnes and Polvani (2013) evaluate changes in the annual mean mid-latitude jets in the CMIP5 ensemble, finding consistent poleward shifts in both hemispheres under RCP8.5 for the end of the 21st century. In the NH, the poleward shift is ~1°, similar to that found for the CMIP3 ensemble (Woollings and Blackburn, 2012). [...]
>
> Although the poleward shift of the tropospheric jets are robust across models and likely under increased GHGs, the dynamical mechanisms behind these projections are still not completely understood and have been explored in both simple and complex models (Chen et al., 2008; Lim and Simmonds, 2009; Butler et al., 2010) [...]
>
> In summary, poleward shifts in the mid-latitude jets of about 1 to 2 degrees latitude are likely at the end of the 21st century under RCP8.5 in both hemispheres (medium confidence) with weaker shifts in the NH and under lower emission scenarios.

It is easy to see there are no predictions about average wind speed change, only a change of the jet stream position in RCP8.5 by about 1 deg northward (this is the border between blue and yellow at about 40N in the bottom of the right graph above).

However, this is not the end of the story. It is getting worse: we need to make a climatology not for the whole North Hemisphere but only North Atlantic and for each month, instead of the whole year. It is well known that the smaller area and shorter time period becomes, the lower id the ratio of climate signal to (internal variability) noise. So we can expect not to have even one sigma anywhere in the results. And I do have support also for this statement. There was on paper on the North Atlantic results of the CMIP5 ensemble, Barnes and Polvani, 2015, "CMIP5 Projections of Arctic Amplification, of the North American/North Atlantic Circulation, and of Their Relationship", Journal of Climate, 28, 5254-5271, http://dx.doi.org/10.1175/JCLI-D-14-00589.1 . It was based on 27 CMIP5 models and the RCP4.5 emission scenario. The results for the 2076-2099 period minus 1980-2004 are summed up in Figure 4. I copy here the relevant part of it:



The colors are seasons (explained in the left panel) and the vertical bars denote the 10th-90th range. The graph most relevant here is "(b) zonal winds".

It is obvious that no robust predictions for average seasonal (or annual) wind speeds can be made for the North Atlantic sector. The "x" sign show outlier models and we have as many predicting increase as decrease of zonal wind speeds.

To sum it up, we believe that making climatology from this ensemble would be even worse than pointless. It would be misleading because in reality there are no robust predictions. At first glance it would seem that using the climatology for 1980-2004 would do (it is not possible to falsify the null hypothesis that 2076-2099 zonal winds will be identical) however it is not correct: it is not that we know the winds will not change. The truth is we do not know even the sign of its future change. (Which is no surprise as we do not know even whether NAO will be on average more positive or negative in a warmer world: not only models disagree but theory disagrees with the models which do not include all the necessary processes – there exists a large literature on that).

And finally, all the above discussion has nothing to do with the subject of the paper which is the uncertainty coming from the choice of gas transfer velocity formula for existing data. Repeating the analysis for the end of 21th century would have also to involve using predicted data for SST, surface salinity, and partial CO_2 pressures in both the sea and atmosphere (for every month in the year). We believe they are even less constrained than the future winds. Predicting all the parameters for year 2099 (and therefore the CO_2 air-sea interaction fluxes) would be a great feat (if it were possible!), worth most probably a separate Nature or Science paper.

We did try to do one test, changing the southern border of the North Atlantic region from the arbitrary

30 N position (moving it north- and southward). This was meant to check how the results depend on the choice. This could be treated as a proxy of moving the jet stream area (the only semi-robust prediction of the CMIP5 ensemble), because it changes the ratio of areas with stronger and weaker winds (however we admit it is a kind of poor man's proxy). The changes caused the average fluxes (that per surface unit) to change by about 2% with a move of the southern border by one degree (with increase when the border was moved northwards, as expected). At the same time such changes caused the ratios of fluxes for different wind speed powers to differ one order of magnitude less, that is much below the 1% precision we reported them with (with decreasing difference as the southern border moves North which actually supports our conclusion about the North Atlantic). To be more precise a difference of 10,0% would become 10,1% with the south border moving from 30 N to 31 N. In other words, a change of the jet stream by about one degree should not affect our conclusions as given in the manuscript. However, we chose not to add this information to the manuscript because of the circumspect and indirect way it has been derived but we hope it makes the reviewer more comfortable.

L210-214: I am not entirely convinced by this analysis and I think that this is not entirely true (e.g.based on Fig 4d and 3d). From my visual inspection, the spatial variations in Fig.4 actually resemble very much Fig. 2, and therefore, I would assume the difference shown in Fig. 4 to a large extent stems from the amplitude of delta pCO2 values, rather than the wind speed. Assuming that atmospheric CO2 has very little spatial variability, the delta pCO2 patterns will be very similar with Fig. 2.

It can't be right because the delta pCO2 difference is identical in the case of every *k* parameterization. Therefore the *ratio* of two parameterizations is caused only by *k* formula differences (the delta pCO2 part cancels out if we calculate *ratios* of different flux parameterization values). And *k* in all the parameterizations used is parameterized solely by wind speed. By the way this is exactly why we normalized all of the parameterizations by one of them!

Some of the existing analysis can be illustrated in a simpler and compacted way, such as put values from Figs. 6 and 7 into one single table (Line 234-237, the values from the global estimates, which are missing from Fig. 7 should be added to this table as well). In addition, to better illustrate regions where different wind parameterization could be important for air-sea fluxes estimation, I recommend showing a map of standard deviation/variance of gridded fluxes computed using the different equations.

Minor comments:

L30-32: I suggest revising this sentence into something like: "The seasonal flux in the Arctic computed from the two climatology data sets are opposite to one another, possibly due to insufficient spatial and temporal data coverage, especially in the winter."

L31-33: 'We also compare the available $p\mathrm{CO_2}$ datasets (Takahashi and SOCAT), $p\mathrm{CO_2}$ discrepancies in the annual fluxes values of 8% in the North Atlantic and 19% in the European Arctic. Seasonality of the flux changes in the Arctic are opposite to one other in both datasets, most likely caused by insufficient data coverage, especially in winter.'

L55: Studying the rate of the ocean CO2 sink…
L55: … especially its long-term trends, one needs …
L60: parameterization
L83 … European Space Agency funded OceanFlux …
L98: calculations.
L98 users can choose
L101: configure them in ..
L111: Bakker et al.
L137: the solubility unit should be g m-3 microatm-1
L195: NIghtingale et al. (2000)
L195: Takahashi et al., (2009)
L175: in water])
L205: are shown
L219 "… by up to 30% …"

All the above mentioned errors have been corrected.

L85: Maybe better to replace "opening sourced" with "publicly available"

We used the phrase of the very authors of the software. However we see the point. To make the text both precise and communicable we changed the phrasing to "became publicly available under an open source license".

L86: The software became publicly available in March 2016 under an open source license, but at the time we started this study we did not have more information about it than is included in the paper describes a set of tools (Shutler et al., 2016).

L99: how is ice age affect air-sea flux? clarify. Do you mean sea-ice area?

We mean % of ice cover. Thanks for catching that. We corrected that removing the ice age.

Just to explain that, FluxEngine uses the Takahashi et al 2009 approach as described in Shutler at al. 2016.

L100 ….% of sea ice cover from monthly model data ECMWF air pressure, whitecapping (Goddijn-Murphy et al., 2011), from monthly climatology as pCO$_2$, SST, salinity) and configure them in a various way…

L100: clarify what is meant by "whitecapping"?

Whitecapping refers to the steepness-induced wave dissipation in deep water which some air is entrained into the near-surface water, forming an emulsion of water and air bubbles (foam) that appears white. It occurs when the velocity of individual water particles near the wave crest exceed the phase speed of the wave, causing the front face of the wave to became too steep and "break". Whitecapping is an essential process for air-sea gas exchange. Whitecapping dissipation rates have been estimated from observations, using the equilibrium range theory developed by Phillips (1985), and are well correlated with both wind speed and acoustic backscatter observations.

FluxEngine uses the whitecap parameterization of Goddijn-Murphy L., Woolf D. K., Callaghan A. H.: Parameterizations and Algorithms for Oceanic Whitecap Coverage, J. Phys. Oceanogr., 41, 742-756, 2011 for the rain-driven flux calculations. The author list makes it clear it has been created just for the purpose of air-sea gas fluxes. We added this reference to make it clear.

L100: ….% of sea ice cover from monthly model data ECMWF air pressure, whitecapping (Goddijn-Murphy et al., 2011), from monthly climatology as pCO_2, SST, salinity) and configure them in a various way

L388: **Goddijn-Murphy L., Woolf D. K., Callaghan A. H.: Parameterizations and Algorithms for Oceanic Whitecap Coverage, J. Phys. Oceanogr., 41, 742-756, 2011.**

L117. clarify what is meant here by "preprocessed"?

In that meaning pre-processed, means that before put it into FluxEngine software, the FluxEngine authors pre-processed this into the format required by the FluxEngine (using *in situ* tool). We used this to make it clear this was not done by us, the end users of this toolset.

L133: What is meant by "sea state"?

Sea state in oceanography means general condition of the free surface on a large body of water-with respect to wind waves and swell-at a certain location and moment. We use that term identically to Goddijn-Murphy et al. 2011 (see above) as the set of conditions whitecap is a proxy of. There is a large body of air-sea interaction literature that uses it this way. See for example the very title of Wu, J.,: *Oceanic whitecaps and sea state.* J. Phys. Oceanogr., 9, 1064–1068, 1979, one of the classical papers in the field.

L184: do you mean "Ocean Flux GHG project"?

We mean OceanFlux Greenhouse Gases (GHG) Evolution project funded by ESA, the one being part of the title of the paper describing the methodology we cite (Goddijn-Murphy et al. 2015).

L197: replace "all-season" with "annual"

We used all-season map in the contest maps of every season. Therefore we changed it into "seasonal maps".

L200: The area, as a whole, is a carbon sink but even the seasonal maps show that some regions close to North Atlantic Drift and East Greenland Current are net sources.

L269:271, again not clear to me where are 33% and 50% quantities derived from?

From the FluxEngine calculations. We report here a result derived directly from the flux calculations for the described regions.

L257-258, where does the 9% value come from? Fig. 7 only shows values for North Atlantic and Arctic.

Similarly this is the results we report, coming directly from flux calculations, not from the figures being illustration of the calculation records.

Double check the units throughout the text. I believe most of the [Tg] should be [Pg].

Yes!!! Thanks for catching it. It seems reviewers #3 are useful (no one noticed this obvious error!). We corrected that everywhere and also added a comment on the choice of unit at the end of Section 2.1.
L232, 234,251,252,254,255,587,590,760,774

Figs caption: would be useful to state that positive values represent outgassing and negative otherwise, when showing the air-sea flux of CO2 values.

This is correct. In the literature of the theme it is determine that the air-sea fluxes uses the wording: source by a flux of negative (downward) and sink in the positive flux (upward). By using such terms, it is known what we had in mind. We added a sentence explaining it at the end of section 2.1.

L131: We use everywhere the convention of sources (upward ocean-to-atmosphere fluxes) being positive and sinks (downward atmosphere-to-ocean fluxes) being negative. We give all results of carbon fluxes in the SI unit of Pg (numerically identical to Gt).

Fig1. the color bar can be improved, e.g., use ranges from -20 to 20 since there seems to be no grid points showing values of -40.

You have right, but the range of color bar from -30 to 20 is better to show the fluxes.

Fig. 1. Add zero contour lines to distinguish uptake and outgassing areas.

There is no necessary add zero contour lines, because zero is colored by green, so the differences are easy to see.

Fig. 5 shows that the differences in wind speed parameterization is largest between McGills and Nightingale parameterizations. Based on this, it would be more informative to show the difference between fluxes computed based on these formulations in Fig. 4, as an indicator for upper uncertainty range.

Well, then a reviewer could ask us why we over emphasize the differences choosing the outermost parameterizations. Every choice has its downside. We believe our choice shows the difference well enough. An interesting twist is that both the formulas have been widely used and both… come from the same author.

Fig. 6a and 6b are identical, and the units in x-axes are wrong as well.

Corrected.

Fig. 7. Clarify what is meant by normalized here. Shouldn’t it be unit-less instead of Tg/year?

That’s correct. Figure 7 has no units. The error in the caption has been corrected (units removed).

L597, 806

We also added the definition of "normalized" (divided by it) in the manuscript text where this figure is discussed.
L234: Figure 7 shows the same data "normalized" to the N2000 (divided by value), the parameterization results in a lower absolute flux values to visualize the relative differences.

Fig. 8: units should be in Pg instead of Tg.

Corrected!

Last paragraph of introduction: In addition to formula in the gas transfer velocity, the selection of wind product also contribute to the uncertainty, therefore would be good to mention this, citing the study by Gregg et al., 2014 (Ocean Modelling: Sensitivity of simulated global ocean carbon flux estimates to forcing by reanalysis products.)

A mention of this (and the citation) has been added at the end of last but one paragraph of the Introduction.

L388: added **Goddijn-Murphy L., Woolf D. K., Callaghan A. H.: Parameterizations and Algorithms for Oceanic Whitecap Coverage, J. Phys. Oceanogr., 41, 742-756, 2011.**

L404: added **Gregg, W. W., Casey N. W., Rosseaux C. S,: Sensitivity of simulated global ocean carbon flux estimates to forcing by reanalysis products, Ocean Modelling, 80, 24-35, doi: 10.1016/j.ocemod.2014.05.002, 2015**.