We thank both reviewers for so many very helpful comments and appreciate for constructive feedback on the paper, which has helped to improve the manuscript. All their comments have been addressed and changes have been included the revised version the manuscript (see below).

Because some of the comments overlap, we would like to start with responding to the two most important topics rose in **both reviews**: the purpose of the paper and the toolset used.

The purpose of the paper is to constrain the uncertainty resulting from the choice of the gas transfer velocity (*k*) parameterization in the case of the North Atlantic. This is a region best covered with measurements and one with stronger winds than average over the world ocean. We started the study convinced the relative uncertainty will be larger than elsewhere. However it turned out it is smaller. This is a previously unpublished finding, important not only for the ocean carbon budget but especially to the gas transfer velocity community. The North Atlantic is a place where multiple experiments aiming at containing the *k* parameterization were performed. Knowing (thanks to our results) that typical North Atlantic winds are the environment least suited for choosing between parameterizations of different wind speed power should be taken into account when planning future experiments of this kind. The feedback we had presenting early results at several meetings (EGU and SOLAS conferences, a SOCAT workshop and a grant meeting involving most of the key European researchers in the field). The feedback was very encouraging. It is the feedback which made us include some of the text the reviewers had comments on (like the Arctic seasonality which was treated by us as a curiosity until we heard feedback at the PICO session at EGU). The discussion we had at the meetings, in one case so much we thanked its author with a citation (see also below). Also the manuscript revision showed to us that the gap in literature (lack of papers which show comparisons of resulting fluxes for multiple parameterizations, especially the recent ones) makes the manuscript even more relevant.

The other subject raised in many comments is the FluxEngine toolset. It has been developed by researchers we cooperate with but during a previous project we were not a part of. A paper describing the toolset has been recently published in peer reviewed literature (Shutler et al. 2016, doi:10.1175/JTECH-D-14-00204.1, available also free of charge on ResearchGate). The manuscript has been available to us when we were working on this study but otherwise we worked as end-users. Therefore we reply (below) to questions about FluxEngine according to our best knowledge. However, because we have not yet seen the actual source code (it will be open sourced in near future), we know only as much as stated in the documentation (online in the tool and in the paper). In fact, the decision not to bombard the toolset authors with email questions was one of the additional aims of the study, the first one performed by persons other than its authors. We wanted to check whether the toolset is "user ready". We have to add that it did work and therefore we plan to use FluxEngine in our next projects, including ones involving fluxes that are not yet included (this should be easier after it becomes open sourced).

Comments from referee#1:

Major comments:

(a) Many of the mentioned gas transfer formulations have been developed for different wind products (e.g. NCEP, CCMP), whereas the authors only use one wind product. There are some major differences between wind products. I am not convinced that, If you would consider using a certain transfer formulations in combination with the wind product it was initially calculated for, that you would still get the same difference in the results. I believe this aspect has to be thoroughly discussed.

A: We used only one wind product, an Altimeter Global Monthly Wind Field on a 1°x1° geographical grid from GlobWave L2P because at the time it was the only one available in the FluxEngine toolset (actually one could choose instead ASCAT Global Monthly Wind Field but… it was not yet ready). However we do not think it is a major problem because the point of the study was to constrain uncertainty caused by the choice of the gas transfer velocity parameterization. We did not want to repeat the analysis done within the same ESA project and presented in two submitted papers (Woolf et al 2015a and Woolf at al 2015b) which focused on other sources of uncertainty (the wind field is one of them).

(b) The authors mention the use of both Takahashi and SOCAT climatology. While the Takahashi et al 2009 climatology fills data gaps using an advection based algorithm, the SOCAT climatology to the extent of my knowledge does not use any gap filling methods at all. The authors report a difference between the climatologies of 8% (NA) and 19% (Arctic), whereas it is not clear if this number truly stems from the difference in the climatologies or simply the difference between gap-filled and not gap-filled estimate.

A: In both cases we used the FluxEngine toolset which has its own tools for interpolation. They were used for both the datasets. The details are given in the Appendix to Shutler et al 2016 (available online). This fact has been added in the revised text.

(c) I am struggling a bit to find the importance of this work – i.e. what do you add to our scientific understanding of the topic that has not been known before. It is well known that there are differences in the formulations, but if the intention of this paper is to quantify this difference, then I believe you need to quantify point (a) above. Furthermore, many of the gas transfer formulations are developed using data collected over a somewhat narrow wind range (mainly between 5-12m/s), which explains large differences at the edge or outside the sampled wind range. This aspect also needs to be discussed.

A: The fact that some parameterizations were created using only low winds makes it, in our opinion, even more important to compare the results of their usage overseas with high winds, such as the North Atlantic. We hope the reviewer agrees that it makes it even more surprising (and publishable) that the differences in the net fluxes are smaller in such a basin than globally. We are therefore grateful for the remark.
We added text giving this very motivation as the last sentence of Methods in the revised manuscript and in the Discussion section when mentioning the result.

(d) In the introduction page 2593 lines 22-24, the authors mention that the uncertainty of the flux has been recently discussed in Woolf et al 2015 a and b. But there is no discussion of the results of Woolf in comparison to this study. In general I am missing a proper comparison to prior studies, e.g. Sweeney et al 2007, who found that the gas transfer parametrization leads to a 30% uncertainty in the flux, whereas Landschützer et al 2014 find 37% (also including measurement uncertainty and gridding error), or any other recent study. How do previous studies compare to this one? Does this new estimate fundamentally change our current estimates?

A: This is exactly because the topic of this manuscript is supposed to fill something we believe is a gap in the generally very comprehensive analysis Woolf at al. 2015a and 2015b papers. We are not coauthors of them and although they are available to us within the ESA project, we are not authorized to present the results which have not yet been published (apart from the fact that the papers do not cover the parameterization choice issue and therefore we cannot directly compare the results of the papers with the present study).

As concerns the comparison with previous papers, we will mention the global results suggested by the reviewer in the revised manuscript but it has to be stressed that the point of the paper is a regional study and we show the global results only for comparison. However, we do agree that it will improve the manuscript if we mention that our global uncertainties due to the choice of gas transfer velocity formula are similar to the previously published estimates. However, direct comparison is impossible here because Sweeney et al. 2007 compared two quadratic parameterizations (his and Wanninkhof 1992) we did not use, choosing instead some more recent ones namely, in the case of quadratic formulas, Ho et al. 2006 and Wanninkhof 2014. The difference of flux between formulas with the same wind power is equal to the difference of the constant coefficient (transfer resistance factor) only so there is no need to integrate them with wind fields to know how much the resulting fluxes will differ. The interesting part is to compare parameterizations with different wind speed dependence (which has been the purpose of the manuscript).

Landschützer et al. 2014 unfortunately showed only the combined uncertainty “stemming from ΔpCO_2 and the transfer velocity, using square root of the sum squares propagation [which] yields an uncertainty of ±0.53 Pg C yr^{-1}” (by the way they also use only one wind product!). This result also cannot be directly compared with ours. In fact this shows that we presented something which had not been previously shown: the uncertainly coming solely from the transfer velocity formula choice.

(e) Throughout the manuscript, Flux engine is sometimes spelled “Flux Engine” and sometimes “FluxEngine”. In this review I will spell it the way of its first appearance, i.e. Flux Engine.

A: Corrected – FluxEngine is the right form

Abstract line 2: The authors mention the importance for the anthropogenic budget, but there are some issues with this term. Surface observation based flux estimates, like those calculated in this work do not provide an anthropogenic sink estimate, but a contemporary sink estimate. The anthropogenic sink can only be determined by the pre-industrial state of the ocean, which is estimated to be a source of natural CO2 to the atmosphere due to river input of carbon.

A: Well, the term is established and there are many papers about anthropogenic part of the carbon budget (we mention some of them later on, such as Le Quéré et al. 2105 or Orr et al 2001). However we agree with the reviewer that it may be controversial and we actually do not need the word “anthropogenic” in the abstract (we never differentiate this part of CO2 flux in the paper). Therefore we drop it in the revised manuscript replacing it with “global carbon budget”.

Abstract line 3: “uncertainties in”
Abstract line 4: remove “sink”.
Abstract line 5: “parameterization of THE CO2 gas transfer velocity”
Introduction, page 2593 line 1: There is a spurious “Le” in the reference list before “Landschützer et al 2014”. Presumably this belongs to “Quéré et al 2015
Introduction page 2593 lines 5-8: The word “interdecadal” might be not appropriate here, as Schuster and Watson 2007 report results from the mid 1990s to the early 2000s, i.e. only 1 decade. More appropriate would be interannual or intra-decadal.
Methods page 2596 line 1: “ignore the difference” - please provide a reference
Methods page 2596 line 4: change “taken” to “referred to as”

A: All done

Introduction page 2593 lines 16-19: The authors list a number of potential sources for flux uncertainty, yet later in the manuscript, only one is considered, namely the transfer parametrization. As a reader I would like to know what is the most important of these uncertainties? Is there any literature regarding this topic besides Takahashi 2009?

A: We agree that citing Takahashi in this place was not a good choice (the point was to show the climatology we used, not the literature on uncertainty). We have corrected it now listing Landschützer et al. (2014) and the two submitted Woolf et al. papers (which are discussion of the very topic), deleting the sentence about them at the end of the paragraph.

Methods page 2594: I am not familiar with the Flux Engine software, so a bit more detail would be appreciated (e.g. what reanalysis and model data are included? Are there other wind products available to test? Is it publicly available, and if yes, is there a URL?)

A: FluxEngine is not yet publicly available but should be open sourced by the time the paper is published (the condition was publication of the FluxEngine paper which is already online), possibly within weeks from the moment this response is written. Therefore we added the URL and some additional information

> Line 8 …FluxEngine which is available on the http://www.ifremer.fr/cersat1/exp/oceanflux/
>
> Line 12 …the toolbox that can be use by the scientific community and to aid the…
>
> Line 14 ….gridded flux products with 1ºx1º spatial resolution. The output files contained twelve sets (one set per month) in a NetCDF files. Each data set includes the mean (first order moment), median, standard deviation and the second, third and fourth order moments calculated for each calendar month. There is also information about origin of data inputs as well as results of our calculated. Input data users can chose from all available on the FluxEngine program (perhaps from monthly EO data: rain intensity and event, wind speed and direction, % of ice age and thickness, from monthly model data ECMWF air pressure, whitecapping, from monthly climatology as pCO_2, SST, salinity) and configurable them in a various way. The user needs to choose different components in a calculation process as a way of computed transfer velocity, parametrization to the wind speed calculation, corrections etc.. The FluxEngine has been developed not only to support the study of the air-sea flux of CO_2 but also to aid the study of other gases as DMS and N_2O (Land et al., 2013; Shutler et al., 2016).

Methods page 2594 line 21-22: The authors mention that both SOCAT and Takahashi climatology are calculated for 2010. Takahashi et al 2009 is calculated for a reference year 2000 and to the extend of my knowledge, the SOCAT climatology does not have a reference year. Have they been recalculated, and if yes how?

A: This is correct, as concerns the original papers. However the climatologies were calculated within FluxEngine tool set for the same year (the user has a choice of year). A short explanation has been added to the manuscript text.

Methods page 2595 lines 1-3: I assume the wind speed data are at a height of 10 meters above surface. To the extend of my knowledge, all parametrizations used use the 10 meter above surface wind speed. In general, how has the second and third moment of the wind speed been calculated. There is an interesting discussion in Wanninkhof et al. 2013 where the authors caution that it is essential to use <u2> not <u>2. Hence some information how (if so) the wind product has been averaged.

A: The definition of U10 was already provided below equations (4-8). There was a small language error (now corrected).

As concerns the calculation of wind speed moments, we cannot be sure before FluxEngine is open sourced (we are end-users ourselves even if insider user-ends and we never saw the source code). However, we assume they are actual moments, not powers of the mean value because this is how they are described in Shutler at al. 2016 (we now paraphrased the fragment of the paper to beef up the toolset description in the revised manuscript as described above).

Methods page 2595 lines 8-10: I was wondering what the motivation was to separate North Atlantic and Arctic at 64N? Furthermore, please state how far north the Arctic estimate extents, and how you have dealt with ice covered areas. From Figure 1 it seems like the surface area changes from season to season - this is relevant information for your final flux estimate that is currently missing in the text.

A: The 64N choice was rather arbitrary. The motives were to cover all the areas of the annual Arctic cruise of the IOPAN ship R/V Oceania for later study. All calculation and corrections were made in FluxEngine toolbox within FluxEngine software. The algorithm of which “pixels” to include in every month is based on percentage ice cover for each month (Shutler et al. 2016). However the air-sea flux on sea-ice covered area is zero anyway and therefore we believe this is the correct approach. From the same reasons we believe that plotting the ice masks for each month is not really relevant for the purposes of the paper.

Methods page 2596 lines 25-26: Please explain in more detail what “wind driven and radar backscatter driven” mean.

A: Wind driven and radar backscatter driven are versions of algorithm using either U10 or directly the wave slopes from scatterometers as described in Goddijn-Murphy et al., 2012. We decided not to copy the whole explanation from the paper but just add similar explanation as the one above and to reference the paper.

Results page 2597 and Figure 1 and 2: I do not understand why there are gaps (white areas) in the Takahashi et al based pCO2 and flux estimate in the North Atlantic e.g. in the center of the basin between 40-50N? I could not identify such gaps in the Takahashi et al 2009 publication. Do they result from k and if yes, then why? Please explain.

A: This is another question about FluxEngine which is not easy to answer not being its authors and not having access to the source code. Some of the gaps are obviously caused by the transition from the 5ºx5º grid Takahashi used to the FluxEngine 1ºx1º grid and the ice and shore masks (Rockall Island is a visible example). This is mentioned in Shutler et al. 2016. However we do not know the reasons for every missing pixel in every month. We added explanation in the captions of Figures 1-4.

Discussion page 2599 lines 9-10: please quantify what "within the experimental uncertainty" means.

A: This is actually what the authors of the three parameterizations (Ho, Nightingale and Wanninkhof) said during the Kiel SOLAS session on the very subject (described in the next sentences). We wrote the paragraph just after the session so it is as close to actual quote from the authors as possible. The meaning is experimental data we have in hand cannot distinguish between the three. The report from the session (available online http://goo.gl/TrMQkg) supports our memory stating that:

> For gas transfer of CO2 over the oceans the relationships proposed in Nightingale et al. (2000), Sweeney et al. (2007), Ho et al. (2006), and Wanninkhof et al. (2009) are recommended. They are very similar and fall within the overall uncertainty of DT measurements.
> The relationships by Liss and Merlivat (1986) Wanninkhof 1992 and McGill is et al. (2001) do not agree with current constraints.

We did not use Sweeney et al 2007 (as mentioned above) therefore we mention only the other three. We also cite a newer Wanninkhof (2014) paper but the formula it uses can be really found to the Wanninkhof et al 2009. However it is hidden among many other formulas so we believe the 2014 citations is clearer to the reader.

The "our" in the manuscript sentence was supposed to refer to the scientific community, not the manuscript authors. This has been rewritten to make it clearer and a citation of the session report by Nightingale (2015) has been added.

Discussion page 2601 line 4: "Takahashi and SOCAT pCO2 climatologies". SOCAT reports fCO2 How has this been converted to pCO2? Via the Flux Engine software?

A: Yes, the data from SOCAT website were pre-processed into the format required by the FluxEngine software. Actually it works the other way: FluxEngine recalculates pCO2 to fugacities (we mention its use of fugacity in the new paragraph on FluxEngine). A sentence about this pCO2 → fugacity recalculation has also been added to the manuscript.

Figure comments:

Figure 1,2,3 and 4: I am wondering where the data gaps come from? Also, please increase the font of the plot, as numbers, e.g. from the colorbar are difficult to read.

A: We added information about gaps to the describe under figures (see also above) as well as change the scale (put big one for all print in one figure) for better view. Unfortunately we cannot change the font of the numbers (software problem). We hope that in the print paper the scale and maps will be bigger that in OS Discussion paper.

Figure 5,6,7 and 8: There is contradicting information in the caption compared to the y-axis or the figure title. In the caption, the authors report units of g/m2/day whereas on the y-axis/title they report Tg. Which one is correct? In case Tg is correct, should it not be Tg/yr?

A: Thanks for spotting this. We have corrected this in the captions (to Tg/year)

Figure 6: please increase the font to make it better readable.

See above

Figure 6 and 7: It is remarkable that without a few exceptions, the majority of the parametrizations are within the standard deviation of all parametrizations. Using the standard deviation as an uncertainty criterion, this would suggest that you based your statement page 2599 lines 9-10 on this figure. Is that correct? If so, please state this more explicitly.

A: This is the standard deviation (SD) of the values of fluxes calculated with different parameterizations, a simple value of the spread of the results (variability). If the results obeyed normal distribution, 2/3 of them would be, by definition, within one standard deviation from the average. They obviously are not (the sample is small) but still, the fact that the majority of results are within one SD from the average results directly from the definition of standard deviation.

We show the SD value exactly as a measure of variability of all the results. We do not place to much stress on the value as it is calculated from both the parameterizations believed (see the discussion of the Kiel 2015 SOLAS session on the subject) to be close to the best experimental results, and formulas which are not (but still are found in the literature and sometimes used). In the discussion we tried to differentiate the two. The source of the p. 2999, l. 9-10 statement was the very Discussion Session. We added information on this (see above), including the link to its minutes (Report by Phil Nightingale) where the statement is given explicitly as one of the session recommendations.

Figure 8: Again, it is important to understand for the reader how the SOCAT climatology has been created. If it is a climatology from the cruise tracks only as provided on the SOCAT website, than it is not directly comparable to the Takahashi climatology. If it is a climatology created by a gap filling method, then please explicitly explain how it has been created. Otherwise figure 8 is more misleading than helpful.

A: We added a statement that the SOCAT data were interpolated using the FluxEngine toolset (actually in two places: in the Methods and Results).

We presented the results at two conferences (EGU and SOLAS) and this difference between the Takahashi and SOCAT results, especially in the Arctic where they have inverse seasonal variability, was commented by many experts in the field as one of the most interesting results. We have to add that this result was also shown as a short presentation at a special SOCAT/SOCOM workshop ("Surface Ocean CO2 Atlas & Surface Ocean pCO2 Mapping

Intercomparison”) one day before the Kiel conference, and the discussion showed it was deemed an interesting and important result. This is exactly why we felt obliged to be including it in the manuscript.

The question raised by the figure is which data set (Takahashi vs. SOCAT) is right. As much as we believe SOCAT (as the more complete one) is more accurate, as the one using more Arctic data, we have no way of concluding this from just comparing the resulting fluxes. Only additional experimental data can settle this, and we stated as much in the manuscript (in the last sentence of the conclusions).